POLITECHNIKA WROCŁAWSKA

WYDZIAŁ MECHANICZNY

KIERUNEK: Automatyka i Robotyka

SPECJALNOŚĆ: Systemy Produkcyjne

PRACA DYPLOMOWA

MAGISTERSKA

TEMAT\_PO\_POLSKU

TEMAT\_PO\_ANG

AUTOR:

Kamil Kozdrowiecki

PROMOTOR:

Dr inż. Krzysztof Chrapek

W10/I24

OCENA PRACY:

WROCŁAW 2016

Spis treści

[1. Cel i zakres pracy 3](#_Toc451285699)

[2. Założenia projektowe 3](#_Toc451285700)

[3. Teoria stereowizji 3](#_Toc451285701)

[3.1. Obrazowanie punktu z przestrzeni trójwymiarowej 3](#_Toc451285702)

[3.2. Triangulacja w stereowizji 5](#_Toc451285703)

[4. Dobór osprzętu 6](#_Toc451285704)

[4.1. Jednostka obliczeniowa 6](#_Toc451285705)

[4.2. Kamery 6](#_Toc451285706)

[4.3. Biblioteki programowe 7](#_Toc451285707)

[5. Komunikacja jednostka obliczeniowa - robot 7](#_Toc451285708)

[5.1. Stawiane wymagania 7](#_Toc451285709)

[5.2. TCP 7](#_Toc451285710)

[6. Oprogramowanie jednostki obliczeniowej 8](#_Toc451285711)

[6.1. Struktura 8](#_Toc451285712)

[6.2. Klasa CStereoCalib 11](#_Toc451285713)

[6.3. Klasa CStereoVision 11](#_Toc451285714)

[6.4. Klasa CTCPConnection 11](#_Toc451285715)

[7. Oprogramowanie robota 11](#_Toc451285716)

[7.1. Struktura 11](#_Toc451285717)

[8. Testowanie oprogramowania 11](#_Toc451285718)

[8.1. Kalibracja kamer 11](#_Toc451285719)

[8.2. Filtrowanie 12](#_Toc451285720)

[8.3. Analiza dokładności 16](#_Toc451285721)

[9. Podsumowanie 16](#_Toc451285722)

[10. Bibliografia 16](#_Toc451285723)

[11. Spis ilustracji 16](#_Toc451285724)

[12. Spis tabel 16](#_Toc451285725)

[13. Spis załączników 16](#_Toc451285726)

# Cel i zakres pracy

Coraz większego znaczenia w przemyśle nabierają układy, które w łatwy sposób można zaadaptować do zmian parametrów procesowych wymuszonych przez modyfikacje wytwarzanego przedmiotu, czy też usprawnień w technologii procesu, montażu. Większość takich elastycznych maszyn wymaga od operatora wiedzy i doświadczenia, nabytych przez kosztowne kursy i szkolenia, aby dostosować nastawy lub programy urządzeń nawet do pomniejszych zmian w procesie. Biorąc pod uwagę powyższe tendencje rozwoju technologii, za cel niniejszej pracy magisterskiej postawiono opracowanie intuicyjnego systemu programowania punktów i prostych funkcji dla robotów przemysłowych. Wprowadzenie takiego rozwiązania przyspieszy procedurę przystosowania elastycznego systemu wytwórczego do nowych produktów, szczególnie w przypadku małoseryjnych produkcji.

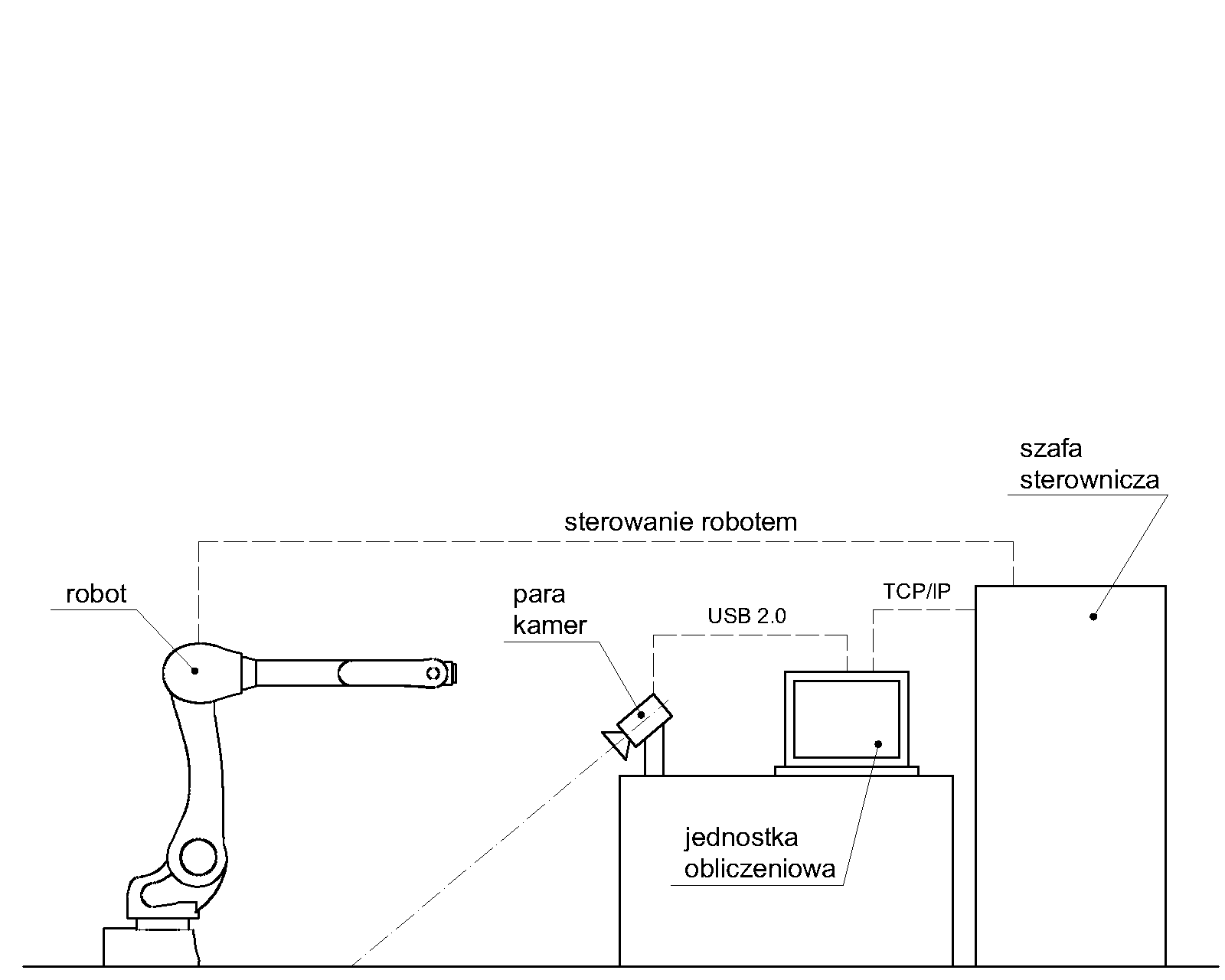
# Założenia projektowe

Do rozpoznania i analizy położenia zadanego punktu w przestrzeni wykorzystano dwie jednakowe kamery umieszczone na jednej linii. Wykorzystując technikę stereowizji opisaną w rozdziale 3 możliwe jest wyznaczenie głębi punktu z dwóch obrazów, a co za tym idzie, określenie przestrzennej pozycji danego piksela. Korzyści wynikające z zastosowania tej metody to m. in.:

* Względnie ekonomiczne rozwiązanie w porównaniu do innych technik, np. skaner laserowy.
* Niewielki rozmiar urządzenia
* Możliwość dostosowania wykrywanych punktów do urządzenia wskazującego przez duży zakres metod filtrowania
* Możliwość wykorzystania pobranych obrazów do innych operacji np. identyfikacja obiektu, wyznaczenie rozmiaru obiektu

Przyjęto, że urządzeniem wskazującym zadane położenia będzie wskaźnik laserowy o czerwonym świetle. Decyzja ta została podjęta na podstawie kilku czynników:

* Ułatwienie procedury filtrowania i wydobycia odpowiednich pikseli, ze względu na dużą jasność punktu, charakterystyczny względem tła kolor
* Możliwość wskazania punktu z daleka, nie będąc bezpośrednio w strefie roboczej robota
* Łatwo dostępne, tanie, kieszonkowe urządzenie wskazujące, nie wymagające dużej mocy zasilania

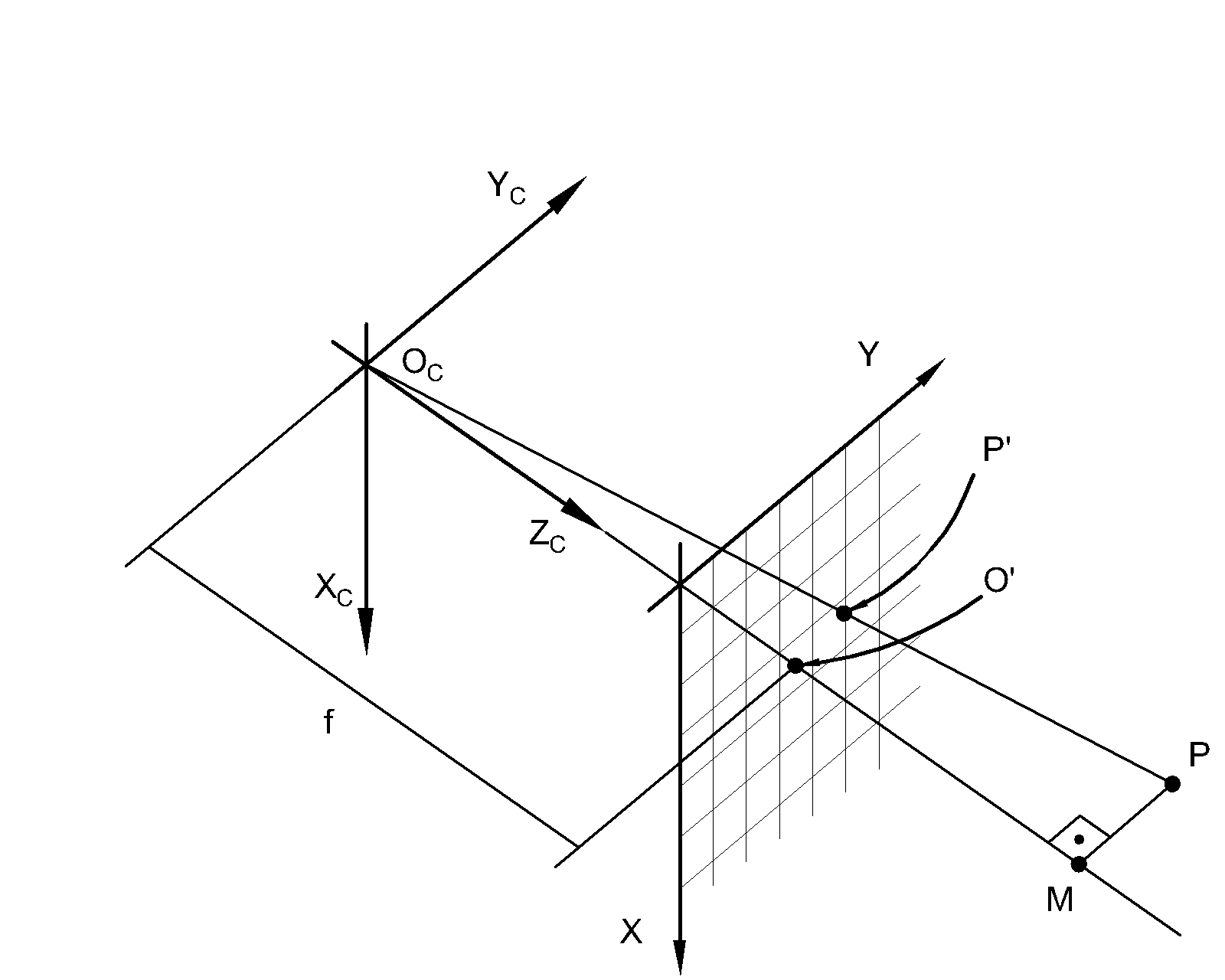


# Teoria stereowizji

Już jako niemowlę człowiek przyswaja sobie praktyczną umiejętność wykorzystania stereowizji do funkcjonowania w naturalnym środowisku. W biologicznym aspekcie znana jest jako „ widzenie binokularne”. Technika, czerpiąc pomysły garściami z przyrody, wykorzystała również i tę metodę określania głębi obrazu.

## Obrazowanie punktu z przestrzeni trójwymiarowej

Układ współrzędnych kamery wyznaczają osie , , oraz punkt centralny . O odległość ogniskową od płaszczyzny przesunięta jest płaszczyzna obrazowa z lokalnym układem współrzędnych i siatką pikselową, której rozmiar elementu podstawowego wynosi . Punkt na płaszczyźnie obrazowej jest rzutem punktu wzdłuż osi . Prosta wyznaczona przez punkt i jest osią optyczną. Zobrazowanie punktu położonego w przestrzeni trójwymiarowej polega na perspektywicznym zrzutowaniu tego punktu wzdłuż prostej przechodzącej przez punkt centralny układu kamery i punkt na płaszczyznę obrazową, wyznaczając w ten sposób punkt . Wykorzystując podobieństwo trójkątów i można utworzyć następujące zależności [1]:



Biorąc pod uwagę fakt, że rzeczywista kamera posiada niedokładności w postaci błędów wykonania elementów oraz ich montażu, do równania NUMER\_RÓWN wprowadza się parametry korekcyjne.

Przekształcając powyższe równanie do postaci macierzowej otrzyma się:

Macierz określana jest jako macierz kamery i jest dla niej niezmienna. Wartości określają przesunięcie środka płaszczyzny obrazowania względem osi optycznej, wynikające z nieidealnego montażu matrycy światłoczułej. Kształt matrycy wymusza wprowadzenie osobnych ogniskowych dla odpowiednich kierunków.

Inną nieprawidłowością powodującą radialne zniekształcenie obrazu nasilające się wraz z odległością od osi optycznej jest dystorsja sferyczna [OBRAZEK?] . Aby zredukować jej wpływ wprowadza się kolejną korektę w postaci [2]:

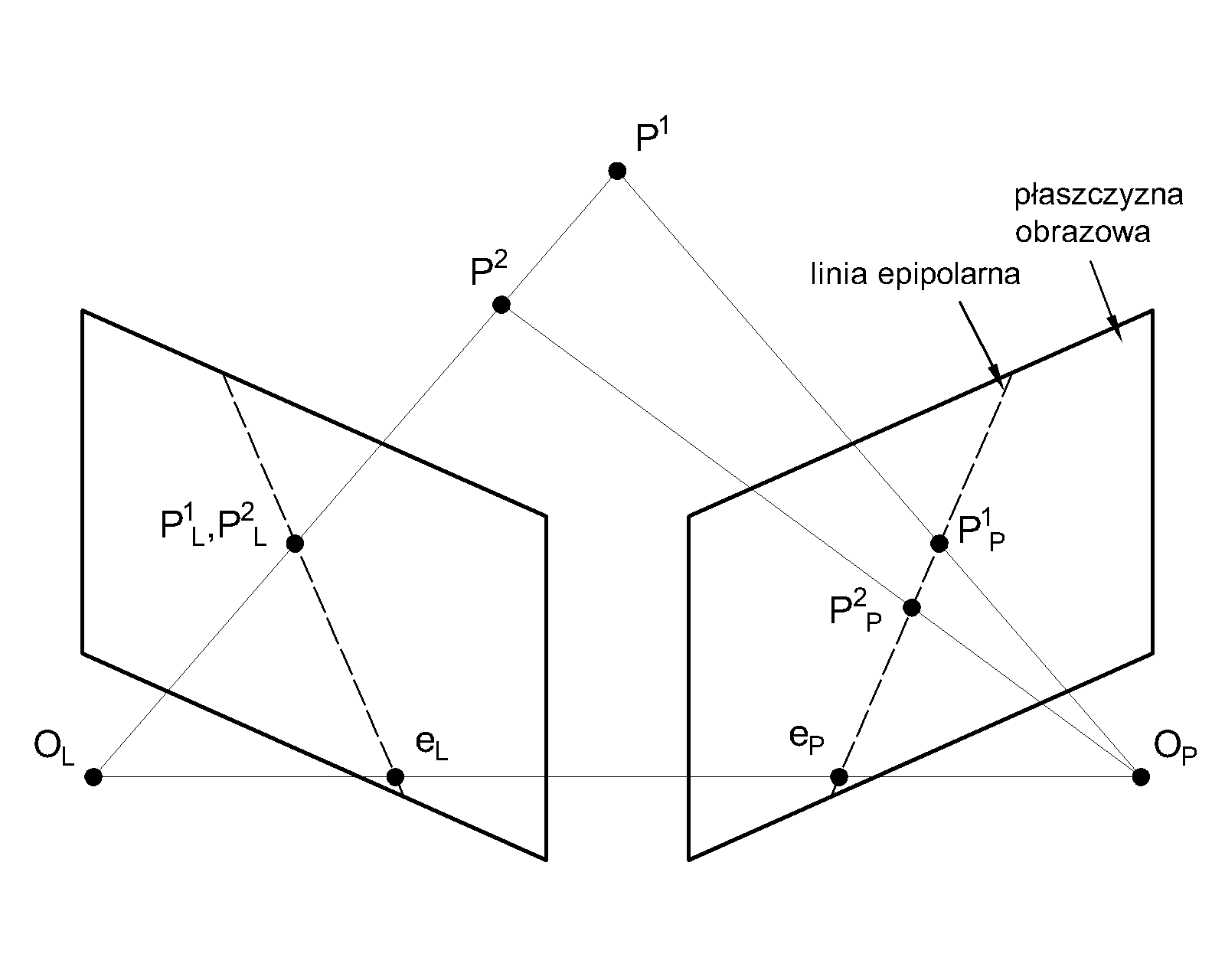
,gdzie:

Dystorsja tangencjalna jest kolejnym efektem deformującym obraz, choć w mniejszym stopniu niż dystorsja sferyczna. Wprowadzenie poniższego równania   
minimalizuje skutki zniekształcenia [2]:

## Geometria epipolarna

Podstawową koncepcją związaną ze stereowizją jest geometria epipolarna. Na rys. XX zaznaczono dwa punkty położone w przestrzeni trójwymiarowej. Rzuty tych punktów na płaszczyzny obrazowań kamery lewej i prawej odpowiadają kolejno punktom . Rzut prostej przechodzącej przez punkt centralny danej kamery i punkt w przestrzeni na płaszczyznę obrazową drugiej kamery nazywamy linią epipolarną. Punkty epipolarne są punktami przebicia płaszczyzn obrazowych kamer przez odcinek łączący punkty centralne kamer. W przypadku kanonicznego układu kamer, gdy płaszczyzny obrazowań obu kamer leżą w tej samej płaszczyźnie, punkty epipolarne znajdują się w nieskończoności.

Używając wyłącznie lewej kamery w sytuacji przedstawionej na rys. XX wiadome jest, że punkty leżą na tej samej prostej przechodzącej przez środek projekcji , jednak określenie odległości każdego punktu od płaszczyzny obrazowej jest niemożliwe.. Aby mieć podstawę do odróżnienia obu punktów i wyznaczenia ich pozycji wymagany jest punkt odniesienia. Odnajdując rzut punktu płaszczyźnie obrazowej prawej kamery można wyznaczyć linię epipolarną, na której musi leżeć również rzut punktu , na bazie którego możemy ocenić różnice w położeniach obu punktów w przestrzeni.



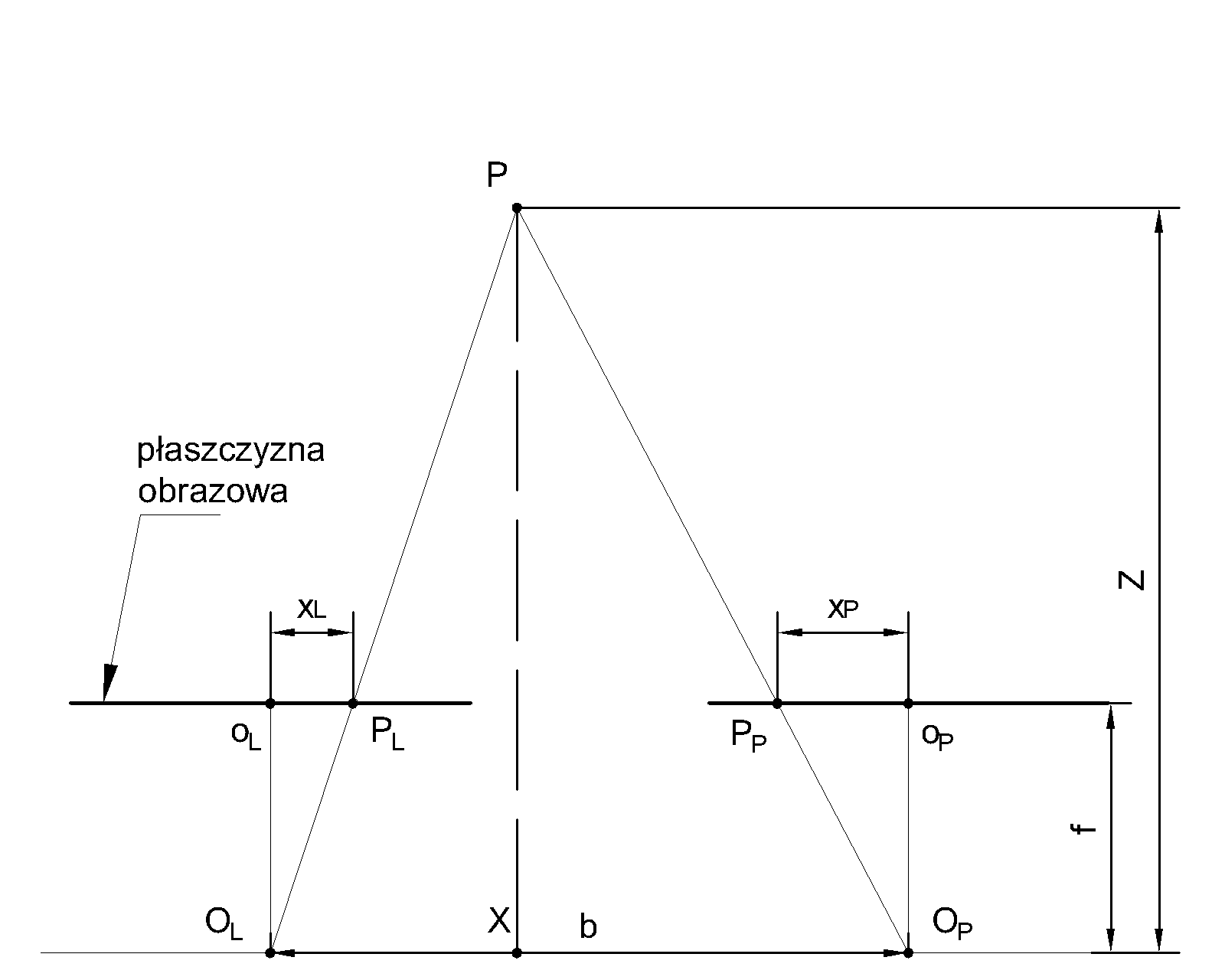
Na podstawie powyższego przykładu można zauważyć, że każdy punkt obrazowy punktu P z przestrzeni trójwymiarowej znajduje się na płaszczyźnie obrazowej wyłącznie na odpowiednich liniach epipolarnych[1]. Jest to jedno z ważniejszych założeń geometrii epipolarnej, dzięki któremu można zredukować przeszukiwanie odpowiadających punktów na obu obrazach z dwóch wymiarów do jednego, znacząco zmniejszając potrzebny czas na obliczenia.

## Triangulacja w stereowizji

Analiza głębi w stereowizji opiera się na wyznaczeniu różnicy w horyzontalnym położeniu danego punktu na obrazach z obu kamer. Wyznaczona głębia danego punktu jest odwrotnie proporcjonalna do uzyskanej różnicy w położeniu horyzontalnym zrzutowanego punktu na płaszczyzny obrazowe. Zależność ta wynika z następującego rozumowania:

Na rys. XX przedstawiono schematycznie projekcję punktu P na płaszczyzny obrazowe dwóch kamer w układzie kanonicznym, który zakłada, że:

* Płaszczyzny obrazowania kamer są do siebie równoległe
* Osie optyczne kamer są do siebie równoległe
* Odległości ogniskowych kamer są sobie równe



Różnica w położeniu zrzutowanego punktu P na płaszczyznach obrazowych kamer wynosi [1]:

Ze względu na podobieństwo par trójkątów PBA i AED oraz PBC i CFG:

Po przekształceniu otrzymuje się:

gdzie:

– różnica w horyzontalnym położeniu od środka matryc zobrazowanego punktu P

– odległość ogniskowa kamer

– odległość między kamerami

– odległość punktu od kamer

Powyższe równanie jest prawdziwe dla kamer o idealnie równoległych płaszczyznach obrazowania. Oczywiście w praktyce bardzo trudno to osiągnąć, dlatego bierze się pod uwagę translację i rotację kamer względem siebie.

# Dobór osprzętu

## Jednostka obliczeniowa

Wstępny projekt uwzględniał wykorzystanie popularnych w ostatnich latach komputerów jednopłytkowych takich jak: Raspberry Pi, Banana Pi, czy BeagleBone. Pomimo niewielkich rozmiarów, posiadają niezbędne do projektu interfejsy 2xUSB i Ethernet. Przegląd podobnych rozwiązań [wskaż które], w których również wykorzystano takie jednostki wykazał, że cechują się one za niską mocą obliczeniową w stosunku do złożoności algorytmów przetwarzania obrazów. Z tego powodu porzucono te założenia na korzyść komputera PC. Ostatecznie wszystkie testy zostały przeprowadzane na laptopie HP Pavilion dv6, którego główną specyfikacje przedstawiono w tab. 4.1.

|  |  |
| --- | --- |
| System operacyjny | Windows 7 Home x64 |
| Procesor | 8 x Intel Core i7 2.20 GHz |
| Pamięć | 8 GB |
| Karta graficzna | AMD Radeon HD 6700M |

## Kamery

Dobór kamer był kluczowym elementem przy wstępnych założeniach. Głównymi parametrami, które miały wpływ na wybór urządzenia były:

* Rozmiar matrycy
* Maksymalna rozdzielczość
* Liczba klatek na sekundę
* Możliwość konfiguracji parametrów ( ekspozycja, wzmocnienie)
* Typ przewodu

Rozdzielczość obrazu wpływa na dokładność wyznaczanej głębokości punktu ale jednocześnie zwiększa wymaganą moc obliczeniową do przetworzenia większej ilości danych. Z tych względów przyjęto, że kamera będzie pobierać obraz w rozdzielczości 640x480. Istotną i często pomijaną cechą jest możliwość kontrolowania ustawień kamery. Zazwyczaj producenci kamer internetowych implementują do swoich produktów automatyczne dopasowanie wzmocnienia lub ekspozycji, których użytkownik nie może kontrolować i przy pewnym oświetleniu często skutkuje to pojawianiem się szumów. Możliwość sterowania tymi parametrami daje kolejne stopnie swobody np. przy filtrowaniu obrazu. Liczba klatek na sekundę nie jest już tak istotną cechą, ze względu na długi okres przetwarzania pojedynczej klatki, przez co trudno jest wykorzystać całkowity potencjał szybkości pobierania obrazu przez kamerę.

Podczas projektowania i testowania systemu korzystano z kamery *Logitech C905*. Charakteryzuje się możliwą maksymalną rozdzielczością większą od zakładanej, przez co nie wykorzysta się w pełni jej potencjału mając na uwadze ograniczoną moc obliczeniową. Układ optyczny wykonała firma *Carl Zeiss* znana z wytwarzania wielorakiego sprzętu optycznego. Dodatkowo producent udostępnia oprogramowanie umożliwiające ręczny dobór parametrów takich jak: ekspozycja, wzmocnienie, nasycenie Domyślnie kamera pracuje na ustawieniach automatycznych, które zawsze można przywrócić po zmianach.

.

|  |  |
| --- | --- |
| Ilość pikseli | 2 Mpix |
| Maks. rozdzielczość | 1600 x 1200 |
| Liczba klatek na sekundę | 30 |
| Typ przewodu | USB 2.0 |

## Biblioteki programowe

Projektowanie i pisanie programów analizujących obrazy na jednostkę centralną opierało się o bibliotekę funkcji OpenCV. Jest to jedna z najpopularniejszych bibliotek programowych służąca do szeroko pojętego przetwarzania obrazów: filtrowania, transformacji, kalibracji, analizy ruchu, wykrywania i śledzenia obiektów. Wiele przykładowych programów, kilka napisanych poradników[5][6], jak i dobrze rozwinięta dokumentacja[7] poszczególnych modułów i funkcji ułatwiają proces zapoznawania się z oprogramowaniem i jego możliwościami. Dodatkowo OpenCV wydane jest na bardzo liberalnej licencji BSD, według której można modyfikować i korzystać z biblioteki zarówno w zakresie edukacyjnym, jak i komercyjnym zezwalając na rozprowadzanie napisanych programów w postaci zbudowanej i kodu źródłowego, a nawet włączać je do zamkniętego oprogramowania[3].

Główne funkcje biblioteki, wykorzystywane w kalibracji i wyznaczaniu trójwymiarowej reprezentacji punktów zawarte się w module *calib3d*. W czasie projektowania i pisania niniejszej pracy magisterskiej korzystano z wersji OpenCV 3.1.

# Komunikacja jednostka obliczeniowa - robot

## Stawiane wymagania

Wymogi odnoszące się do komunikacji jednostki obliczeniowej z robotem uwzględniają następujące cechy:

* Dostępność portów, złączy – takie same porty lub złącza w sterowniku robota i jednostce obliczeniowej eliminują potrzebę stosowania konwerterów
* Rodzaje obsługiwanych protokołów – sterownik robota nie jest tak elastyczny jak jednostka obliczeniowa i ogranicza zbiór możliwych protokołów komunikacyjnych
* Funkcje programowe – łatwość i przejrzystość funkcji inicjujących komunikację, odbierania i wysyłania danych
* Obsługa błędów – protokół z zaimplementowaną kontrolą danych i ew. retransmisją pozwala uniknąć dodatkowej obsługi błędów na poziomie aplikacji
* Szybkość transmisji danych – mniej istotna cecha, ze względu na okresową, niewielką ilość przesyłanych informacji

Wykorzystywany sterownik robota E4X w wersji podstawowej posiada trzy interfejsy: USB, Ethernet ( 100BASE-TX) i RS232C, z czego USB przeznaczone jest do zapisywania i odczytywania danych z nośników zewnętrznych. RS232C jest szeregowym standardem komunikacyjnym, cechującym się prędkością transmisji danych do 20 kb/s na maksymalną odległość 15 m. Pomimo wystarczających parametrów interfejs ten jest zazwyczaj przeznaczony do programowania robota przez terminal komputera PC. Ostatnim możliwym wyborem, co nie znaczy że najgorszym, jest standard Etherner. Technologia 100BASE-TX wykorzystuje jako medium nieekranowaną skrętkę miedzianą zakończoną złączami 8P8C, dla których odpowiadające porty znajdują się obecnie w każdym komputerze. Możliwa prędkość przesyłania danych wynosi 100 Mb/s, przy czym na pewno nie zostanie wykorzystany w pełni ten potencjał.

## Protokół TCP/IP

Uwzględniając stawiane wymagania co do sposobu komunikacji jednostki obliczeniowej z robotem zdecydowano się na wybór standardu Ethernet 100BASE-TX. Jest to pierwsza, fizyczna warstwa modelu sieci OSI będąca częścią całej techniki sieciowej Ethernet i określająca medium oraz interfejsy sieciowe. W przypadku trzeciej warstwy, sieciowej, zarządzającej połączeniem na poziomie jego ustanawiania, utrzymywania i rozłączania oraz dobierającej optymalną trasę dla danych, nie ma większego wyboru protokołu poza najpopularniejszym IPv4 identyfikującym adresatów po adresie IP składającym się z czterech, ośmiobitowych liczb oddzielanych kropkami. Do kolejnej, wyższej warstwy transportowej zalicza się protokoły TCP lub UDP. Wysyłane paczki danych przekazywane są przez pośrednie węzły sieci, często o różnych przepustowościach, opóźnieniach, czy topologiach. Te różnice mogą wpływać na opóźnienie w dostarczeniu danych, jak również powodować złą kolejność otrzymywanych pakietów danych, bądź ich utratę. Protokół TCP został tak opracowany, aby zapewnić dostarczenie danych w niezmienionej formie do adresata. Korzysta przy tym z takich procedur jak: kontrola przepływu, potwierdzenie odbioru i retransmisja. Gwarancja odebrania identycznej informacji jak przy nadawaniu może wiązać się z kosztem zwiększonego całkowitego czasu przesyłania informacji. Drugi protokół warstwy transportowej skupia się na dzieleniu całej wiadomości na mniejsze, podstawowe fragmenty określane jako datagramy i strumieniowym wysyłaniu do adresata. Ponadto sam adresat nie musi wcześniej zaakceptować połączenia, tak jak w przypadku TCP, by odbierać informacje. Umożliwia to nadawanie wiadomości do wielu odbiorców równocześnie, tzw. broadcasting. Istotną cechą protokołu UDP odróżniającą go od TCP jest brak mechanizmów potwierdzenia i retransmisji datagramów w przypadku wadliwego odbioru, co oznacza, że jeśli chcemy mieć pewność odbioru oryginalnej, niezmienionej informacji należy w wyższej warstwie zaimplementować algorytm kontrolny odbieranych datagramów. Protokół UDP znajduje zastosowanie tam, gdzie wymagana jest duża szybkość transmisji danych np. wideokonferencje, gry sieciowe, strumieniowanie dźwięku[9].

W wykonywanym systemie, informacje o przypuszczalnej, niewielkiej wadze kilkudziesięciu bajtów będą wysyłane okresowo co ~1 s, co jest w stanie zapewnić praktycznie każda sieć, więc protokoły o szybszej transmisji danych nie są wymagane. Z tego względu jako warstwę transportową przyjęto protokół TCP, dzięki któremu można uniknąć implementowania dodatkowych procedur kontrolnych, aby mieć pewność dostarczenia prawidłowych informacji.

# Oprogramowanie jednostki obliczeniowej

## Struktura

W jednostce obliczeniowej zaimplementowano dwa programy:

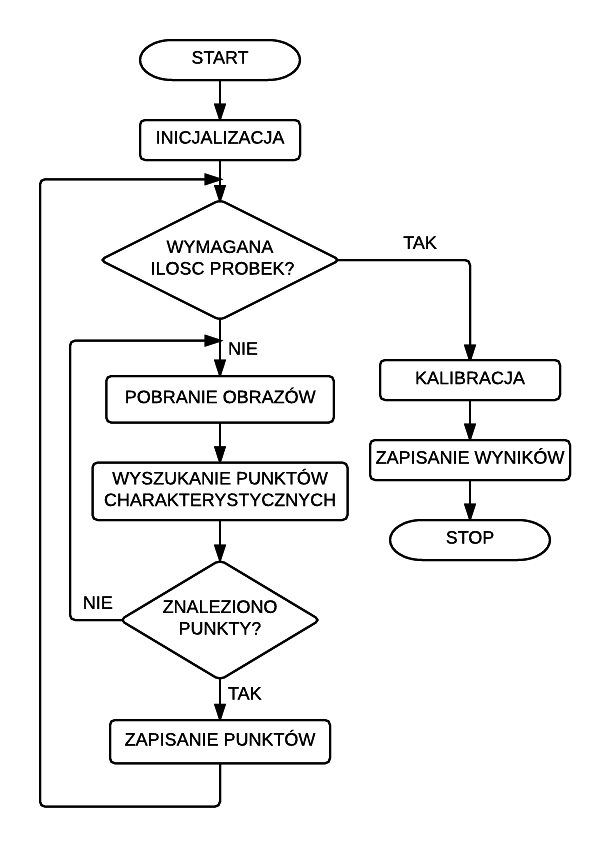
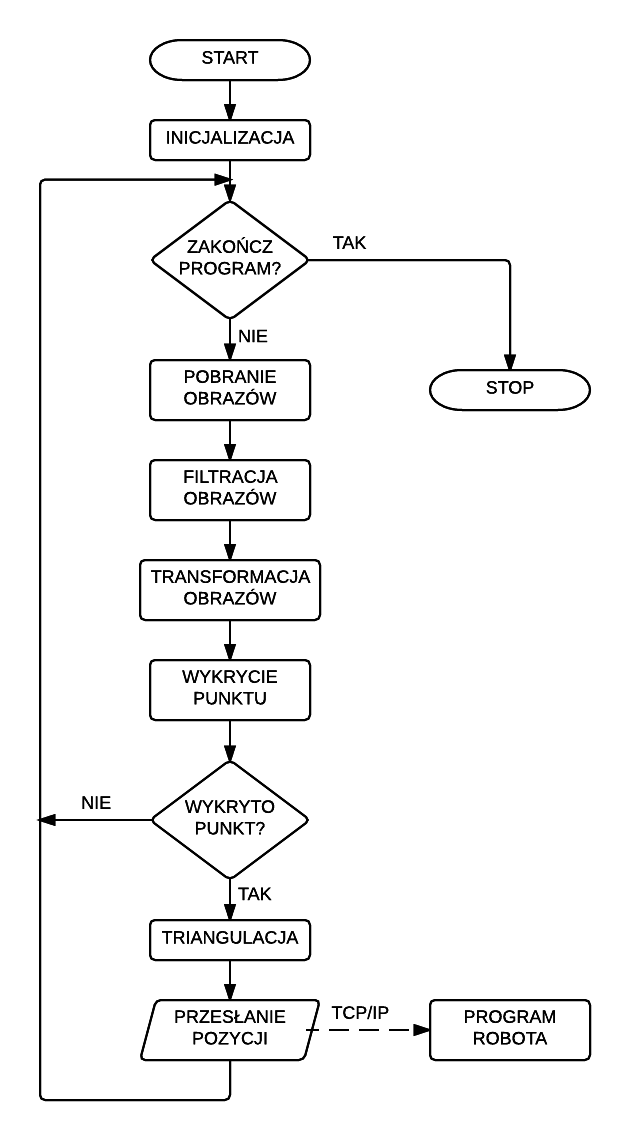
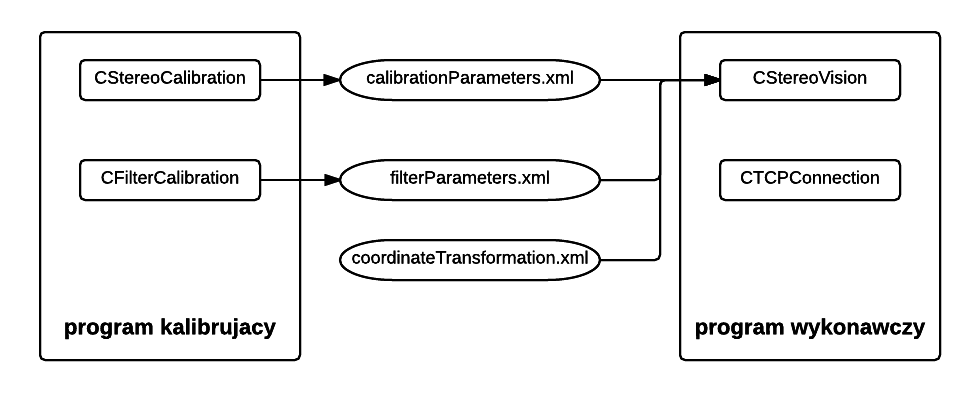
* Kalibrujący - z dwiema opcjami:
  + Kalibracja kamer – przeprowadzenie procesu kalibracji z wykorzystaniem szachownicy w celu wyznaczenia i zapisania do zewnętrznego pliku współczynników i macierzy niezbędnych do przeprowadzenia poprawnej analizy obrazów w programie wykonawczym
  + Ustawienie parametrów filtrowania – empiryczne wyznaczenie progów filtrowania obrazów w jednej z dwóch metod: RGB lub HSV
* Wykonawczy – główny program realizujący ogólną sekwencję: pobranie obrazów, analiza obrazów, wyznaczenie położenia punktu oraz przesłania informacji o punkcie do robota

Druga opcja kalibracji – ustawienie parametrów filtrowania – w kroku inicjalizacyjnym pyta użytkownika o numer ID kamery oraz przestrzeń kolorów, według której ma być realizowanie progowanie ( RGB lub HSV).

Zasadę działania głównego programu wykonawczego, analizującego położenie punktu wskaźnika opisuje schemat blokowy na rys. NUMER. Podobnie jak w przypadku kalibracji, pierwszym krokiem jest inicjalizacja, podczas której sprawdzany jest dostęp do pliku wynikowego programu kalibrującego, odczytywane są dane i weryfikuje się połączenie z kamerami. Następnie pobierana jest para obrazów, która poddawana jest szeregowi operacji:

* Filtracji - wyodrębnienie punktu świetlnego wskaźnika od tła
* Transformacji – kompensacja dystorsji obrazów, rektyfikacja
* Wykrycia punktu – uzyskanie położenia punktu wskaźnika z przekształconych wcześniej obrazów

Jeśli wykryto punkty na obu obrazach to na podstawie ich położeń rekonstruuje się przestrzenne położenie punktu wskaźnika. Uzyskana pozycja przesyłana jest do robota, który podejmuje odpowiednio zaprogramowaną akcję.



## Kalibracja kamer

Przed przystąpieniem do kalibracji należy zaopatrzyć się we wzorzec, w postaci czarno – białej szachownicy o znanej ilości pól wszerz i wzdłuż oraz długości boku pojedynczego pola. Należy pamiętać, aby wzorzec znajdował się na płaskiej powierzchni. Wszelkie zafalowania, wybrzuszenia powierzchni mogą skutkować błędnie wyznaczonymi położeniami punktów specyficznych, a co za tym idzie dużym błędem kalibracji.

Ogólna procedura kalibracji kamer przedstawiona została na rys. NUMER. Po uruchomieniu programu kalibrującego prowadzony jest w konsoli dialog z użytkownikiem, w celu wprowadzenia niezbędnych informacji do wykonania kalibracji:

* Szerokość i wysokość szachownicy – wartości liczone w ilościach pół szachownicy używanej do znajdowania specyficznych punktów
* Długość boku pola szachownicy – długość w mm pojedynczego pola szachownicy
* Liczba wymaganych próbek – określa minimalną ilość obrazów z wykrytymi punktami charakterystycznymi która jest wymagana przy procedurze kalibracji. Większa liczba próbek podwyższa jakość kalibracji, jednak znacząco wydłuża czas jej wykonywania. Przeważnie wystarczająca wartość to 15.
* Numery identyfikacyjne obu kamer – zazwyczaj są to liczby 1 i 2. Jeśli nie można otworzyć którejkolwiek kamery, program wyświetli stosowną informację i zapyta ponownie o numer ID.

Jeśli wszystkie wartości zostaną poprawnie wprowadzone, pojawiają się dwa okna wyświetlające obrazy z obu kamer. W tym momencie należy tak manipulować wzorcem, aby znalazł się on w kadrach obu kamer. Poprawne wykrycie szachownicy sygnalizowane jest zaznaczeniem specyficznych punktów w postaci kolorowych okręgów na obrazach kamer. Jeżeli wykryto punkty na obu obrazach, zapisywane są do bufora położenia tych punktów i oznaczane jako kolejna próbka. Przed dodaniem kolejnej próbki występuje dwusekundowa przerwa pozwalająca uniknąć zduplikowania próbki i daje czas użytkownikowi na ustawienie wzorca pod inną orientacją lub położeniem.

Po tym, jak zdobyta będzie wymagana ilość próbek, określona przy inicjalizacji programu, następuje proces analizy otrzymanych danych i obliczenie na tej podstawie parametrów kamer. Procedura trwa kilka minut i trwa tym dłużej im więcej jest próbek. Zakończenie pracy oznajmiane jest stosownym komunikatem w konsoli informującym również o czasie obliczeń i orientacyjnej wartości błędu kalibracji. Na zakończenie użytkownik ma możliwość zapisu parametrów do pliku, jeśli jest usatysfakcjonowany małym błędem, lub odrzucenia wyników kalibracji w przeciwnym wypadku.

|  |  |
| --- | --- |
|  |  |

## Ustawienie parametrów filtrowania

Uruchamiając program kalibracyjny można wybrać drugą funkcję – ustawienie parametrów filtrowania. Po wybraniu odpowiedniej opcji, program pyta o kolejne wymagane informacje:

* Numer identyfikacyjny kamery – wymagana jest tylko jedna kamera, która zazwyczaj posiada ID o wartości 1. Jeśli nie można uruchomić kamery, program ponownie pyta o ID.
* Metoda filtrowania –określa przestrzeń kolorów, dla której w kolejnym etapie określane są granice progowania obrazu wejściowego. Podaje się liczbę przypisaną do danej przestrzeni: RGB lub HSV.

Wprowadzenie poprawnych danych wejściowych pozwala przejść do etapu ustalania wartości progowania obrazu. Pojawiają się dwa okna:

* *cam* – wyświetla w czasie rzeczywistym niezmieniony obraz pobrany z kamery
* *filtered* – wyświetla obraz po filtrowaniu; piksele mogą przyjmować jedną z dwóch wartości 0 lub 255

W oknie *filtered* znajdują się suwaki służące do zmiany wartości progów. Obok każdego z nich wyświetlana jest nazwa progu w zależności od wybranej przestrzeni barw wraz z aktualną wartością. Zmiany wartości przez przesunięcie suwaków są na bieżąco realizowane i uwidocznione w oknie wynikowym filtrowania. Rys. X przedstawia okno *cam* z obrazem wejściowym pobranym przez kamerę, natomiast okno *filtered* na rys. Y uwidacznia efekt filtrowania obrazu wejściowego w przestrzeni HSV na podstawie parametrów dobranych w taki sposób, aby uwidocznić czerwone obiekty.

Jeżeli użytkownik jest usatysfakcjonowany z wybranych parametrów, wciśnięcie dowolnego klawisza przy aktywnym jednym z obu okien powoduje zatrzymanie aktualizacji obrazów i wyświetlenie komunikatu w konsoli o możliwości zapisu parametrów i metody filtrowania do pliku.

|  |  |
| --- | --- |
|  |  |

## Program wykonawczy

Zgodnie z ogólnym schematem działania głównego programu przedstawionym na rys. XX pierwszym etapem po uruchomieniu programu jest inicjalizacja. W tym kroku podobnie jak w programie kalibrującym, należy podać numery ID kamer. Oprócz tego wczytywane są także dane z plików zewnętrznych, wymagane przy późniejszych procesach. W przypadku braku któregokolwiek z pliku program informuje w konsoli o braku danego pliku i kończy swoje działanie. Drugim etapem jest nawiązanie połączenia z robotem. Prowadzony jest dialog w konsoli, w celu uzyskania skonfigurowanego adresu IP robota i numeru portu zdefiniowanego w programie robota. Po poprawnym wprowadzeniu wszystkich danych i wczytaniu plików konfiguracyjnych program przechodzi do głównej pętli programu.

Pierwszym krokiem jest pobranie z obu kamer obrazów, które następnie zostają skorygowane w procesie transformacji uwzględniającej parametry wczytane z wynikowego pliku kalibracji kamer. W kolejnym kroku skorygowane obrazy zostają poddane filtracji, która wykonuje progowanie na podstawie metody i wartości uzyskanych z pliku ustawień filtrowania. Jeżeli na obu wynikowych, dwustanowych obrazach zostanie wykryty punkt charakterystyczny, funkcja triangulacji wyznacza jego położenie przestrzenne względem układu kamer. Pierwsze dwie próbki są ignorowane, aby dać możliwość operatorowi wskazania laserem odpowiedniego miejsca. Kolejne próbki włączane są do bufora, a po przekroczeniu wymaganej ilości próbek liczona jest średnia wartość położenia, która następnie jest transformowana do układu współrzędnych robota. Położenie punktu dla każdej z trzech osi zamieniane jest na ciąg znaków, który jest przesyłany do robota zgodnie z protokołem TCP/IP, pod warunkiem poprawnie ustanowionego połączenia.

## Klasa CStereoCalibration

## Klasa CFilterCalibration

## Klasa CStereoVision

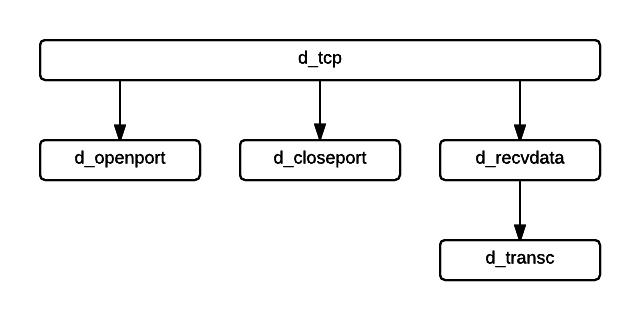
## Klasa CTCPConnection

# Oprogramowanie robota

## Struktura

Oprogramowanie robota składa się z głównego programu *d\_tcp* i 4 podprogramów odpowiedzialnych za:

* d\_openport – poprawne otworzenie portu TCP i nasłuchiwania na prośbę połączenia
* d\_closeport – zatrzymanie nasłuchiwania i zamknięcie portu
* d\_recvdata – odbieranie otrzymywanych pakietów danych
* d\_transc – interpretacje otrzymanych danych i wykonanie odpowiedniej akcji



# Testowanie oprogramowania

## Kalibracja kamer

Pozyskiwany z kamer obraz nie jest idealnym dwuwymiarowym rzutem monitorowanej przestrzeni. Głównymi przyczynami zaburzeń obrazu są niedokładności w wykonaniu oraz pozycjonowaniu względem siebie elementów urządzenia. O ile centralny obszar obrazu nie jest zdeformowany, o tyle zewnętrzna część jest tym bardziej zniekształcona, im dalej od centrum obrazu. Skutkami tego może być zaniżona dokładność w wyznaczaniu przestrzennego położenia punktu przez triangulacje, dlatego przed uruchomieniem głównego programu stereowizyjnego należy skrupulatnie przeprowadzić proces kalibracji obrazu.

Kalibracja kamer w układzie stereowizyjnym polega na wyznaczeniu współczynników korekcyjnych omówionych w rozdziale 3. W tym celu wymagane jest pobranie próbek obrazów, na których będzie możliwe wyznaczenie specyficznych punktów, znając jednocześnie rzeczywiste powiązania wymiarowe pomiędzy tymi punktami. Idealnym wzorcem, którego punkty można łatwo wykryć i zmierzyć jest szachownica. Ilość pól wzdłuż i wszerz oraz długość boku pola w mm są niezbędnymi parametrami, które należy wprowadzić jako argumenty do programu kalibrującego. Podczas pobierania próbek obrazu, wyświetlane są obrazy z obu kamer, aby ułatwić poprawną manipulacje szachownicy w kadry kamer. Gdy program wykrywa specyficzne punkty – wierzchołki kwadratów – na obrazie wejściowym, w odpowiednich miejscach rysowane są kolorowe okręgi. Pomiędzy pobraniem próbek obrazów występuje odstęp czasowy ( domyślnie dwusekundowy), aby zapobiec zduplikowaniu próbek, co niekiedy może prowadzić do błędnych wyników kalibracji.

Ilość próbek, na podstawie których przeprowadzane są obliczenia, decyduje o dokładności wyników i długości przetwarzania danych. Na wykresie [NUMER] przedstawiono wyniki badanej zależności czasu obliczeń od ilości próbek. Należy mieć na uwadze, że system operacyjny przeznaczał część czasu procesora na inne procesy, przez co wyniki mają zaniżoną dokładność, jednak widoczna jest tendencja wzrostowa.

## Filtrowanie

Wyznaczenie efektywnej metody filtrowania i jej parametrów jest jednym z najtrudniejszych procesów przetwarzania obrazów. Wydobycie z obrazu obiektów zainteresowania utrudnia wiele czynników, m.in.:

* Niejednorodność tła
* Zmienne oświetlenie
* Zmienne parametry kamery
* Szumy

Niektóre z powyższych problemów można rozwiązać stosując dodatkowe źródło stałego oświetlenia, zapewniające jednakową jasność monitorowanego obszaru lub wykorzystać jednokolorowe maty, podłoża ułatwiające wykrycie i odrzucenie tła. Dopracowany algorytm filtrowania pozwala często uniknąć stosowania dodatkowych procedur pochłaniających wiele zasobów i czasu, mających na celu sprecyzowanie położenia szukanego obiektu. Z tego powodu istotne jest przeprowadzenie poprawnej i dokładnej filtracji na obrazie wejściowym.

Testowany filtracja przebiega na dwóch poziomach. Pierwszym poziomem są wewnętrzne parametry kamery Logitech C905, które można zmieniać dzięki oprogramowaniu producenta. Niestety nie można wprowadzać, ani odczytywać wartości liczbowych konkretnych parametrów, a jedynie ustalać ukryte wartości przez pasek przesuwny. Istotne jest, aby ustawienia na obu kamerach były jednakowe. W przeciwnym razie kolejna warstwa filtrowania będzie operować na różnych obrazach wejściowych, co może doprowadzić do wykrycia błędnych punktów, bądź też wyeliminowania poprawnych. Wyróżniono dwie grupy parametrów: ustalone automatycznie przez oprogramowanie kamery oraz ręcznie.

|  |  |
| --- | --- |
| Nastawy automatyczne:  Ekspozycja, wzmocnienie i równowaga bieli dobierane są na podstawie analizy pobranego, aktualnego obrazu. |  |
| Nastawy ręczne:  Wszystkie parametry są dobierane ręcznie i pozostają stałe. |  |

Druga warstwa odnosi się do sposobu zapisu przestrzeni kolorów do postaci cyfrowej. Jednym z modeli często wykorzystywanym w informatyce jest trójwymiarowy układ RGB, gdzie punkt reprezentowany jest przez trzy współrzędne barw przyjmujące wartości od 0 do 255: R – czerwonej, G – zielonej oraz B – niebieskiej. Kombinacja współrzędnych w odpowiednich proporcjach daje dowolny kolor. Innym modelem jest układ HSV, który dzieli się na współrzędne: H – barwa, S – nasycenie oraz V – jasność i jego reprezentacja geometryczna jest w postaci walca. Oba wyżej wymienione sposoby reprezentacji koloru zostały wykorzystane do progowanie obrazu. Jeśli współrzędne punktu w danej przestrzeni barw znajdowały się w przedziale określonych wartości to punkt ten przyjmował wartość 255 w jednowymiarowym obrazie:

|  |  |
| --- | --- |
|  |  |

Celem testu filtrowania było stworzenie obrazu z wyodrębnionym obszarem plamki światła lasera o długości 690 nm i mocy 35 mW oraz odrzucenie pozostałych, nieistotnych informacji tła. Przetworzony dwuwymiarowy obraz zawiera piksele, które mogą przyjąć jeden z dwóch stanów: 0 – informacja nieistotna lub 255 – szukana informacja. Po ustawieniu wewnętrznych parametrów kamery, uruchomiono program, który w czasie rzeczywistym pobierał i progował obraz z kamer w wybranej przestrzeni barw. Doświadczalnie dobierano parametry progowania, śledząc na bieżąco wyświetlany obraz wynikowy, tak aby uzyskać wyłącznie obszar wskazany przez światło lasera. Wyniki testowania przy parametrach opisanych w TAB\_X przedstawiono w TAB\_Y, gdzie czerwonym okręgiem zaznaczono szukany, czerwony punkt.

Dla automatycznych nastaw, pomimo różnych kombinacji zakresu progowania, na żadnym obrazie z obu przestrzeni barw nie udało się całkowicie odrzucić tła. W przypadku obrazu RGB niektóre z białych, jaśniejszych pikseli pozostało na wynikowym obrazie. Progowanie przestrzeni HSV wyeliminowało zdecydowaną większość tła, pozostawiając jedynie kilka pikseli, kosztem utraty dużej części obszaru światła lasera, przez co trudno odróżnić szukany punkt od kilku nieodfiltrowanych pikseli tła.

Druga seria obrazów wejściowych, została pobrana przez kamerę o ręcznie ustawionych parametrach. Zdecydowanie zmniejszono ekspozycję i wzmocnienie, aby zmniejszyć ilość nadmiernie jasnych pikseli. Jednocześnie zwiększono intensywność kolorów, by zintensyfikować widoczność czerwonej plamki światła. Progowanie obrazu wejściowego skutecznie wyeliminowało tło, pozostawiając jedynie piksele odpowiadające laserowi. W przestrzeni HSV utracono nieco więcej informacji, niż w układzie RGB.

Najlepszy rezultat filtrowania został uzyskany stosując ręczne nastawy kamery dla danych warunków otoczenia oraz progując wartości pikseli w przestrzeni RGB, gdzie udało się odfiltrować tło, bez odrzucania pikseli przypadających na światło lasera. Spostrzeżono, że zwiększenie intensywności kolorów w parametrach kamery powoduje okresową destabilizacje wynikowego obrazu progowania

|  |  |  |
| --- | --- | --- |
|  | Automatyczne nastawy kamery | Ręczne nastawy kamery |
| Obraz wejściowy |  |  |
| Progowanie w przestrzeni RGB |  |  |
| Progowanie w przestrzeni HSV |  |  |

Parametry progowania :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | B | G | R |  | H | S | V |
| Automatyczne  nastawy kamery | MIN | 0 | 0 | 255 | 7 | 0 | 255 |
| MAX | 210 | 224 | 255 | 26 | 133 | 255 |
| Ręczne  nastawy kamery | MIN | 0 | 0 | 96 | 100 | 0 | 114 |
| MAX | 161 | 162 | 255 | 200 | 255 | 255 |

Powyższe metody filtrowania zakładają, że plamka światła wskaźnika będzie najintensywniejszym obszarem na monitorowanej scenie. Podczas wielokrotnego testowania oprogramowania filtrującego w różnych warunkach oświetleniowych zauważono, że silne refleksy świetlne np. na gładkich metalowych powierzchniach skutecznie zakłócają procedurę triangulacji, z powodu wielu niewłaściwie wykrytych punktów. Metody progowania w obu przestrzeniach kolorów nie pozwoliły wyeliminować tego efektu.

## Analiza dokładności

# Podsumowanie

# Bibliografia

1. Cyganek Bogusław, „Komputerowe przetwarzanie obrazów”
2. <http://docs.opencv.org/3.0-beta/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html>
3. <https://pl.wikipedia.org/wiki/Licencja_BSD>
4. D. Rzeszotarski, P. Strumiłło, P. Pełczyński, B. Więcej, A. Lorenc, „System obrazowania stereoskopowego sekwencji scen trójwymiarowych”
5. G. Bradsky, A. Kaehler, „Learning OpenCV”
6. J. Howse, S. Puttemans, Q. Hua, U. Sinha, “OpenCV 3 Blueprints”
7. <http://docs.opencv.org/3.0-beta/modules/refman.html>
8. <https://pl.wikipedia.org/wiki/RS-232>
9. https://pl.wikipedia.org/wiki/User\_Datagram\_Protocol

# Spis ilustracji

# Spis tabel

# Spis załączników