



**WYDZIAŁ MATEMATYKI
i INFORMATYKI**

Uniwersytet Łódzki

Gabriel Ozeg

Nr albumu: 395263

System antykolizyjny na mikroprocesorze Raspberry Pi

Praca magisterska
na kierunku Informatyka

Praca wykonana pod kierunkiem
prof. dr hab. Paweł Zajączkowski
Katedra Opiekuna

Łódź, 2025

Słowa kluczowe: pierwsze, drugie, trzecie, czwarte

Title in English: Title in English

Keywords: first, second, third, fourth

Spis treści

1. Wstęp	5
2. Podstawowe pojęcia	7
2.1. Definicje i własności	7
2.2. Przykłady	7
3. Część główna	9
3.1. Czym jest stereowizja?	9
3.1.1. Stereowizja w robotach do pomieszczeń czystych	9
3.2. Model kamery	10
3.2.1. Ogniskowa obiektywu	10
3.2.2. Zniekształcenie obiektywu	10
3.2.3. Kalibracja za pomocą OpenCV	10
3.3. Obrazowanie stereoskopowe	10
3.3.1. Wyjaśnienie	10
3.3.2. Triangulacja	11
3.3.3. Geometria epipolarna	12
3.3.4. Macierze podstawowe i fundamentalne	12
3.3.5. Macierz obrotu i wektor przesunięcia	13
3.3.6. Rektyfikacja stereo	13
3.3.6.1. Algorytm Hartley’a	13
3.3.6.2. Algorytm Bougueta	14
4. Rozdział badawczy	15
4.1. Funkcjonalność programu do obrazowania stereo	15
4.1.1. Wykorzystane pakiety	16
4.1.2. Główna pętla	16
4.1.3. Funkcjonalność programu „Take-images-for-calibration.py”	16

4.1.3.1.	Wektory kalibracyjne	17
4.1.3.2.	Pozyskiwanie obrazów do kalibracji	17
4.1.4.	Funkcjonalność programu „Main-Stereo-Vision-Prog.py”	17
4.1.4.1.	Kalibracja zniekształceń	17
4.1.4.2.	Kalibracja kamery stereo	17
4.1.4.3.	Obliczanie mapy rozbieżności	17
4.1.4.4.	Zastosowanie filtra WLS (ważonych najmniejszych kwadratów)	17
4.1.4.5.	Pomiar odległości	17
4.1.4.6.	Możliwe ulepszenia	17
5.	Zakończenie	19
	Bibliografia	21

Rozdział 1

Wstęp

We wstępie pracy dyplomowej powinien znaleźć się opis wkładu własnego studenta w uzyskanie przedstawianych wyników a także informacje o podstawowych źródłach, na podstawie których student przygotował pracę.

Rozdział 2

Podstawowe pojęcia

2.1. Definicje i własności

W niniejszej części pracy podane zostaną pojęcia niezbędne w późniejszych rozważaniach (patrz [2] lub [?]).

Definicja 2.1.1. *Niech G będzie niepustym zbiorem. Działaniem w G nazywamy dowolne odwzorowanie $\circ : G \times G \rightarrow G$.*

Definicja 2.1.2. *Niech G będzie niepustym zbiorem, \circ działaniem w G . Element $e \in G$ nazywamy neutralnym (działania \circ), jeśli dla każdego $a \in G$ mamy $a \circ e = e \circ a = a$.*

Lemat 2.1.1. *Jeśli działanie \circ w G posiada element neutralny, to jest on jeden.*

Dowód. Niech $e, e' \in G$ będą dwoma elementami neutralnymi. Wtedy

$$e = e' \circ e = e'. \quad (2.1)$$

Zatem element neutralny jest jeden. ■

2.2. Przykłady

Działaniem w zbiorze liczb naturalnych jest dodawanie, natomiast działaniem w tym zbiorze nie jest odejmowanie.

Rozdział 3

Część główna

3.1. Czym jest stereowizja?

Stereoskopia to proces polegający na przechwyceniu 2 obrazów tej samej sceny w celu utworzenia mapy dysproporcji sceny. mapa rozbieżności sceny. Z tej mapy rozbieżności można zmierzyć odległość do obiektu i utworzyć mapę 3D sceny. do obiektu i utworzyć mapę 3D sceny.

3.1.1. Stereowizja w robotach do pomieszczeń czystych

Celem tego dużego projektu jest opracowanie robota do pomieszczeń czystych, który może mierzyć cząsteczki w całym pomieszczeniu czystym. w całym pomieszczeniu czystym. Aby przeprowadzić pomiar bez żadnych problemów, robot musi orientować się bez kolizji. robot musi orientować się bez kolizji. Do tej orientacji zdecydowaliśmy się użyć tylko kamer, a do pomiaru odległości od obiektu potrzebujemy co najmniej dwóch kamer co najmniej dwóch kamer (stereowizja). Dwie kamery znajdują się w odległości 110 mm między sobą.

Im większa jest ta odległość, tym lepiej można ocenić odległość do sceny dla obiektów znajdujących się daleko. obiektów, które są daleko. W tym projekcie zaimplementowano program Python wykorzystujący bibliotekę OpenCV do kalibracji kamer i pomiaru odległości do obiektów na scenie. obiektów w scenie. (Patrz załącznik)

Kamera stereo składa się z dwóch kamer Logitech Webcam C170.

Widoki wideo są optymalne przy obrazach 640x480 pikseli. Ogniskowa: 2,3 mm. Wbudowana kamera stereo:

3.2. Model kamery

Kamery rejestrują promienie świetlne z naszego otoczenia. Zasadniczo kamera działa jak nasze oko, odbite promienie światła z naszego otoczenia docierają do naszego oka i są zbierane na siatkówce. Są one gromadzone na naszej siatkówce. „Kamera otworkowa” jest najprostszym modelem. Jest to dobry uproszczony model do zrozumienia zrozumieć, jak działa kamera. W tym modelu wszystkie promienie światła są zatrzymywane przez powierzchnię. powierzchnię. Tylko promienie przechodzące przez otwór są przechwytywane i rzutowane w odwrotnej kolejności na powierzchnię w kamerze. powierzchnia w kamerze. Poniższa ilustracja wyjaśnia tę zasadę

Źródło: <https://funsizephysics.com/use-light-turn-world-upside/>

Zasada ta jest bardzo prosta, ale nie jest to dobry sposób na uchwycenie wystarczającej ilości światła przy szybkiej ekspozycji. szybkiej ekspozycji. Dlatego soczewki są używane do zbierania promieni światła w jednym miejscu. Problem polega na tym, że taki obiektyw powoduje zniekształcenia. zniekształcenia. Istnieją dwa różne rodzaje zniekształceń: - zniekształcenie promieniowe - zniekształcenie styczne Zniekształcenie promieniowe wynika z kształtu samego obiektywu, a zniekształcenie styczne wynika z geometrii kamery. Obrazy można następnie skorygować za pomocą metod matematycznych. metod matematycznych. Proces kalibracji umożliwia stworzenie modelu geometrii kamery i modelu zniekształceń obiektywu. zniekształcenia obiektywu. Modele te tworzą parametry wewnętrzne kamery.

3.2.1. Ogniskowa obiektywu

3.2.2. Zniekształcenie obiektywu

3.2.3. Kalibracja za pomocą OpenCV

3.3. Obrazowanie stereoskopowe

3.3.1. Wyjaśnienie

Stereo Vision umożliwia rozpoznawanie głębi na obrazie, wykonywanie pomiarów na obrazie i przeprowadzanie lokalizacji 3D. Między innymi należy znaleźć punkty, które należy znaleźć punkty, które pasują do siebie między dwiema kamerami. Można to następnie wykorzystać do odległość między kamerą a punktem. Wykorzystywana jest geometria systemu systemu w celu uproszczenia obliczeń.

Te cztery kroki są wykonywane podczas obrazowania stereo: 1. usuwanie zniekształceń

promieniowych i stycznych za pomocą obliczeń matematycznych obliczenia. W ten sposób powstają nieznieskształcone obrazy. 2. rektyfikacja kąta i odległości obrazów. Na tym etapie oba obrazy są obrazy współpłaszczyznowe na osi Y, co ułatwia wyszukiwanie korespondencji. łatwiejsze i wystarczy szukać tylko na jednej osi (osi X). (a mianowicie na osi X). 3. znajdź tę samą cechę na prawym i lewym obrazie. Daje to mapę mapę dysproporcji pokazującą różnice między obrazami na osi x. 4. Ostatnim krokiem jest triangulacja. Mapa rozbieżności jest przekształcana w odległości za pomocą triangulacji. Krok 1: Usuwanie zniekształceń Krok 2: Prostowanie Krok 3: Znalezienie tej samej cechy na obu obrazach Krok 4: Triangulacja

3.3.2. Triangulacja

W ostatnim kroku, triangulacji, zakłada się, że oba obrazy projekcji są współpłaszczyznowe i że poziomy rząd pikseli lewego obrazu jest wyrównany z odpowiadającym mu obrazem prawego. są współpłaszczyznowe i poziomy rząd pikseli lewego obrazu jest wyrównany z odpowiednim poziomym rzędem pikseli lewego obrazu. poziomy rząd pikseli lewego obrazu jest wyrównany. Poniższy obraz można teraz skonstruować przy użyciu poprzednich hipotez.

Punkt P leży w środowisku i jest pokazany na lewym i prawym obrazie na p_l i p_r , z odpowiadającymi im współrzędnymi odpowiadającymi współrzędnymi x_l i x_r . To pozwala nam wprowadzić nową wielkość $d = x_l - x_r$. Można zauważyć, że im dalej punkt P, tym mniejsza staje się wielkość d . Dysproporcja jest zatem odwrotnie proporcjonalna do odległości. odległości.

Do obliczenia odległości można użyć następującego wzoru można obliczyć: $Z = f \cdot T / (x_l - x_r)$

Można zauważyć, że istnieje nieliniowa zależność między rozbieżnością a odległością. istnieje. Jeśli rozbieżność jest bliska 0, małe różnice w rozbieżności prowadzą do dużych różnic w odległości. różnice w odległości. Zjawisko to ulega odwróceniu, gdy rozbieżność jest duża. Małe różnice małe różnice dysproporcji nie prowadzą do dużych różnic odległości. Na tej podstawie można wywnioskować, że stereowizja ma wysoką rozdzielczość głębi, tylko dla obiektów znajdujących się blisko kamery. które znajdują się blisko kamery.

Diese Methode funktioniert aber nur wenn die Konfiguration der Stereo-Kamera ideal ist. In der Realität ist dies jedoch nicht der Fall. Deswegen wird das linke und rechte Bild mathematisch parallel ausgerichtet. Natürlich müssen die Kameras zumindest approximativ physisch parallel positioniert werden. Bevor wir die Methode zur Mathematischen Ausrichtung der Bildern erklären, müssen wir zuerst die Epipolare Geometrie verstehen.

3.3.3. Geometria epipolarna

Das oben dargestellte Bild zeigt uns das Modell einer nicht-perfekten Stereo-Kamera die aus zwei Pinhole-Kameramodelle besteht. Durch die Kreuzung der Linie der Projektierungszentren (O_l , O_r) mit den Projektierungsebenen entstehen die Epipolarpunkte e_l und e_r . Die Linien (p_l , e_l) und (p_r , e_r) werden Epipolarlinien genannt. Das Bild aller möglichen Punkte eines Punkts auf einer Projektierungsebene ist die Epipolarlinie die auf der anderen Bildebene liegt und durch den Epipolarpunkt und dem Gesuchten Punkt geht. Dies ermöglicht, die Suche des Punkts auf einer einzigen Dimension zu begrenzen anstatt auf einer ganzen Ebene. Man kann also folgende Punkte zusammenfassen:

- Jeder 3D Punkt in der Sicht einer Kamera ist im Epipolaren Plan enthalten
- Ein Merkmal in einer Ebene muss sich auf der entsprechenden Epipolarlinien der andere Ebene befinden (Epipolarbedingung)
- Eine zweidimensionale Suche eines entsprechenden Merkmals wird zu einer eindimensionalen Suche umgewandelt, wenn man die Epipolargeometrie kennt.
- Die Reihenfolge der Punkte wird behalten, d.h. dass zwei Punkte A und B in derselben Reihenfolge auf der Epipolarlinien einer Ebene gefunden werden, wie auf der, der anderen Ebene.

3.3.4. Macierze podstawowe i fundamentalne

Aby zrozumieć, w jaki sposób obliczane są linie epipolarne, musimy najpierw wyjaśnić macierze podstawowe i macierze fundamentalne (odpowiadające macierzom E i F). Macierz podstawowa E zawiera informacje o tym, jak fizycznie rozmieszczone są obie kamery. są fizycznie rozmieszczone. Opisuje ona lokalizację drugiej kamery względem pierwszej za pomocą parametrów translacji i rotacji. względem pierwszej kamery za pomocą parametrów translacji i rotacji. Parametry te nie są bezpośrednio odczytywane w macierzy Parametrów tych nie można odczytać bezpośrednio w macierzy, ponieważ jest ona używana do planowania projektu. W sekcji Kalibracja stereo wyjaśnimy, jak obliczyć R i T (macierz rotacji i wektor translacji). Macierz F zawiera informacje z podstawowej macierzy E , fizyczny układ kamer i informacje o kamerach. kamer i informacje o wewnętrznych parametrach kamer. Relacja między rzutowanym punktem na lewym obrazie p_l i rzutowanym punktem na prawym obrazie p_r jest zdefiniowana następująco obraz p_r jest zdefiniowany następująco:

$p_r^T E p_l = 0$ Można by pomyśleć, że ta formuła w pełni opisuje związek między lewym i prawym punktem. prawym punktem. Należy jednak zauważyć, że macierz 3×3 E jest rzędu jest rangi 2. Oznacza to, że wzór ten jest równaniem prostej. Aby w pełni zdefiniować relację między punktami, parametry wewnętrzne. parametry wewnętrzne. Pamiętajmy, że $q = M p$, z macierzą wewnętrzną M . Podstawienie do poprzedniego równania daje wynik $q_r^T (M_l^{-1})^T E M_l^{-1} q_l = 0$ Podstawienie: $F = (M_l^{-1})^T E M_l^{-1}$ W ten sposób otrzymujemy $q_r^T F q_l = 0$

3.3.5. Macierz obrotu i wektor przesunięcia

Teraz, gdy wyjaśniliśmy już macierz podstawową E i macierz podstawową F , musimy musimy zobaczyć, jak obliczyć macierz obrotu i wektor translacji. Zdefiniujemy następujące oznaczenia: - P_l i P_r definiują pozycje punktu w układzie współrzędnych odpowiednio lewej i prawej kamery. prawa kamera - R_l i T_l (lub R_r i T_r) definiują obrót i translację z kamery do punktu w otoczeniu dla lewej (lub prawej) kamery. - R i T to obrót i translacja układu współrzędnych prawej kamery w układzie współrzędnych lewej kamery. R i T to obrót i translacja układu współrzędnych prawej kamery w układzie współrzędnych lewej kamery. Daje to następujące wyniki $P_l = R_l P_r + T_l$ i $P_r = R_r P_l + T_r$ Mamy również: $P_l = R T (P_r - T)$ Z tych trzech równań ostateczny wynik to $R = R_r R_l^T$ $T = T_r - R T_l$

3.3.6. Rektyfikacja stereo

Dotychczas zajmowaliśmy się tematem „kalibracji stereo”. Chodziło o opis geometrycznego rozmieszczenia obu kamer. Zadaniem rektyfikacji jest rzutowanie dwóch obrazów tak, aby leżały dokładnie w tej samej płaszczyźnie i precyzyjne wyrównanie rzędów pikseli tak, aby linie epipolarne stały się poziome w celu zapewnienia zgodności punktu na dwóch obrazach. aby znaleźć zgodność punktu na dwóch obrazach w sposób bardziej losowy. W wyniku procesu wyrównywania obu obrazów uzyskuje się 8 wyrażeń, po 4 dla każdej kamery. kamery: - wektor zniekształceń - macierz rotacji R_{rect} , która musi zostać zastosowana do obrazu - wyprostowana macierz kamery M_{rect} - nierektyfikowana macierz kamery M OpenCV pozwala nam obliczyć te warunki za pomocą dwóch algorytmów: algorytmu Hartleya i algorytmu Bougueta.

3.3.6.1. Algorytm Hartley'a

Algorytm Hartleya wyszukuje te same punkty na obu obrazach. Próbuje on stara się zminimalizować rozbieżności i znaleźć homografie, które ustawiają epipole w nieskończoności. nieskończoność. Dzięki tej metodzie nie jest więc konieczne obliczanie parametrów wewnętrznych dla każdej kamery. dla każdej kamery. Zaletą tej metody jest to, że kalibracja jest możliwa tylko dzięki obserwacji punktów w scenie. punktów na scenie. Główną wadą jest jednak brak skalowania obrazu Masz tylko informacje o względnej odległości. Nie można dokładnie zmierzyć jak daleko obiekt znajduje się od kamer.

3.3.6.2. Algorytm Bougueta

Algorytm Bougueta wykorzystuje obliczoną macierz obrotu i wektor translacji aby obrócić obie rzutowane płaszczyzny o pół obrotu, tak aby znalazły się w tej samej płaszczyźnie. tej samej płaszczyźnie. Sprawia to, że główne promienie są równoległe, a płaszczyzny współpłaszczyznowe. ale nie są jeszcze wyrównane w rzędach. Zostanie to zrobione później. W projekcie wykorzystaliśmy algorytm Bougueta.

Rozdział 4

Rozdział badawczy

Praca powinna spełniać wymogi formalne, merytoryczne i redakcyjne opisane w Regulaminie Studiów (Rozdział IX) oraz w uchwale nr 184 Rady Wydziału Matematyki i Informatyki UŁ z dnia 25.09.2019 ze szczególnym uwzględnieniem wymogu, aby była ona samodzielnym opracowaniem zagadnienia naukowego lub praktycznego albo dokonaniem technicznym, prezentującym ogólną wiedzę i umiejętności studenta, związanym ze studiami na danym kierunku, poziomie i profilu oraz umiejętności samodzielnego analizowania i wnioskowania (Ustawa 2.0 Art. 76 p. 2)

Praca dyplomowa będąca pracą inżynierską powinna zawierać samodzielne opracowanie praktycznego problemu i może mieć charakter projektu, studium porównawczego lub opracowania analitycznego.

4.1. Funkcjonalność programu do obrazowania stereo

Jak już wspomniano, program jest kodowany w Pythonie i wykorzystywana jest biblioteka OpenCV. jest używana. Zdecydowaliśmy się na język Python i bibliotekę OpenCV, ponieważ mieliśmy już z nimi doświadczenie i ponieważ istnieje wiele dokumentacji na ich temat. Innym argumentem za tą decyzją jest to, że chcieliśmy pracować tylko z bibliotekami „open source”. biblioteki. Na potrzeby tego projektu opracowano dwa programy w języku Python. Pierwszy z nich, „Take-images-for-calibration.py”, służy do robienia dobrych zdjęć, które są później wykorzystywane do kalibracji obu kamer. Później są one wykorzystywane do kalibracji obu kamer (kalibracja zniekształceń i kalibracja stereo). kalibracja. Drugi program, a tym samym główny program „Main-Stereo-Vision-Prog.py” jest używany do obrazowania stereo. jest używany do obrazowania stereo. W tym programie kalibrujemy kamery za pomocą wykonanych zdjęć, generujemy mapę dysproporcji i dzięki doświadczalnemu równaniu równania, które zostało znalezione eksperymentalnie, możemy zmierzyć odległość dla każ-

dego piksela. pomiar. Na końcu używany jest filtr WLS, aby lepiej rozpoznawać krawędzie obiektów. rozpoznać krawędzie obiektów.

4.1.1. Wykorzystane pakiety

Do programu zaimportowano następujące pakiety: - Wersja OpenCV.3.2.0 z opencv-contrib (zawiera funkcje stereo) jako „cv2” w Pythonie, zawiera: o bibliotekę do przetwarzania obrazów o funkcje do stereowizji - Numpy.1.12. o Używany do operacji na macierzach (obrazy składają się z macierzy) - Skoroszyt z openpyxl o Pakiet umożliwiający zapisywanie danych w pliku Excel - „normalizacja” biblioteki sklearn 0.18.1 o sklearn umożliwia uczenie maszynowe, ale w tym projekcie używany jest tylko filtr WLS

4.1.2. Główna pętla

Aby pracować z kamerami, należy je najpierw aktywować. Funkcja `cv2.VideoCapture()` aktywuje obie kamery poprzez wprowadzenie numeru portu każdej z nich. kamery (w programie tworzone są dwa obiekty korzystające z metod klasy `cv2.VideoCapture()`). klasy `cv2.VideoCapture()`.

Aby uzyskać obraz z kamer, używana jest metoda `cv2.VideoCapture().read()` Wyjściem jest obraz sceny, którą kamera ogląda w momencie wywołania tej funkcji. funkcja jest wywoływana. Aby uzyskać obraz wideo, należy wywołać tę metodę w nieskończonej pętli. Aby być bardziej wydajnym, zaleca się aby przekonwertować obrazy BGR na obrazy w odcieniach szarości, odbywa się to za pomocą funkcji `cv2.cvtColor()` funkcja.

Aby wyświetlić wideo na komputerze, używana jest funkcja `cv2.imshow()`. aby otworzyć okno, w którym można wyświetlić wideo.

Przerwanie służy do wyjścia z nieskończonej pętli. Staje się ona aktywna za każdym razem, gdy użytkownik naciśnie spację. Rozpoznanie, że klawisz został naciśnięty jest rozpoznawane dzięki funkcji `cv2.waitKey()`. Wreszcie, dwie używane kamery muszą zostać dezaktywowane za pomocą metody `cv2.VideoCapture().release()`, a otwarte okna są niszczone za pomocą funkcji `cv2.destroyAllWindows()`.

4.1.3. Funkcjonalność programu „Take-images-for-calibration.py”

Po uruchomieniu tego programu obie kamery stają się aktywne i otwierane są dwa okna. aby użytkownik mógł zobaczyć, gdzie na obrazach znajduje się szachownica.

4.1.3.1. Wektory kalibracyjne

Funkcja `cv2.findChessboardCorners()` wyszuka określoną liczbę narożników szachownicy i wygenerowane zostaną następujące wektory: - `imgpointsR`: zawiera współrzędne narożników na prawym obrazie (w przestrzeni obrazu) - `imgpointsL`: zawiera współrzędne narożników na lewym obrazie (w przestrzeni obrazu) - `objpoints`: zawiera współrzędne narożników w przestrzeni obiektu. Precyzja współrzędnych znalezionych narożników jest zwiększana za pomocą funkcji `cv2.cornerSubPix()`.

4.1.3.2. Pozyskiwanie obrazów do kalibracji

4.1.4. Funkcjonalność programu „Main-Stereo-Vision-Prog.py”

4.1.4.1. Kalibracja zniekształceń

4.1.4.2. Kalibracja kamery stereo

4.1.4.3. Obliczanie mapy rozbieżności

4.1.4.4. Zastosowanie filtra WLS (ważonych najmniejszych kwadratów)

4.1.4.5. Pomiar odległości

4.1.4.6. Możliwe ulepszenia

Rozdział 5

Zakończenie

Bibliografia

- [1] OpenCV, <https://opencv.org>.
- [2] Aleksiej Kostrykin, *Wstęp do algebry. Podstawy algebry*, Warszawa, Wydawnictwo Naukowe PWN, 2022.
- [3] John Lambert, <https://johnwlambert.github.io/stereo/>, April 4, 2025.
- [4] Rajesh Rao, *Lecture 16: Stereo and 3D Vision*, <https://courses.cs.washington.edu/courses/cse455/09wi/L>, University of Washington.