**politechnika częstochowska**

**WYDZIAŁ INŻYNIERII PRODUKCJI I TECHNOLOGII MATeRIAŁÓW**

**INSTYTUT FIZYKI**

A description...

Adam Świącik

nr albumu: 118375

PRACA DYPLOMOWA MAGISTERSKA

Studia stacjonarne

**Temat:**

**Wykorzystanie fotogrametrii do wyznaczania geometrii obiektów**

Promotor:

**dr inż. Rafał Dobrakowski**

**Streszczenie**

Praca zawiera opis implementacji oprogramowania komputerowego przeznaczonego do cyfrowej rekonstrukcji obiektów rzeczywistych, poprzez kolejno: wczytanie serii zdjęć badanego przedmiotu wykonanych z różnej perspektywy, wyznaczenie za ich pomocą struktury przedmiotu, a następnie rekonstrukcję powierzchni i wizualizację wyników. Badania przeprowadzono na kilku obiektach sprawdzając zachowanie się algorytmu w różnych realnych sytuacjach.

Część merytoryczna pracy została podzielona na cztery rozdziały.

Wstęp, obejmuje charakterystykę poruszanego w pracy problemu, motywację do podjęcia właśnie tej tematyki oraz przegląd używanych w celach rekonstruktorskich technik pomiarowych, z wyszczególnieniem technik fotogrametrycznych. Rozdział zamyka przegląd literatury oraz przedstawienie celu pracy.

Rozdział drugi stanowi wprowadzenie teoretyczne do poruszanego problemu. Każdy podrozdział poświęcony jest jednemu z etapów składających się na proces rekonstrukcji, zaczynając od akwizycji obrazu i otrzymania chmury punktów, po jej przetwarzanie, rekonstrukcję i wizualizację

W rozdziale *Realizacja* znajduje się, opis programu w tym diagram klas, a także informacja o użytych podczas jego budowy bibliotek programistycznych. Opisano w nim przebieg samego badania, prezentację otrzymanych wyników oraz ich krótkie porównanie z wynikami autorskiego skanera laserowego.

Ostatni rozdział to wnioski z przeprowadzonych badań, podsumowanie wykonanych prac oraz komentarz dotyczący dalszych planów rozwojowych nad programem.

Spis treści

[1 Wstęp 5](#_Toc446708229)

[1.1 Przegląd technik pomiarowych 6](#_Toc446708230)

[1.2 Fotogrametria 8](#_Toc446708231)

[1.3 Przegląd literatury 10](#_Toc446708232)

[1.4 Cel pracy 11](#_Toc446708233)

[2 Wprowadzenie teoretyczne 12](#_Toc446708234)

[2.1 Akwizycja obrazu 13](#_Toc446708235)

[2.1.1 Matematyczny model kamery 13](#_Toc446708236)

[2.1.2 Dystorsja 15](#_Toc446708237)

[2.1.3 Kalibracja 17](#_Toc446708238)

[2.2 Rekonstrukcja geometrii 18](#_Toc446708239)

[2.2.1 Detekcja cech 18](#_Toc446708240)

[2.2.2 Obliczenie deskryptorów 20](#_Toc446708241)

[2.2.3 Dopasowanie punktów charakterystycznych 21](#_Toc446708242)

[2.2.4 Wyznaczenie macierzy projekcji 23](#_Toc446708243)

[2.2.5 Triangulacja 23](#_Toc446708244)

[2.3 Rekonstrukcja powierzchni 26](#_Toc446708245)

[2.3.1 Usunięcie punktów odstających 26](#_Toc446708246)

[2.3.2 Uproszczenie 27](#_Toc446708247)

[2.3.3 Wygładzenie 27](#_Toc446708248)

[2.3.4 Wyznaczenie wektorów normalnych 27](#_Toc446708249)

[2.3.5 Próbkowanie 28](#_Toc446708250)

[2.3.6 Rekonstrukcja Poissona 28](#_Toc446708251)

[2.4 Wizualizacja 29](#_Toc446708252)

[2.4.1 Operacja na wierzchołkach 30](#_Toc446708253)

[2.4.2 Przycinanie i budowanie prymitywów 30](#_Toc446708254)

[2.4.3 Rasteryzacja 31](#_Toc446708255)

[2.4.4 Operacje na fragmentach 31](#_Toc446708256)

[3 Realizacja 31](#_Toc446708257)

[3.1.1 Schemat działania 31](#_Toc446708258)

[3.1.2 Narzędzia 31](#_Toc446708259)

[3.1.3 Opis aplikacji 31](#_Toc446708260)

[3.1.4 Problemy 31](#_Toc446708261)

[3.1.5 Wyniki badań 31](#_Toc446708262)

[4 Wnioski 31](#_Toc446708263)

[5 Bibliografia 32](#_Toc446708264)

# Wstęp

W dobie automatyzacji i autonomizacji, istnieje rosnący trend wskazujący na coraz większą potrzebę gromadzenia możliwie jak największej ilości danych o otaczającym nas świecie i o nas samych. Do niedawna za sprawą komputeryzacji proces ten dotyczył głównie danych tekstowych. Obecnie jest to nie wystarczające źródło informacji, dlatego coraz częściej jego miejsce zajmuje obraz.

Na wzór rozwiązań które zaproponowała natura, od dawna buduje się systemy wizyjne, których zadaniem jest akwizycja, a następnie analiza i dalsze przetwarzanie danych pochodzących z obrazu. Mimo powstania dotychczas wielu takich systemów nadal nie spełniają one wielu założeń inżynierów przez co nadal jest to jeden z ważniejszych problemów z jakimi mierzy się obecnie chociażby robotyka.

Analiza obrazu często pozwala uzyskać geometryczne informacje o obserwowanym obiekcie. Jest to możliwe dzięki użyciu technik fotogrametrycznych, które stanowią narzędzie do rekonstrukcji geometrii badanego przedmiotu, a ta reprezentowana jest najczęściej w postaci chmury punktów. Uzyskane w ten sposób dane mogą następnie posłużyć do rekonstrukcji powierzchni i przedstawienia rzeczywistego obiektu w formie modelu 3D. Z takiej możliwości korzysta się coraz częściej, m.in.:

* w policji, w celu dokumentowania wypadków drogowych i przestępstw [1] [2] [3],
* w medycynie, do badania wad postawy [4], szeroko rozumianej protetyki i podczas badań z użyciem tomografii komputerowej czy rezonansu magnetycznego [5].
* w archeologii i muzealnictwie pozwala na digitalizację cennych obiektów historycznych [6] oraz tworzy trójwymiarową dokumentację dzieł sztuki [7] [8].
* W inżynierii odwrotnej która stała się w ostatnim czasie bardzo popularna, przy wykorzystaniu skanerów 3D znajduje zastosowanie chociażby do rekonstrukcji geometrii, a następnie modelowania 3D [9] [10], planowania ergonomicznych stanowisk pracy [11] i inwentaryzacji pomieszczeń biurowych [12]
* W budownictwie, do badania deformacji elementów konstrukcyjnych [13] , oceny stanu technicznego obiektów budowlanych [14]

Jednym z ważniejszych zastosowań wspomnianych systemów wizyjnych i cyfrowej rekonstrukcji otoczenia, jest budowa autonomicznych samochodów, robotów ratowniczych, i im podobnych. Jest to bardzo aktualny kierunek badań, ponieważ istnieje duże zapotrzebowanie na tego typu technologię. Dowodem na to są sponsorowane przez amerykańską agencję rządową DARPA (Agencję Zaawansowanych Projektów Badawczych w Obszarze Obronności) konkursy i zawody, z udziałem wysoce zaawansowanych i zautomatyzowanych robotów. Elementarnym problemem podczas budowy takich robotów jest otrzymanie, analiza i interpretacja informacji o ich najbliższym otoczeniu, a źródłem tych informacji w głównej mierze jest właśnie obraz.

Wymienione tutaj przykłady zastosowań oraz wskazane kierunki rozwoju były główną motywacją do podjęcia w pracy tematyki związanej z cyfrową rekonstrukcją obiektów w oparciu o analizę obrazu.

## Przegląd technik pomiarowych

W celu uzyskania informacji o geometrii obiektu należy dokonać jego zbadania poprzez pomiar, a ten można wykonać na szereg różnych sposobów. Zasadniczo techniki pomiarowe podzielone są na stykowe/dotykowe oraz bezkontaktowe/optyczne.

Rysunek 2.1.1.2 Podział technik pomiarowych

W przypadku pierwszej grupy, tj. metod dotykowych podczas dokonywania pomiaru następuje kontakt aparatury mierzącej z badanym przedmiotem, przez co zalicza się tę grupę to metod inwazyjnych. Cechą pomiarów dotykowych jest wysoka dokładność, a zarazem długi czas pomiaru. Korzystające z nich maszyny są generalnie drogie, mało mobilne i skomplikowane w użyciu, przez co znajdują zastosowanie z reguły w procesach wymagających pomiarów o szczególnie dużej precyzji. Najczęściej są to maszyny CNC oraz współrzędnościowe maszyny pomiarowe wyposażone w precyzyjne sondy zamontowane na ruchomych ramionach.

W metodach bezdotykowych, jak sugeruje nazwa, pomiarów dokonuje się w sposób pośredni, głównie z zastosowaniem szeroko rozumianego promieniowania elektromagnetycznego. Począwszy od fal radiowych wykorzystywanych w radarach, po podczerwień użytą w dalmierzach laserowych i promienie rentgena z których korzysta m.in. tomografia komputerowa. Z racji tego, że wyłącznie światło widzialne znajduję się w kręgu tematycznym tej pracy, stąd zastosowanie tylko tego promieniowania będzie dalej omawiane.

Optyczne metody pomiarowe

W grupie pomiarów optycznych, ze względu na rodzaj użytego podczas badania światła, rozróżnia się techniki aktywne i pasywne.

Pomiary pasywne w przeciwieństwie do technik aktywnych nie wykorzystują żadnych dodatkowych źródeł światła oprócz tego które dostępne jest w danej chwili. W tej grupie znajdują się wszystkie techniki fotogrametryczne.

Techniki aktywne korzystają z własnego źródła światła, które zostaje skierowane na badany przedmiot, odbite od jego powierzchni i zarejestrowane przez specjalny detektor. Obraz z detektora zostaje następnie poddany analizie i dopiero na jej podstawie dokonuje się niezbędnych pomiarów. Do tej grupy, zalicza się takie techniki jak:

* Triangulacja laserowa

Ogólna zasada pomiaru z użyciem triangulacji laserowej sprowadza się do oświetlenia badanego przedmiotu wiązką promieniowania laserowego i zarejestrowanie obrazu powstałej na powierzchni przedmiotu krzywej. Jej kształt jest ściśle związany z topologią powierzchni przedmiotu i z perspektywą z której jest obserwowana. Badany przedmiot umieszczony jest na stole obracanym przez silnik krokowy o stały kąt. Na podstawie serii tak wykonanych zdjęć i przy wykorzystaniu triangulacji wyznacza się położenie punktów które na pojedynczym obrazie widoczne są w formie jasnej linii.

Do zalet triangulacji laserowej należy zaliczyć fakt, że do pewnego stopnia nasłonecznienia możliwe jest użycie jej przy naturalnym oświetleniu. Konstrukcja zbudowanych w oparciu o nią urządzeń jest stosunkowo prosta, przez co ich obsługa jest ułatwiona, a koszt zakupu mniejszy w porównaniu do konkurencyjnych rozwiązań.

Jej wady to mały obszar skanowania i mała ilość uzyskanych jednorazowo punktów co jest jednoznaczne ze stosunkowo długim czasem skanowania.



Rysunek 2.1.1.3 Triangulacja laserowa. Str.15 [8]

* Technika światła strukturalnego

W technice światła strukturalnego, w przeciwieństwie do w/w konstrukcji, miejsce lasera zajmuje projektor, który rzutuje na dany obiekt specjalny wzorzec. Ten podobnie jak w przypadku skanera laserowego jest rejestrowany przez kamerę. Następnie dokonuje się analizy zniekształceń wzorca po odbiciu od powierzchni przedmiotu i wyznacza współrzędne trójwymiarowe punktów.

Charakterystyczne dla tej metody jest to, że podczas wykonania jednego pomiaru otrzymuje się znaczną ilość informacji ponieważ powierzchnia przedmiotu jest często w całości oświetlona wzorcem. Ponadto gęstość otrzymanych punktów pomiarowych jest równomierna. Ułatwia to znacząco późniejsze prace podczas rekonstrukcji powierzchni. Użycie projektora oprócz wymienionym zalet jest jednocześnie słabszą stroną tej metody, ponieważ urządzenie to musi odznaczać się bardzo dobrymi parametrami technicznymi w tym rozdzielczością i jasnością. Stosowanie tej metody wymaga odpowiedniego zaciemnienia pomieszczenia.



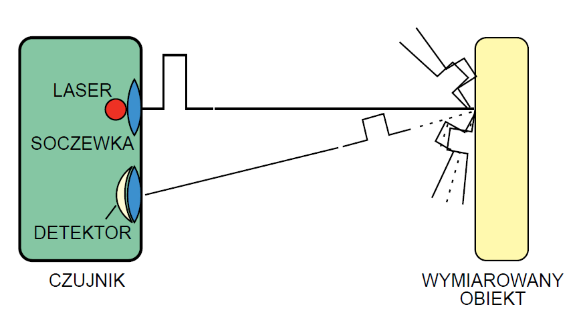
Rysunek 2.1.1.4 Zasada pomiaru w technice   
światła strukturalnego. Str.19 [8]

* Technika czasu przelotu wiązki

Książkowe systemy wykorzystujące technikę czasu przelotu zbudowane są z dwóch elementów. Źródła światła, którym jest laser oraz detektora promieniowania. Ogólna zasada pomiaru jest stosunkowo prosta. Wysyłane przez laser sygnały świetlne zostają odbite od badanego przedmiotu i rejestrowane przez detektor, następnie obliczany jest czas pomiędzy wysłaniem a rejestracją sygnału. Na podstawie czasu zostaje wyznaczona odległość pomiędzy przedmiotem a aparaturą mierzącą.

W praktyce, są to układy bardzo drogie i skomplikowane. Wykorzystywane najczęściej do badania otoczenia poruszających się autonomicznych pojazdów. W takim przypadku, w celu zwiększenia ilości gromadzonych danych, laser wraz z detektorem umiejscowiony jest na obracającym się z dużą szybkością stole. Omiatając wielokrotnie będący w zasięgu widzenia teren.

Jest to technika jak wspomniano wcześniej przeznaczona głównie do badania dużych obszarów, gdzie inne rozwiązania nie zdają egzaminu. Można jej użyć w każdych rozsądnych warunkach oświetlenia. Największą wadą która towarzyszy tej metodzie w porównaniu do poprzednich rozwiązań jest niska dokładność, którą do pewnego stopnia rekompensuje pole powierzchni skanowanego obszaru. W celu zminimalizowania błędu pomiaru wprowadza się do tej techniki różne udoskonalenia. Jednym z nich jest modulacja wiązki światła.



Rysunek 2.1.1.5 Pomiar odległości techniką czasu przelotu wiązki. Str.17 [8]

## Fotogrametria

Nazwa „*fotogrametria”* pochodzi od greckich słów „*photos”* oznaczające światło, „*gramma”* – zapis, oraz *„metreo”* – mierzę.

Fotogrametria jest nauką i technologią pozwalającą otrzymać informacje geometryczne o obserwowanych obiektach i ich otoczeniu, poprzez rejestrację, analizę i interpretację wykonanych zdjęć fotograficznych [15]

Początkowo znalazła głównie zastosowanie w geodezji i kartografii, gdzie za jej pomocą sporządzano mapy topograficzne terenu. Było to związane w tamtym okresie czasu z równoległym rozwojem jej i lotnictwa. Trudno znaleźć obecnie gałąź nauki i techniki, w której fotogrametria nie znajduje zastosowania. Można tu wymienić chociażby architekturę, medycynę, kryminalistykę czy systemy zabezpieczeń. Jako dyscyplina inżynierska odgrywa znaczącą rolę na takich polach jak robotyka czy komputerowa analiza i przetwarzanie obrazów [16].

U podstaw fotogrametrii leży układ wzroku człowieka. Zdolność percepcji głębi, jaką posiadamy ma swój początek właśnie w układzie wzrokowym. Obserwowany przez nas świat wygląda delikatnie inaczej z perspektywy lewego i prawego oka. Zarejestrowane dwa nieznacznie różniące się obrazy, są analizowane przez mózg, który zamienia je w trójwymiarowy model. Mówimy wtedy o stereowidzeniu. Nie jest to bynajmniej jedyny sposób na postrzeganie głębi. Zasłaniając jedno oko nadal posiadamy zdolność widzenia przestrzennego. Jest to wynik wielu różnych procesów z których najważniejsze to mimowolne ruchy głowy i gałki ocznej. Dzięki nim mózg jest w stanie zrekonstruować trójwymiarowy model otoczenia. Robi to analizując sekwencję obrazów rejestrowanych przez jedno oko. Wykonując ruch zmieniamy perspektywę, przez co następujące po sobie obrazy nieznacznie różnią się od siebie. Dalej podobnie jak ma to miejsce w przypadku widzenia stereoskopowego następuje podświadoma rekonstrukcja.

Fotogrametria jest próba przeniesienia tych mechanizmów na pole technicznych zastosowań.

Techniki fotogrametryczne

W ciągu kilkudziesięciu lat od początków fotogrametrii, powstało wiele technik fotogrametrycznych. Należą do nich m.in:

* fotogrametria jednoobrazowa – pozwala wyznaczyć położenie punktów na płaszczyźnie, stąd jej stosowanie jest ograniczone. W przypadku wykorzystania jej w tworzeniu map, rejestrowany teren powinien być płaski. W przypadku dokumentacji obiektów zabytkowych, rejestrowanymi przedmiotami mogą być np. malowidła lub płaskie elewacji budynków [3].
* fotogrametria dwuoobrazowa (stereofotogrametria) – przeznaczona do wyznaczania współrzędnych punktów w przestrzeni.
* fotogrametria lotnicza – korzysta się w niej ze zdjęć fotograficznych wykonanych z powietrza. Duży postęp w tej technice dokonał się początkowo w okresie dwóch wojen światowych i był mocno związany z rozwojem lotnictwa [15].
* fotogrametria naziemna – zdjęcia wykonywane są ze specjalnie przygotowanych stanowisk naziemnych

Profesjonalny sprzęt fotogrametryczny charakteryzuje wysoka precyzja i jakość wykonanych w nim elementów. Dołączana do niego dokumentacja techniczna pozwala wprowadzić dodatkowe korekty w obliczeniach co przekłada się na dokładność otrzymanych wyników. Jest to niestety powiązane z bardzo dużymi kosztami zakupu.

Dla zastosowań nie wymagających tak dużych precyzji, zaczęto szukać bazujących na fotogrametrii tańszych rozwiązań. Jedną z nich i w ostatnim czasie mocno rozwijaną jest technika *Structure from Motion*. Wykorzystuje ona możliwości amatorskiego sprzętu fotograficznego, który wbudowany jest dzisiaj w niemal każdy smartfon czy tablet. Zadaniem SfM jest rekonstrukcja sceny/obiektu na podstawie sekwencji zarejestrowanych obrazów przy założeniu, że zmiana położenia kamery, podczas wykonywania zdjęć jest początkowo nieznana i należy ją wyznaczyć. Kiedy zostanie już wyznaczona, możliwe jest użycie triangulacji w celu otrzymaniu struktury, tj. współrzędnych 3D badanego obiektu.

Cały ten proces można podzielić na kilka etapów:

* Detekcja punktów charakterystycznych
* Dopasowanie w/w punktów
* Wyznaczenie zmiany położenia i orientacji kamery
* Obliczenie współrzędnych 3D

## Przegląd literatury

Temat implementacji i wykorzystania technik fotogrametrycznych jest w ostatnim czasie bardzo popularny. Z tego powodu, ilość dostępnej literatury w tym obszarze jest znacząca. Zawiera się w niej szereg wartościowych prac i publikacji, zarówno z zakresu budowy i użycia systemów wizyjnych opartych o techniki fotogrametryczne jak również literatury ukierunkowanej ściśle na problematykę SfM, rozpatrującą i analizującą osobno każdy jej element.

Jak napisano na wstępie, powstało dotychczas wiele systemów i aplikacji które wykorzystując techniki fotogrametryczne, pozwalają na tworzenie modeli 3D przedmiotów zarejestrowanych na sekwencji obrazów. Mimo to ilość dostępnych rozwiązań na tym polu jest dużo mniejsza od liczby zastosowań samej fotogrametrii, a ta ciągle rośnie.

Christian Larsen [17] proponuje aplikację opartą o SfM która w sposób półautomatyczny pozwala tworzyć modele 3D budynków, te mogą zostać następnie użyte w celach marketingowych. Przedstawia on podobne jak w tej pracy podejście do rekonstrukcji geometrii badanego obiektu. Implementuje ponadto w swoim programie – z czego zrezygnowano w tej pracy - możliwość teksturowania rekonstruowanych modeli, co było zresztą jednym z głównych założeń jego projektu.

Inne bardzo interesujące podejście przedstawił Jim Braux-Zin [18], który użył SfM do zastosowań w rozszerzonej rzeczywistości. W przeciwieństwie do tej pracy, gdzie w celu rekonstrukcji geometrii wykorzystano metodę mocnych cech, czego wynikiem jest dokładna ale rzadka chmura punktów, on użył w swojej aplikacji podobnie jak Adam Dominec [19], mapy głębokości która pozwala otrzymać mniej dokładną ale dużo gęstszą chmurę. Co więcej użył on do rejestracji obrazu dwóch kamer, dzięki temu zamiast analizować sekwencję obrazów, jak ma to miejsce w tej pracy, analizuje on stereopary wykonane przez kamery których położenie względem siebie jest stałe i znane. Zwiększa to znacząco jakość otrzymanych wyników.

Wiele pozycji podejmujących zbliżoną problematykę, obejmuje jedynie część zagadnień poruszanych w tej pracy. Przykładem jest praca Hansa Lango i Mortena Tyldena zajmująca się problemem rekonstrukcji powierzchni na podstawie danych otrzymanych podczas skanowania laserowego. Badania ich pracy są ukierunkowane głównie na samą rekonstrukcję powierzchni podobnie jak ma to miejsce w przypadku pracy Jukki Lankinena [20] który za główny cel obrał budowę aplikacji do tworzenia modelu 3D na podstawie przygotowanej chmury punków nie poruszając problemu jej uzyskania.

## Cel pracy

Celem niniejszej pracy była budowa oprogramowania komputerowego do cyfrowej rekonstrukcji powierzchni badanego przedmiotu na podstawie serii przygotowanych wcześniej zdjęć.

# Wprowadzenie teoretyczne

Realizacja postawionego w pracy celu, wymaga budowy aplikacji która umożliwi rekonstrukcję powierzchni badanego przedmiotu na podstawie serii wykonanych wcześniej zdjęć oraz wizualizację otrzymanych wyników. Z racji tego, że jest to zadanie złożone, zostało podzielone na cztery podzadania. W podobny sposób prezentuje się struktura tego rozdziału.

Rysunek 2.1.1.1 Schemat zaimplementowanego w pracy procesu cyfrowej rekonstrukcji powierzchni   
z wykorzystaniem techniki Structure from Motion.

Pierwszy rozdział, poświęcony jest tematowi akwizycji obrazu i kilku z nim związanych. Wprowadzono w nim na początku matematyczny model kamery, który stopniowo udoskonalano aby opisywał wszystkie najważniejsze zjawiska mające wpływ na proces rzutowania obiektów 3D na płaszczyźnie obrazu. Następnie przedstawiono problem wprowadzanych przez układy optyczne zniekształceń obrazu, w tym dystorsji tangencjalnej i radialnej. Na koniec opisano proces kalibracji kamery, przeprowadzanej w celu niwelacji w/w zniekształceń.

Drugi rozdział prezentuje użytą w pracy, technikę SfM*.*

Rozdział trzeci zajmuje się tematyką rekonstrukcji powierzchni. Przedstawiono w nim przykładowy potok przetwarzania wraz z ilustracjami, które uwidaczniają wpływ każdego z jego etapów na ostateczny wynik.

Rozdział ostatni zawiera opis procesu wizualizacji grafiki 3D na ekranie monitora.

## Akwizycja obrazu

W technikach fotogrametrycznych jedynym źródłem informacji o badanych obiektach i ich otoczeniu są zdjęcia. Te powinna cechować odpowiednia ostrość i rozdzielczość. Są to parametry które znacząco wpływają na ilość i jakość znalezionych na obrazach cech, co może zwiększyć szanse na ich poprawne dopasowanie. Otrzymane zdjęcia nie powinny posiadać zniekształceń których źródłem może być wewnętrzna budowa aparatu. Kolejną ważną kwestią jest zmiana perspektywy pomiędzy obrazami. Zbyt mała może skutkować błędnymi wynikami, a w rezultacie sprawić, że otrzymana chmura punktów będzie całkowicie błędna. Z drugiej strony zbyt duża jej zmiana może nie pozwolić na znalezienie odpowiadających sobie punktów charakterystycznych.

Przed przejściem do konkretnych podrozdziałów należy wyjaśnić pewną istotną z punktu widzenia nomenklatury kwestię. Terminy: aparat i kamera, są w tej pracy używane zamiennie i odnoszą się do urządzenia rejestrującego obraz, zbudowanego z układu optycznego w postaci pojedynczej cienkiej soczewki, oraz matrycy CCD/CMOS, która stanowi płaszczyznę obrazu.

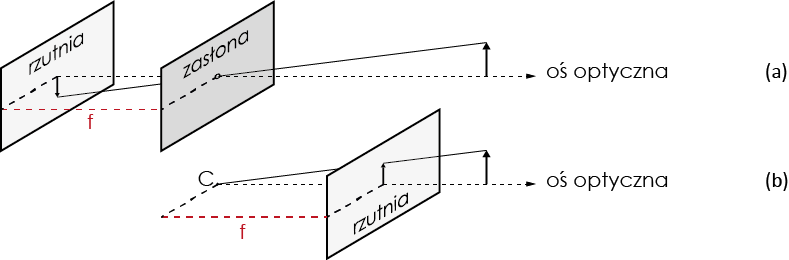
### Model matematyczny kamery

Współczesne aparaty cyfrowe, kamery, czy telefony posiadają często stosunkowo złożone systemy optyczne. Wszystkie te urządzenia znajdują jednocześnie zastosowanie w różnego rodzaju systemach wizyjnych jako elementy akwizycji obrazu. Dla takich zastosowań konieczne jest stworzenie matematycznego modelu, który pozwoli opisać związek pomiędzy obiektem 3D, a jego obrazem zarejestrowanym przez kamerę. Naturalnie im bardziej złożona jest jej wewnętrzna budowa, tym trudniej taki model zbudować. W praktyce więc, często korzysta się z przybliżenia w postaci prostego modelu kamery otworkowej, który następnie w miarę potrzeb uzupełnia się i rozszerza.

Ogólna koncepcja

Kamera otworkowa przedstawiana jest zwykle jako zamknięte pudełko o czarnym wnętrzu i mały otworze na jednej ze ścian. Promienie świetlne, które przez przejdą przez otwór w zasłonie, tworzą wewnątrz pudełka odwrócony obraz obserwowanego przedmiotu. W podobny sposób przedstawia to rysunek 1.1.1.1 a.

W ramach matematycznych uproszczeń, proces rzutowania można zmodyfikować tak, aby obraz przedmiotu nie był odwrócony. W tym celu przenosi się płaszczyznę obrazu między obserwowany obiekt a zasłonę, zachowując jednocześnie odległość f między aperturą/środkiem kamery C i rzutnią.

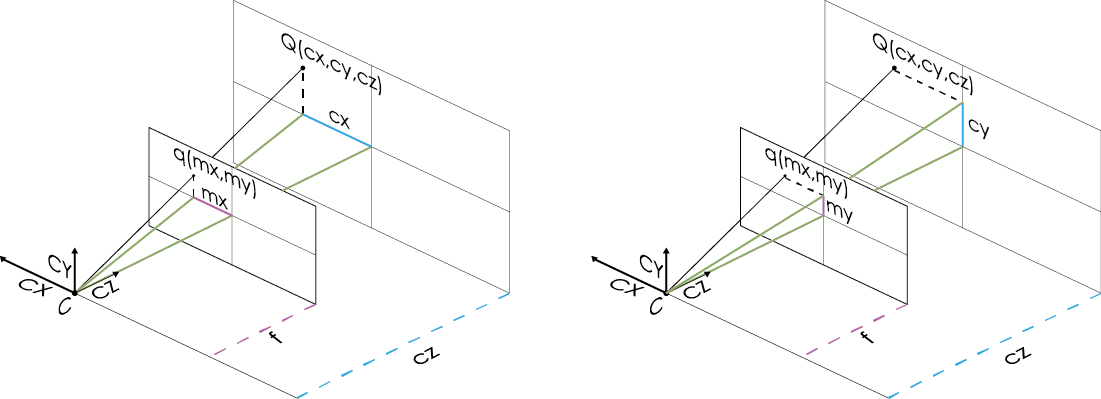


Rysunek 1.1.1.1 Model kamery otworkowej. Promienie przechodzące przez otwór w zasłonie tworzą   
obraz na płaszczyźnie obrazu. f – ogniskowa układu. C - środek projekcji/kamery.

Rzutowanie perspektywiczne

Z rysunku 1.1.1.1 b widać, że podczas rejestracji obrazu obiektów trójwymiarowych dochodzi do mapowania punktów z przestrzeni 3D do 2D. W przypadku modelu kamery otworkowej rzutowanie odbywa się pomiędzy:

* Układem współrzędnych kamery UWK, w początku którego leży środek kamery C, i którego oś CZ pokrywa się z osią optyczną.
* Układem współrzędnych metrycznych obrazu UWMO, związanym z położeniem punktu q, będącym obrazem rzutu punktu przestrzennego Q na płaszczyźnie obrazu.



Rysunek 1.1.1.2 Rzutowanie perspektywiczne. Obraz rzutu punktu przestrzennego Q powstaje na płaszczyźnie obrazu w miejscu q, gdzie promień rzutujący, łączący punkt Q i środek projekcji C przebija rzutnię.

Korzystając z podobieństwa trójkątów, łatwo można wyrazić związek pomiędzy położeniem punktu q na rzutni, a punktem Q w przestrzeni 3D.

( 1.1.1.1.1 )

( 2.1.1.1.2 )

proces projekcji przebiega więc w sposób

( 3.1.1.1.3 )

to samo wyrażenie we współrzędnych jednorodnych wygląda następująco

( 4.1.1.1.4 )

Przechodząc dalej do postaci macierzowej, transformację tę można przedstawić w postaci

( 5.1.1.1.5 )

Przyjmując, że pierwsza macierz jest macierzą projekcji P, związek pomiędzy punktem przestrzennym Q i odpowiadającym mu punktem na obrazie q wyraża się wzorem

( 6.1.1.1.6 )

Macierz P ma wymiary 3x4 i jest złożeniem wszystkich transformacji jakie konieczne są do przejścia pomiędzy układami 3D i 2D.

W trakcie omawiania procesu rzutowania perspektywicznego w tym akapicie, skorzystano ze współrzędnych jednorodnych oraz przekształceń geometrycznych. Zagadnienia te nie były wcześniej omawiane, a ponieważ ich znajomość jest niezbędna do zrozumienia całego procesu rzutowania, dlatego warto przyjrzeć się im bliżej.

Transformacje i współrzędne jednorodne

W szeroko rozumianej grafice komputerowej, robotyce i nie tylko, przekształcenia w przestrzeni i zaliczane są do elementarnych operacji. Do najważniejszych wśród nich należą.:

* Przesunięcie

( 7.1.1.1.7 )

* Obrót

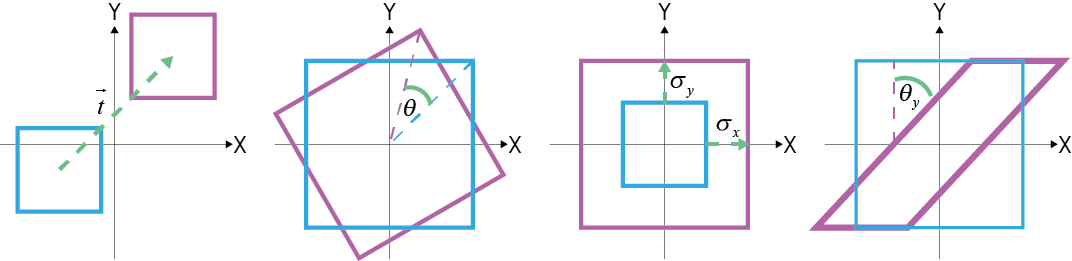
( 8.1.1.1.8 )

* Skalowanie

( 9.1.1.1.9 )

* Pochylenie

( 10.1.1.1.10 )



Rysunek 1.1.1.3 Podstawowe przekształcenia geometryczne na płaszczyźnie.   
Kolejno od lewej: przesunięcie, obrót, skalowanie, pochylenie.

Podczas pracy z przekształceniami geometrycznymi korzysta się głównie z rachunku macierzowego. Jest to podyktowane prostotą zapisu i efektywnością obliczeń, zwłaszcza w obszarze grafiki komputerowej. Położenie punktu można wtedy przedstawić jako wektor liczb. We współrzędnych kartezjańskich jest równe wymiarowi przestrzeni.

W takim układzie współrzędnych, pokazane na rys. 1.1.1.4 przekształcenia można zapisać w formie macierzy transformacji. Wynikiem zastosowania przekształcenia obrotu, skalowania lub pochylenia jest więc iloraz macierzy transformacji i wektora położenia punktu. W przypadku przesunięcia nie ma innej możliwości zapisu jak tylko suma dwóch wektorów: położenia i przesunięcia. Ta niejednolita notacja uniemożliwia wykonanie złożenia np. translacji i rotacji, czyli zapisania tych dwóch przekształceń w postaci jednej macierzy.

Rozwiązaniem tego problemu jest użycie współrzędnych jednorodnych, w których punkt   
opisany przez współrzędnych kartezjańskich reprezentowany jest przez współrzędnych rzutowych/jednorodnych.

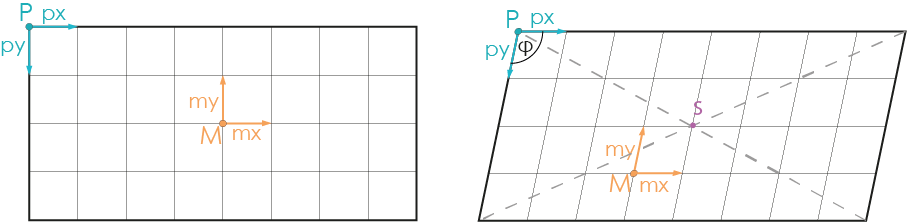
( 11.1.1.1.11 )

Taki sposób zapisu pozwala wyrazić każdą transformację w przestrzeni w postaci macierzy , a w przestrzeni w postaci macierzy . W konsekwencji możliwe jest złożenie dowolnej liczby przekształceń zapisując je jako macierz będącą iloczynem składowych macierzy transformacji.

Układy współrzędnych

Na początku tego rozdziału, omawiając rzutowanie perspektywiczne wprowadzono układ współrzędnych kamery UWK, w którym zdefiniowano położenie kamery jako początek tego układu, a także wyrażono współrzędne punktu przestrzennego Q. Naturalnie, kamera i obiekty które są przez nią rejestrowane najczęściej nie są ze sobą połączone i nie poruszają się razem. Wynika stąd, że punkt Q nie należy do układu UWK, niemniej może – dzięki przedstawionym wcześniej transformacjom – zostać wyrażone w tym układzie jego położenie. Zarówno układ UWK jak i obiekty które są obserwowane przez kamerę definiuje się w globalnym układzie współrzędnych UWG.

Rzutując punkt 3D na płaszczyznę, położenie obrazu jego rzutu można wyrazić za pomocą jednostek metrycznych, tj. układzie UWMO lub za pomocą pikseli – we współrzędnych pikselowych obrazu UWPO. Jest to uzasadnione, ponieważ we współczesnych aparatach, rolę płaszczyzny obrazu pełni najczęściej matryca CCD lub CMOS. W dużym uproszczeniu są to macierze bardzo małych prostokątnych elementów światłoczułych. Zarejestrowany przez nie obraz jest więc gęstą macierzą pojedynczych prostokątów – pikseli.



Rysunek 1.1.1.4 Matryca CMOS/CCD zbudowana z 8x4 prostokątnych elementów światłoczułych. Układ   
współrzędnych UWPO zaczepiony jest w lewym górnym rogu matrycy. Oś optyczna przechodzi przez początek   
układu UWMO, który na rys. a znajduje się w środku płaszczyzny obrazu S, a na rys. b jest względem niego przesunięty.

W idealnym przypadku – rysunek 1.1.1.2. a - matryca CCD/CMOS jest:

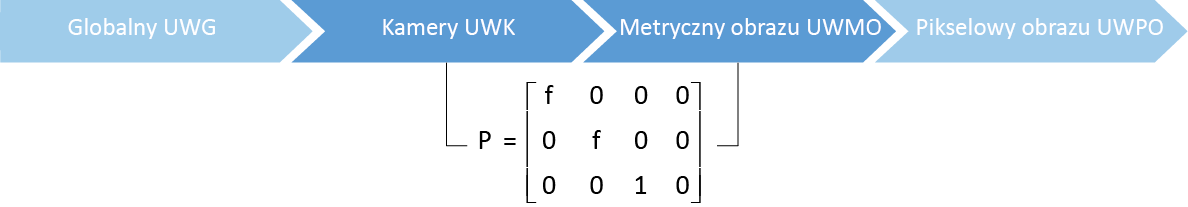
* zbudowana z kwadratowych elementów światłoczułych
* ustawiona prostopadle do osi optycznej, która przechodzi przez jej środek

W rzeczywistości, tak jak ilustruje to rysunek 1.1.1.2 b matryca może:

* posiadać elementy światłoczułe w kształcie równoległoboku - jest to bardzo rzadkie
* nie być ustawiona prostopadle do osi optycznej, co skutkuje pojawieniem się zniekształceń obrazu w rodzaju dystorsji tangencjalnej
* oś optyczna może nie przechodzić przez środek matrycy

Przejście z UWG do UWPO

W poprzedniej części tego rozdziału omówiono cztery układy współrzędnych, jakie towarzyszą w pracy z modelem kamery otworkowej. Przedstawiono również macierz projekcji P która pozwalała przeliczyć współrzędne z układu kamery do układu metrycznego obrazu.



Rysunek 1.1.1.5 Schemat przejścia ze współrzędnych globalnych do współrzędnych pikselowych obrazu.

Docelowo jednak, macierz P powinna wiązać położenie punktu 3D zdefiniowanego w układzie globalnym, z położeniem tego w układzie współrzędnych pikselowych obrazu. W tym celu należy rozszerzyć obecną macierz projekcji o dodatkowe przekształcenia geometryczne które umożliwią transformację pomiędzy UWG, a UWPO.

Transformacja UWMO - UWPO

Chcąc wyrazić położenie punktu na obrazie nie w jednostkach metrycznych jak dotychczas ale w pikselach, konieczna jest transformacja pomiędzy UWMO, a UWPO. Wiąże się to z zastosowaniem:

* skalowania

Matryca aparatu ma swoje wymiary fizyczne, np. 23,6 x 15,7 mm. Jej rozmiar jest również określony poprzez rozdzielczość, np. 5184 x 3456 pikseli. Stosunki tych wielkości, odpowiednio dla pierwszej i drugiej współrzędnej pozwalają określić położenie punktu na obrazie, zarówno w jednostkach metrycznych, jak również w pikselach

( 13.1.1.1.13 )

( 14.1.1.1.14 )

* translacji

Zgodnie z zamieszczonym wcześniej rysunkiem 1.1.1.4 układ współrzędnych UWMO zaczepiony jest w punkcie, gdzie oś optyczna przebija płaszczyznę obrazu. Początek układu UWPO położony jest natomiast w lewym górnym jej rogu. Oba układy są więc przesunięte względem siebie o i .

* pochylenia

Pochylenie, jest w tym przypadku związane z kształtem elementów światłoczułych matrycy, współczynniki pochylenia i zależą od kąta pomiędzy osiami układu UWPO. Zwykle jednak są one prostopadłe względem siebie, więc nie uwzględnia się tej transformacji w obliczeniach.

Uaktualniona macierz projekcji P powinna zatem przybrać postać

( 15.1.1.1.15 )

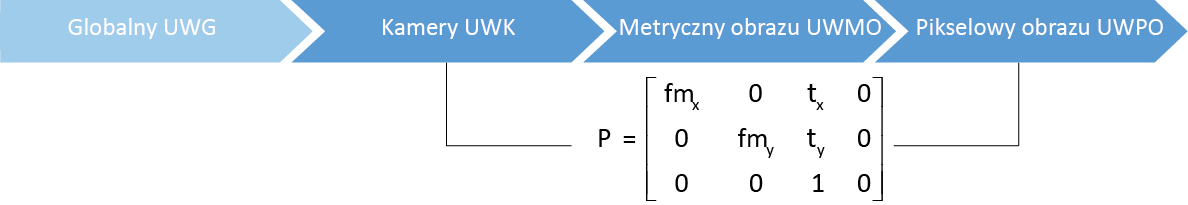
W równoważnej, skróconej formie można zapisać ją jako

( 16.1.1.1.16 )

,gdzie K jest macierzą kalibracji, I macierzą jednostkową 3x3, a 0 zerowym wektorem kolumnowym

( 17.1.1.1.17 )

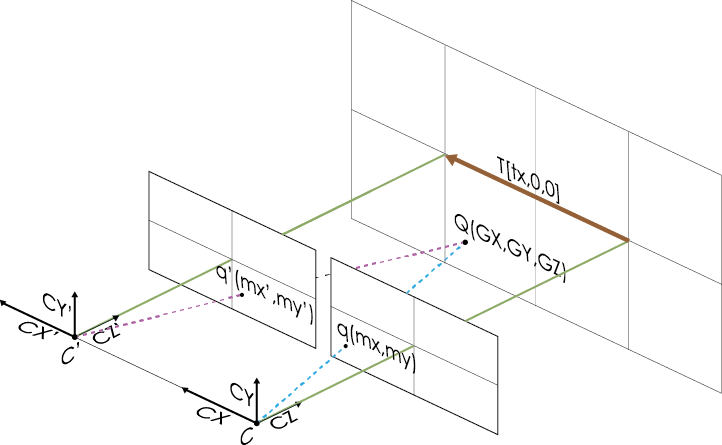
natomiast I to macierz jednostkowa o wymiarach 3x3, a 0 jest zerowym wektorem kolumnowy



Rysunek 1.1.1.6 Transformacja pomiędzy układami współrzędnych UWK i UWPO

Transformacja UWG - UWK

Ostatnim etapem w ramach omawianego dotąd potoku transformacji, jest przejście ze współrzędnych globalnych UWG do układu współrzędnych kamery UWK.



Rysunek 1.1.1.7 Związek pomiędzy dwoma układami współrzędnych

Podążając za rysunkiem, załóżmy, że kamerą C, wykonano zdjęcie punktu Q. Następnie przesunięto kamerę tak, że jej środek znalazł się w punkcie C’ i ponownie wykonano zdjęcie punktu Q. Korzystając z tego, że globalny układ współrzędnych może być dowolnie zdefiniowany w przestrzeni, przyjmijmy więc dla uproszczenia, że jego osie pokrywają się z osiami CX, CY i CZ. W takim przypadku, związek pomiędzy globalnym układem współrzędnych, a układem kamery C’ sprowadza się do określenia jak zmieniła się orientacja i położenie kamery podczas wykonywania pierwszego i drugiego zdjęcia.

Jeżeli za przyjąć współrzędne punktu Q w układzie globalnym, czyli w tym przypadku związanym z kamerą C wykonującą pierwsze zdjęcie, a za jego współrzędne w układzie kamery wykonującej drugie zdjęcie, wtedy związek pomiędzy tymi dwoma układami współrzędnych można wyrazić poprzez wektor przesunięcia i macierz obrotu, a samą transformację z układu UWG do UWK jako

( 18.1.1.1.18 )

,gdzie R to macierzą obrotu określająca orientacje układu UWK względem UWG, natomiast to położenie środka kamery C’ we współrzędnych globalnych.

( 19.1.1.1.19 )

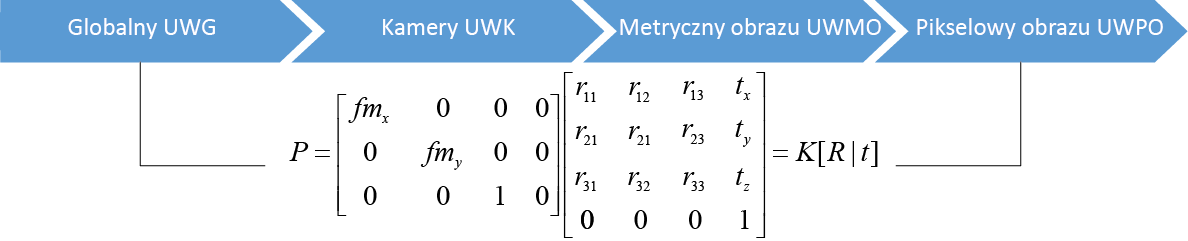
( 20.20 )

Korzystając z rachunku macierzowego, wyrażenie 19.1.1.1.19 można zapisać w postaci

( 21.1.1.1.21 )

Przemnażając przez siebie składowe macierze obrotu i przesunięcia, można uzyskać pojedynczą macierz, która jest złożeniem tych dwóch przekształceń geometrycznych.

( 22.1.1.1.22 )

Ostatecznie więc, transformacja współrzędnych punktu z układu UWG do UWPO przebiega według poniższego schematu.  
  


Rysunek 1.1.1.8 Transformacja z układu współrzędnych UWG do UWPO

### Dystorsja

Jak już wspomniano, matryca aparatu najczęściej nie jest ustawiona równolegle do jej soczewki. W rezultacie powstający na matrycy obraz doznaje zniekształceń. Mowa tutaj o dystorsji tangencjalnej.

Obraz

Matryca CMOS

Soczewka

Klej

Aparat

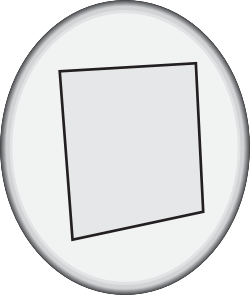
Rysunek 2.1.3 Przykład dystorsji tangencjalnej spowodowanej nierównoległym ustawieniem   
matrycy światłoczułej w stosunku do soczewki. Str.377 [21].

Drugim równie istotnym zniekształceniem obrazu jest dystorsja radialna. Jest ona spowodowana złym wykonaniem układu optycznego, a można ją zaobserwować w postaci chociażby zaokrąglonych krawędzi rejestrowanych przedmiotów. Zniekształcenie to jest tym większe im obraz znajduje się dalej od osi optycznej soczewki. Aktualnie układy optyczne montowane w aparatach i kamerach są znacznie lepsze niż jeszcze kilka lat temu, niemniej jednak ich jakość odbiega znacząco od produktów przeznaczonych ściśle do celów fotogrametrycznych.

Prostokątny   
obiekt

Soczewka

Płaszczyzna obrazu



Rysunek 2.1.2.1 Przykład dystorsji radialnej spowodowany niską   
jakością układu optycznego. Str. 376 [21]

### Kalibracja

Zadaniem kalibracji jest wyznaczenie:

* Macierzy kalibracji
* Macierzy zawierającej informację o wprowadzanej przez układ optyczny dystorsji

Ich znajomość pozwala znaleźć zależności pomiędzy układem współrzędnych kamery, a układem związanym z obserwowanym obiektem

Proces kalibracji polega najczęściej na wykonaniu serii 10-20 zdjęć wzorca kalibracyjnego z różnej perspektywy. Na ich podstawie algorytm wyznacza parametry wewnętrzne kamery i współczynniki dystorsji.

Wszystkie zmienne zawarte w macierzy kalibracji są parametrami wewnętrznymi kamery. Jednym z nich jest odległość ogniskowej. Tu rodzi się problem, ponieważ wykonując serię zdjęć danego obiektu lub nagrywając sekwencję obrazów, często w celu poprawy jakości korzysta się z opcji auto-wyostrzania, która zmienia wartość ogniskowej. Teoretycznie rozwiązaniem tego problemu byłoby

* Przejście do trybu stałej ogniskowej
* Wykonanie kalibracji po każdej zmianie ogniskowej

Oba rozwiązanie są mało praktyczne, dlatego często stosuje się inne podejście. Polega ono na nie wykonywaniu kalibracji z użyciem wzorca kalibracyjnego. Zamiast tego korzysta się z metadanych EXIF które są zwykle dołączane do każdego pliku wykonanego zdjęcia. Dane EXIF zawierają podstawowe informacje o parametrach technicznych aparatu, w tym o jego ogniskowej. Na podstawie tych informacji wyznacza się macierz kalibracji, zakładając przy tym, tak jak w modelu kamery otworkowej, że punkt główny znajduje się dokładnie w połowie wysokości i szerokości matrycy oraz, że wprowadzane przez układ optyczny zniekształcenia są nie wielkie i możne je pominąć.

Kalibracja z użyciem wzorca daje dużo lepsze wyniki, niż przedstawiona przed chwilą metoda tzw. autokalibracji [22]. Nie mniej brak możliwości zmiany ogniskowej jest dużym ograniczeniem. W praktyce każde z tych rozwiązań znajduje zastosowanie.

## Rekonstrukcja geometrii

Celem rekonstrukcji geometrii jest otrzymanie jak największej ilości informacji o kształcie badanego przedmiotu, a te najczęściej są reprezentowane w formie chmury punktów przestrzennych. Proces w którym zostaje ona wyznaczona przy pomocy techniki SfM, jest tematem tego rozdziału.

### Detekcja cech

Punkty charakterystyczne to obszary posiadające unikalne, indywidualne atrybuty które sprawiają, że łatwo jest je znaleźć i porównać z innymi. Punkty te powinna cechować powtarzalność, w praktyce oznacza to, że na dwóch zdjęciach przedstawiających ten sam obiekt ale wykonanych z różnej perspektywy, znalezione punkty kluczowe na pierwszym zdjęciu powinny zostać znalezione również na drugim. Ilość punktów charakterystycznych powinna być odpowiednio duża, a jednocześnie każdy z nich powinien posiadać jak najwięcej indywidualnych atrybutów, które pozwolą na prawidłowe znalezienie odpowiadających im punktów na innym zdjęciu.

W celu lepszego zrozumienia czym są punkty charakterystyczne/kluczowe/cechy warto posłużyć się poniższym rysunkiem. Zawiera on zdjęcie budynku oraz sześć jego fragmentów, które są kandydatami na punkty charakterystyczne.



Rysunek 2.2.1.1 Przykłady punktów charakterystycznych

Na fragmencie *A* nie występuje lokalnie duża różnica jasności, oprócz jednolitego koloru nie posiada on żadnych innych indywidualnych atrybutów które pozwalałyby na szybkie znalezienie jego dokładnego położenia. Fragment B który przedstawia pewien wycinek elewacji również jest trudny do zlokalizowania ponieważ duża część zdjęcia zawiera identyczny regularny wzorzec. Fragmenty C i D wydają się lepszymi kandydatami niż poprzednie A i B, ponieważ zawierają krawędzie. Znacząco ogranicza to obszar poszukiwań, jednak i w tym przypadku dopasowanie ich do konkretnego miejsca na obrazie może nie być proste. Ostatnie dwa E oraz F które prezentują narożniki. Podobnych fragmentów na zdjęciu jest niewiele, dlatego dopasowanie ich do konkretnego miejsca na obrazie jest ułatwione. Są to najlepsi kandydaci z przedstawionego zestawu.

Detekcja punktów charakterystycznych jest jednym z elementarnych zadań podczas analizy i przetwarzania obrazu. Narzędzia które służą do tego celu nazywa się detektorami cech. Istnieje bardzo wiele algorytmów przeznaczonych do wykrywania punktów charakterystycznych. Niektóre z nich odznaczają się większą ilością wykrytych punktów, inne krótszym czasem wyszukiwania, itd. Każdy z nich ma swoje zalety oraz wady i stosowany jest w zależności od postawionych wymagań. Przykładowymi algorytmami tego typu są detektory:

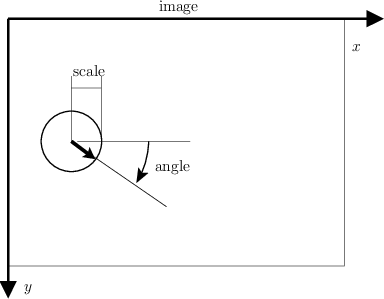
* Harris
* SIFT
* SURF
* FAST
* BRIEF
* ORB

Użytym w pracy został detektor SIFT. Cechuje go niezmienniczość względem rotacji, translacji i skali, a także duży współczynnik poprawnie dopasowanych par [23]. Z tego powodu dalsza część tego rozdziału będzie dotyczyła tylko jego implementacji. Pozostałe algorytmy zostały szczegółowo opisane m.in. w [23] [24].

Pierwszym etapem podczas wykrywania cech z użyciem algorytmu SIFT jest budowa przestrzeni skali. Analizowany obraz poddaje się filtracji Gaussa z różnym parametrem

Punktem kluczowym dla detektora SIFT jest okrągły obszar obrazu posiadający własną orientację, związane z są z nim takie właściwości jak:

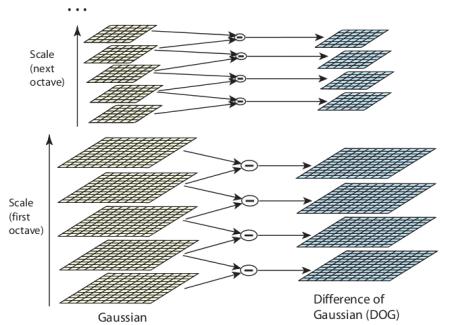
* Współrzędne x i y jego środka
* Skala, będąca promieniem okrągłego obszaru
* Orientacja, tj. kierunek największego gradientu wyrażony w radianach



Rysunek 2.2.1.2 Punkt charakterystyczny SIFT [26]

Pierwszym etapem podczas detekcji cech jest stworzenie piramidy obrazów. Jej konstrukcja polega na zastosowaniu filtru Gaussa z różnym parametrem rozmycia 𝜎 dla obrazu wejściowego w różnej skali, tj. dla różnych rozdzielczości.

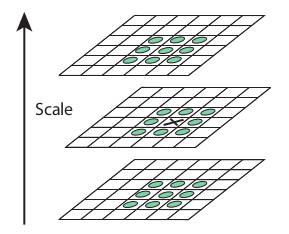
,gdzie I(x,y) - obraz wejściowy  
L(x,y,𝜎) – obraz wyjściowy  
 G(x,y,𝜎) – operator Gaussa =



Rysunek 2.2.1.3 Schemat budowy piramidy obrazów

Piramida obrazów podzielona jest na oktawy. Numer oktawy informuje o rozmiarze obrazu. Im jest on wyższy tym analizowany obraz jest mniejszy. W jednej oktawie znajduje się z kolei kilka warstw, czyli obrazów o różnym poziomie rozmycia.

Następnym krokiem jest obliczenie obrazów różnicowych DoG (ang. Difference of Gaussian). Dla wszystkich obrazów DoG w każdej oktawie z wyjątkiem pierwszego i ostatniego poszukuje się lokalnych ekstremów. Każdy pixel nie brzegowy porównuje się z jego ośmioma sąsiednimi w tej samej skali i dziewięcioma w skali o jeden wyżej i niżej. W przypadku kiedy wartość pixela jest większa lub mniejsza od wartości wszystkich sąsiednich, punkt zostaje zakwalifikowany jako prawdopodobny punkt charakterystyczny.

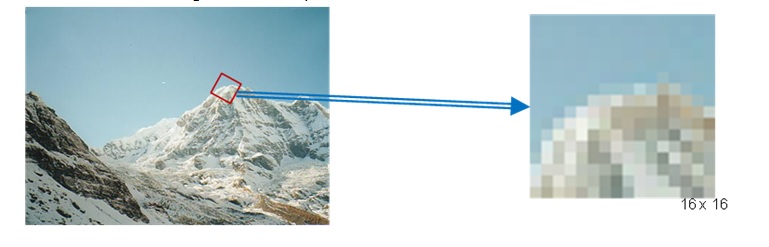


Na koniec usuwa się te które mogą być niestabilne, tj. leżą na krawędziach lub nie spełniają warunku progu kontrastu

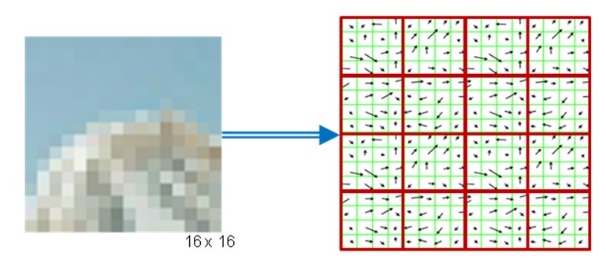
### Obliczenie deskryptorów

Podczas detekcji cech z użyciem detektora SIFT wyznaczana jest ich lokalizacja na zdjęciu, orientacja i skala. Nie jest to wystarczająca ilość informacji aby dokonać ich porównania i znalezienia odpowiadających sobie par. Dlatego dla każdego punktu oblicza się deskryptor który jest 128-wymiarowym wektorem opisujący punkt i jego otoczenie. Ogólna zasada jego obliczenia polega na:

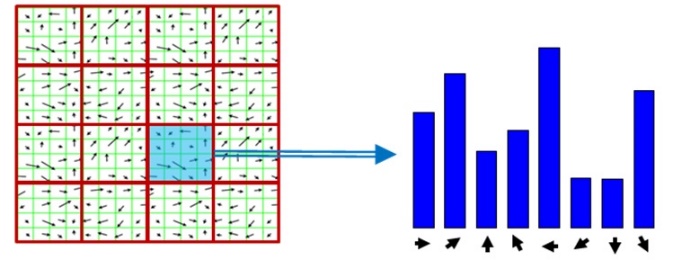
1. Analizie obszaru o wymiarach 16x16 pixeli, zwierającego dany punkt charakterystyczny



1. Obliczeniu gradientu dla każdego pixela w tym obszarze
2. Podzieleniu analizowanego regionu na 16 bloków o wymiarach 4x4 pixele



1. Obliczeniu dla każdego bloku histogramu gradientu w 8 możliwych kierunkach



1. Połączeniu powyższych histogramów i otrzymaniu z nich 128 (4x4x8) wymiarowego wektora.



### Dopasowanie punktów charakterystycznych

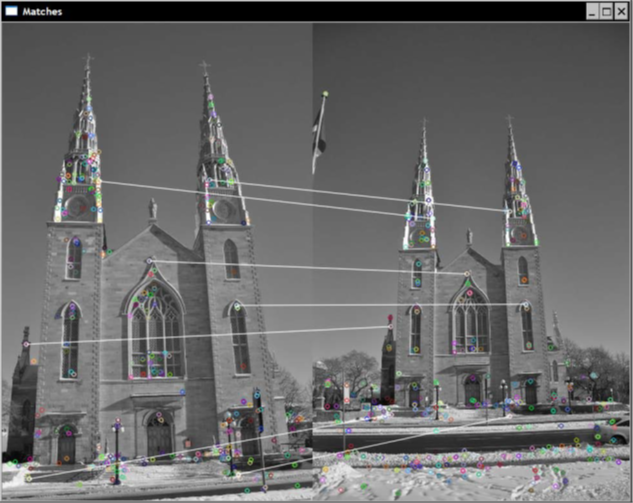
Z szeregu różnych implementacji etapu dopasowywania, często wykorzystywanym jest model w którym początkowo punkty dopasowuje się metodą każdy z każdym, a następnie znalezione pary poddaje się testom. Użycie testów ma na celu usunięcie jak największej ilości złych dopasowań..

W przypadku omawianego wcześniej deskryptora SIFT, wstępne dopasowanie punktów polega na obliczeniu odległości pomiędzy deskryptorem z i-tego i z j-tego zdjęcia. Im mniejsza odległość tym większe prawdopodobieństwo, że opisują one ten sam punkt charakterystyczny. Każdy z deskryptorów to 128 wymiarowy wektor, dlatego aby wyznaczyć odległości między nimi można użyć normy euklidesowej , która jest pierwiastkiem z sumy kwadratów współrzędnych. Tak więc, dla dwóch wektorów i ich norma euklidesowa wynosi

(3.5.1)

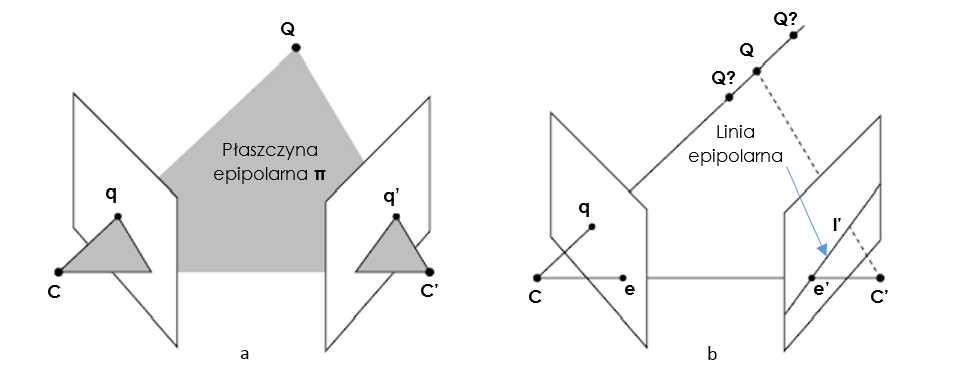
,gdzie *n* dla deskryptora SIFT wynosi 128.

Jednym z testów sprawdzających czy punkty zostały poprawnie dopasowane jest test symetrii. Polega on na znalezieniu dla każdego punktu charakterystycznego na pierwszym zdjęciu jego najlepszego odpowiednika na drugim, następnie postępując analogicznie, dla każdego punktu na drugim szuka się odpowiednika na pierwszym. Na koniec sprawdza się czy w obu przypadkach znalezione dopasowania są identyczne.



Kolejnym, niezwykle ważnym testem jest weryfikacja geometryczna. Do zrozumienia jej działania niezbędna jest znajomość geometrii epipolarnej. Geometria epipolarna zawiera matematyczny opis zależności pomiędzy dwoma różnymi widokami wybranego obiektu. Pełni ona niezwykle ważną rolę w budowie systemów wizyjnych, stąd jej opis będzie pojawiał się również w następnych rozdziałach.

Ogólny sens kryjący się w geometrii epipolarnej łatwo zrozumieć na przykładzie pojedynczego punktu 3D, obserwowanego przez kamerę z dwóch różnych pozycji.



Rysunek 2.2.3.1 Zastosowanie geometrii epipolarnej. Obraz punktu przestrzennego Q powstaje w miejscu przecięcia promienia rzutującego z rzutnią. Są to odpowiednio punkty q i q’ dla pierwszej i drugiej kamery, których środki oznaczono jako C i C’. Wszystkie te punkty leżą na wspólnej płaszczyźnie π., która przecina płaszczyzny obrazu obu kamer i tworzy na nich linie epipolarne l i l’.

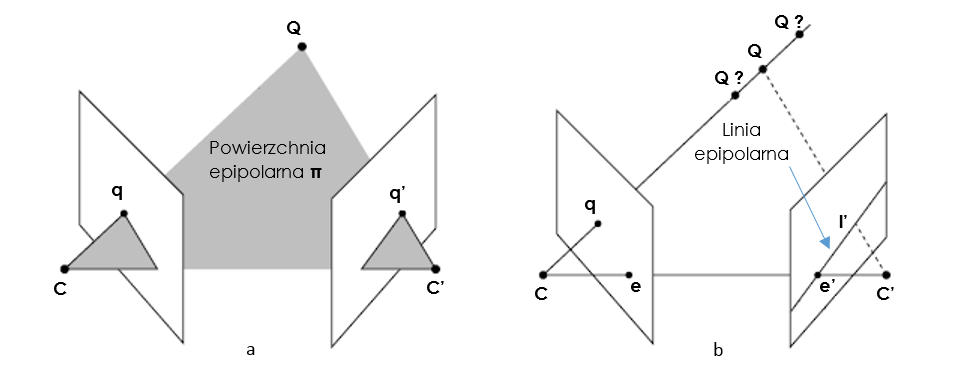
Wykonując zdjęcie przedmiotu za pomocą kamery tracimy informację o jego odległości względem niej. Jest to efekt rzutowania perspektywicznego i związanego z nim mapowania 3D-2D. Sytuacja ulega zmianie kiedy obraz przedmiotu zostaje uchwycony z dwóch różnych perspektyw. Ilustruje to rys 2.2.3.1 na którym punkt przestrzenny Q obserwowany jest przez dwie przesunięte i obrócone względem siebie kamery.

Zgodnie z przykładem a, punkt q powstaje w miejscu w którym promień rzutujący przebija rzutnię i przechodzi przez środek kamery C. W oparciu o punkty q i C, można wyprowadzić równanie prostej w przestrzeni do której powinien należeć również punkt Q tak jak przedstawia to przykład b. Z pozycji drugiej kamery której środek oznaczono jako C’ skonstruowana prosta będzie widoczna w postaci linii epipolarnej. Dysponując wiedzą na temat wzajemnej orientacji i położeniu obu kamer możliwe jest obliczenie pozycji linii l’. Zgodnie z tym co już zaznaczono, możliwe położenie punktu Q ogranicza się do obszaru linii epipolarnej. Tak więc, sprawdzając poprawność dopasowania par punktów charakterystycznych weryfikuje się czy punkty te spełniają powyższe ograniczenie epipolarne.

### Wyznaczenie macierzy projekcji

W poprzednich akapitach omówiono proces detekcji i dopasowania punktów charakterystycznych w oparciu o serię zdjęć wybranego obiektu, wykonanych z różnej perspektywy. Kolejnym krokiem w drodze do rekonstrukcji geometrii z użyciem SfM jest wyznaczenie macierzy projekcji P.

dla każdego z nich. Zgodnie z równaniem (2.1.3), nie jest możliwe wyznaczenie współrzędnych przestrzennych.



W celu wprowadzenia matematycznego opisu rzutowaniaGeometryczne zależności jakie powstają przy rzutowaniu punktu 3D na dwie różniące się położeniem i orientacją płaszczyzny obrazu, opisuje się za pomocą tzw. geometrii epipolarnej.

### Triangulacja

Zadaniem triangulacji jest wyznaczenie współrzędnych przestrzennych punktów na podstawie co najmniej dwóch ich rzutów. Niezbędna w tym celu jest znajomość geometrii kamery, tj. macierzy projekcji dla wszystkich rzutów.

Teoretycznie jest to zadanie trywialne. Obraz punktu przestrzennego P(x,y,z) powstaje na rzutni w miejscu p(x,y) w którym prosta łącząca punkt P i środek projekcji O przebija płaszczyznę obrazu. Tak więc z każdym punktem p i P związany jest promień rzutujący r. Teoretycznie dla każdej pary poprawnie dopasowanych punktów charakterystycznych ich promienie, powinny przecinać się dokładnie w jednym punkcie, punkcie P.

W praktyce jest to wysoce niemożliwe z wielu powodów. Położenie każdego punktu p ma charakter dyskretny, a nie ciągły. Wynika to z faktu, że płaszczyzną obrazu w aparacie jak już wspomniano wcześniej jest macierz elementów światłoczułych o skończonych prostokątnych wymiarach. Ponadto do obliczeń zostają wprowadzone błędy wynikłe z wielu zaokrągleń, błędnie określonych położeń punktów charakterystycznych czy powstałych w wyniku dystorsji.

Promienie rzutujące przetną się, jeżeli układ spełnia kryteria geometrii epipolarnej. W tym celu dodaje się narzędzia (ang. bundle adjustment) których celem jest najlepsze dopasowanie położenia odpowiadających sobie par punktów i położenia punktu przestrzennego P, tak aby umożliwić przeprowadzenie poprawnej triangulacji.

Do metod triangulacji, najczęściej stosowanych należą algorytmy:

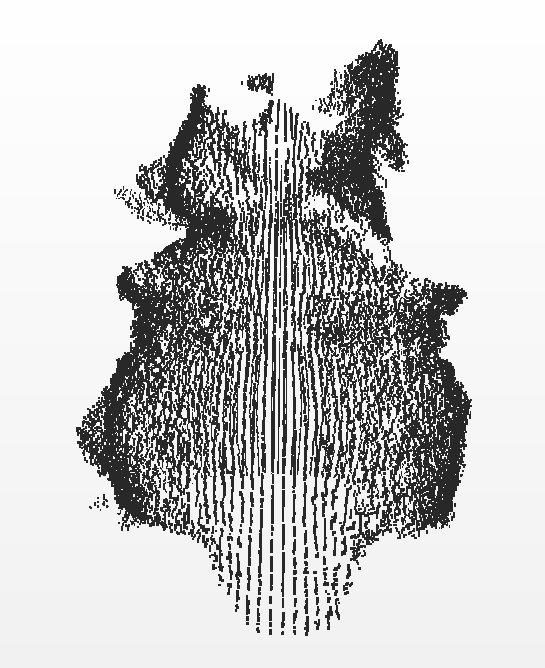
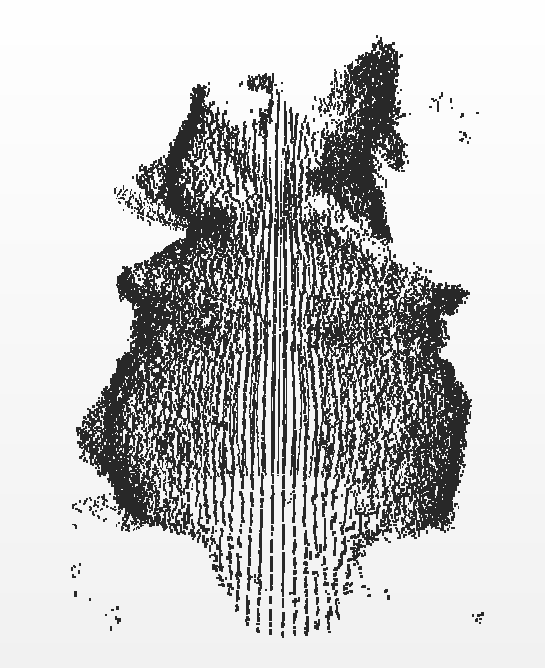
* średniego punktu
* DLT
* Z użyciem macierzy esencjonalnej
* Optymalnej triangulacji

## Rekonstrukcja powierzchni

Otrzymana w wyniku rekonstrukcji geometrii chmura punktów może posłużyć do rekonstrukcji powierzchni, która następnie zostanie zwizualizowana. Aby jednak było to możliwe chmurę należy przedtem poddać analizie i przetworzeniu. Zaimplementowany w tej pracy potok składa się sześciu bloków. Każdy z nich został opisany w osobnym podrozdziale. Ta część pracy została napisana głównie w oparciu o dokumentację techniczną biblioteki CGAL

### Usunięcie punktów odstających

W zależności od użytych narzędzi i technik otrzymana podczas rekonstrukcji geometrii chmura punktów może być bardziej lub mniej spójna. Często zawiera punkty nie związane z rekonstruowanym obiektem, a są one zwykle wynikiem wcześniejszych błędnych obliczeń. Takie punkty należy usunąć ponieważ mogą mieć bardzo negatywny wpływ na jakość zrekonstruowanej powierzchni. Korzysta się w tym celu z algorytmu, który oblicza odległość każdego punktu względem k-najbliższych sąsiednich punktów, a następnie usuwa zadany procent punktów najbardziej odstających punktów.



b

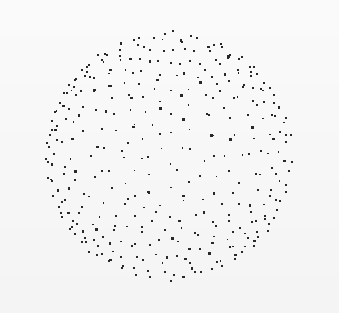
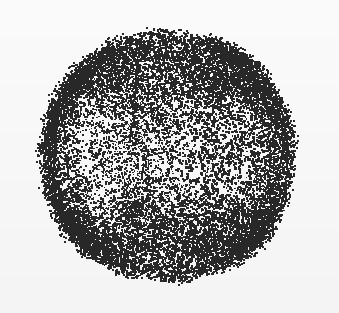
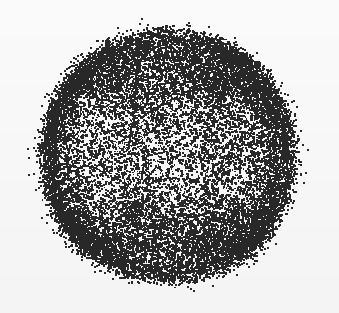
a

Rysunek 2.3.1.1 Przykład użycia algorytmu do usuwania odstających punktów.   
Chmura a została otrzymana podczas skanowania laserowego, b to ta sama chmura po oczyszczeniu.  
 Opracowanie własne

Dopasowując procent usuwanych punktów, oraz ilość najbliższych sąsiadów można uzyskać odpowiedni efekt. Należy jednak zauważyć, że w przypadku zestawu punktów otrzymanych dzięki SfM i rzadkiej rekonstrukcji, algorytm przez błędnie ustawione parametry, może usuwać punkty w obszarach szczególnie rzadkich, a które mają kluczowe znaczenie z punktu widzenia geometrii obiektu.

### Uproszczenie

Na tym etapie dokonuje się usunięcia tych punktów które nie wnoszą znaczącej informacji o geometrii obiektu, zapewniając jednocześnie możliwie równomierne położenie pozostałych punktów. Pozwala to zmniejszyć ilość dalszych obliczeń i uzyskać lepszej jakości powierzchnię.



a

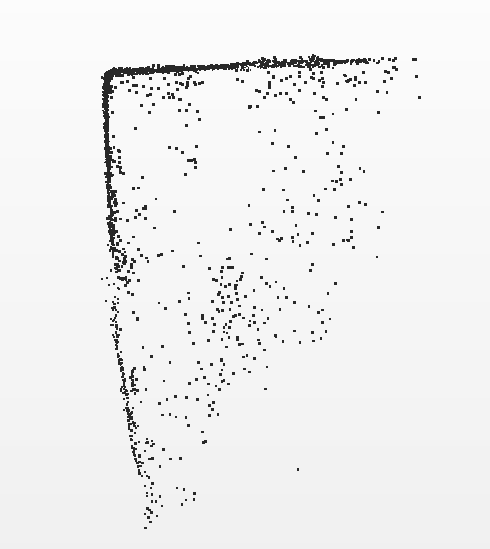
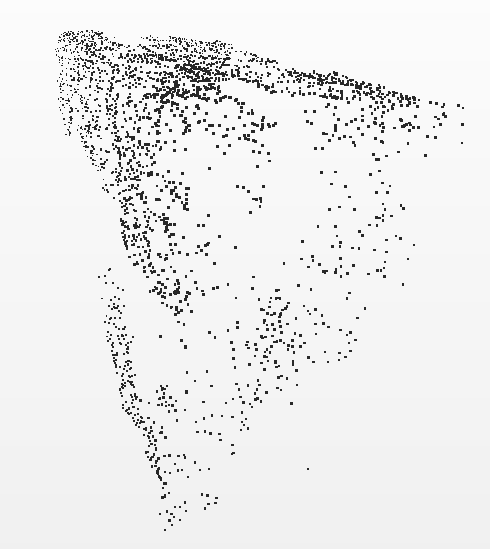
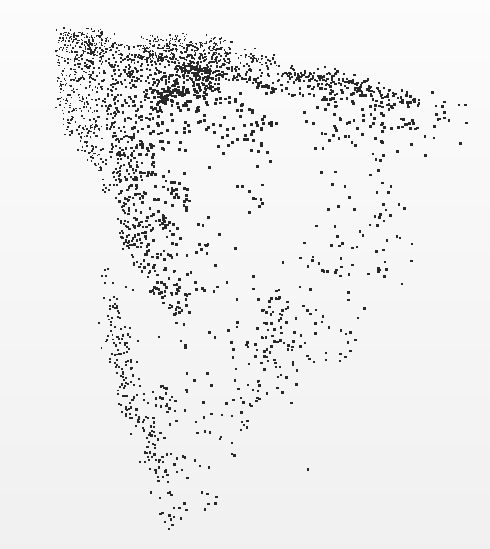
b

c

Rysunek 2.3.2.1 Uproszczenie chmury punktów. Chmura a jest testowym zestawem punktów.  
 Chmura b została oczyszczona z punktów najbardziej odstających, natomiast c ilustruje efekt   
zastosowania algorytmu upraszczającego. Opracowanie własne

### Wygładzenie

Wygładzenie polega na redukcji szumu, tj. próbie ułożenia punktów na gładkiej parametrycznej płaszczyźnie wyznaczonej przez k-najbliższe sąsiednie punkty. Zmniejsza to stopień mogących się pojawić lokalnych zniekształceń, przez co ostateczny wynik prezentuje się wizualnie lepiej.



c

b

a

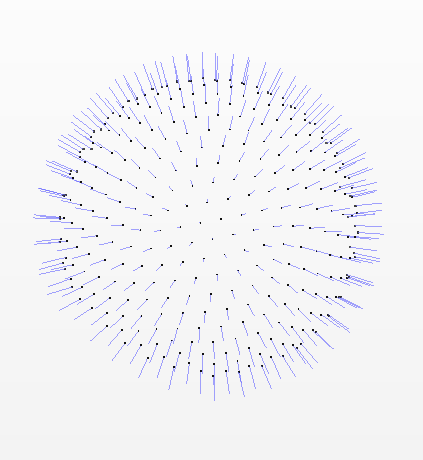
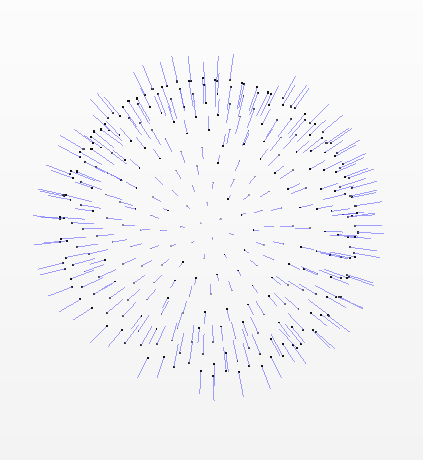
d

Rysunek 2.3.3.1 Wygładzenie punktów. Zestaw a i c stanowią testowy zestaw punktów.   
Zestaw b i d prezentują wynik zastosowania algorytmu wygładzającego. Opracowanie własne

### Wyznaczenie wektorów normalnych

Jednym z możliwych sposobów wyznaczenia wektorów normalnych jest konstrukcja powierzchni utworzonej przez k-najbliższych sąsiednich punktów, a następnie obliczenie wektora prostopadłego do tej powierzchni w punkcie. Istnieją dwa możliwe wektory prostopadłe do powierzchni, dlatego kolejnym krokiem jest ich poprawne zorientowanie. Zostawienie ich w postaci niezorientowanej będzie widoczne podczas wizualizacji jako czarne obszary na zrekonstruowanej powierzchni. Czarny kolor informuje o tym, że światło padające na powierzchnie zostało przez nią całkowicie zaabsorbowane i nawet najmniejsza jego część nie trafia do obserwatora (więcej informacji na ten temat zawiera rozdział 2.4).

W najgorszych wypadku niezorientowane wektory normalne spowodują niezdefiniowane zachowanie algorytmu rekonstrukcji powierzchni. To z kolei będzie widoczne w postaci zmiany topologii powierzchni. Tak więc od wyników tej operacji w dużej mierze zależy postać. Każdy następny algorytm przedstawiony w tym rozdziale wymaga obliczonych i zorientowanych wektorów normalnych.



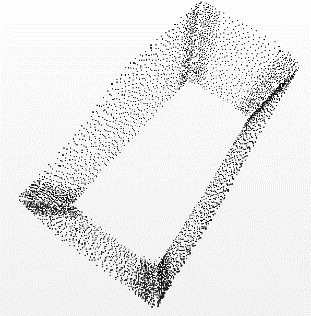
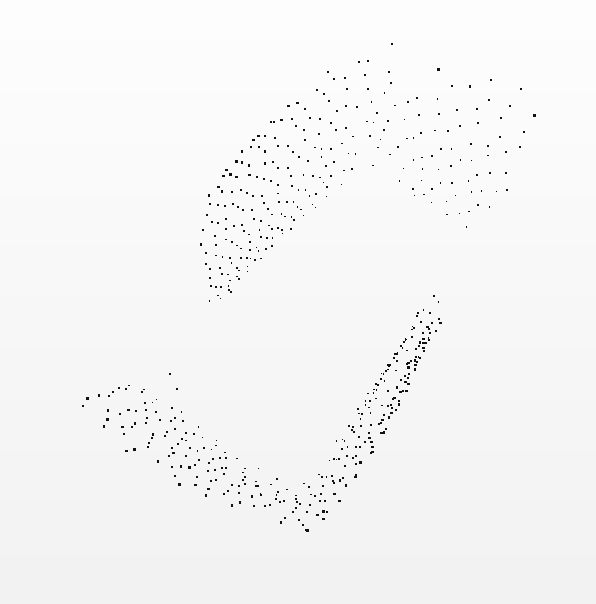
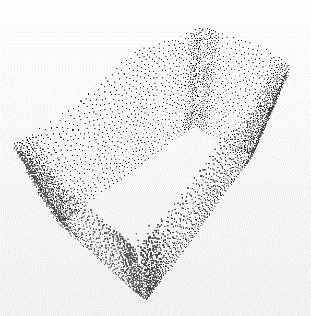
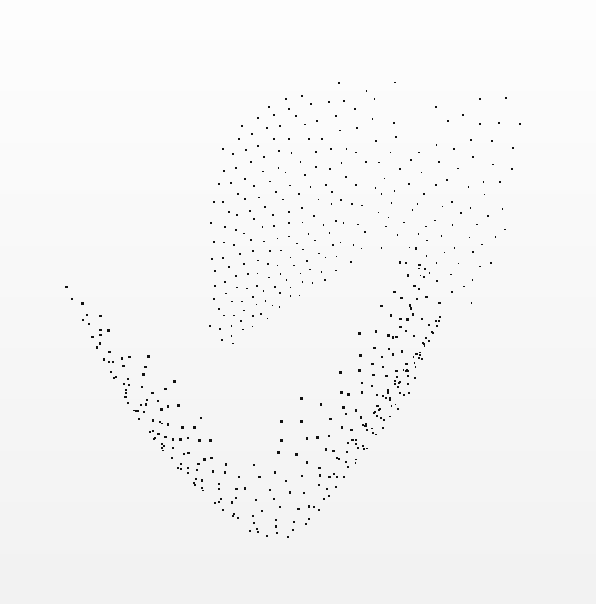
b

a

Rysunek 2.3.4.1 Wyznaczenie wektorów normalnych. Chmura a, z zaznaczonymi   
wektorami normalnymi, chmura b po ich poprawnym zorientowaniu

### Próbkowanie

Podczas tego etapu uzupełnia się miejsca w których występują znaczne ubytki punktów. Taka sytuacja ma często miejsce przy pozyskiwaniu chmury punktów metodą SfM i rzadkiej rekonstrukcji geometrii lub podczas skanowania laserowego. Chmury punktów z takimi ubytkami stanowią problem podczas rekonstrukcji powierzchni, ponieważ zachowanie się algorytmu w tych miejscach może być nieprzewidywalne. Rozwiązaniem tego problemu jest użycie techniki próbkowania powierzchni, opisanej w [25]. Pozwala ona wypełnić do pewnego stopnia brakującą przestrzeń punktami i jednocześnie zwiększyć ich gęstość.



d

c

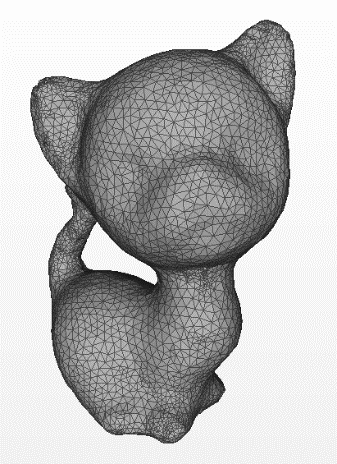
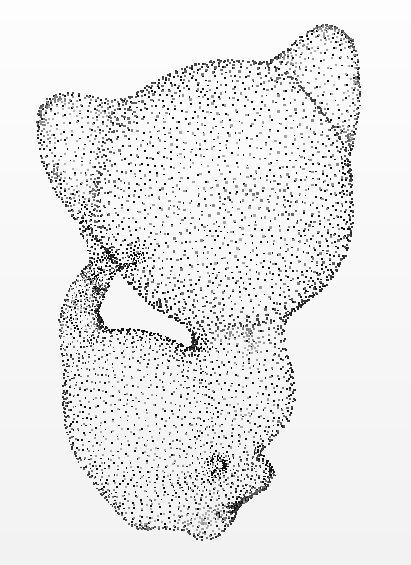
b

a

Rysunek 2.3.5.1 Próbkowanie powierzchni. Zestawy a i c ilustrują przykładową chmurę punktów   
z różnej perspektywy. Zestawy b i d to ta sama chmura po zastosowaniu próbkowania.   
Widoczny jest wzrost gęstości punktów oraz wypełnienie stosunkowo dużych braków

### Rekonstrukcja Poissona

Ostatnim krokiem jest stworzenie siatki wielokątów na podstawie przygotowanej wcześniej chmury punktów. Jednym z algorytmów przeznaczonych do tego celu jest algorytm Poissona.



b

a

Rysunek 2.3.6.1 Rekonstrukcja poissona.   
Przygotowana chmura punktów a i zrekonstruowana powierzchnia b z widoczną siatką wielokątów.

## Wizualizacja

Trudno ocenić poprawność otrzymanych wyników podczas cyfrowej rekonstrukcji bez ich odpowiedniej wizualizacji. Proces przerysowania modelu 3D na ekran monitora jest procesem złożonym, a odpowiedzialność za niego spoczywa obecnie w głównej mierze na karcie graficznej komputera. Do wyświetlenia na ekranie otrzymanej w poprzednim podrozdziale siatki wielokątów niezbędne są narzędzia które na to pozwolą. Jednym z nich jest biblioteka programistyczna OpenGL.   
Z powodu obszerności i złożoności poruszanego tutaj tematu, omówiono jedynie najważniejsze elementy potoku przetwarzania, który towarzyszy renderowaniu obiektów 3D.

Często nawet najbardziej skomplikowane modele trójwymiarowe są w zbudowane z gęstej siatki połączonych ze sobą trójkątów. Przykładem tego jest rys 4.5.1. W zależności od formatu w jakim siatka została zapisana oraz ilością danych jakimi dysponujemy na jej temat, może mieć różną strukturę wewnętrzną, bardziej lub mniej skomplikowaną. W najprostszym jednak przypadku, zawiera:

* Współrzędne wierzchołków
* Współrzędne wektorów normalnych
* indeksy

Karta graficzna odpowiedzialna za renderowanie grafiki tworzy ją posługując się tzw. prymitywami. Są to m.in.: punkty, linie i trójkąty. Z tego powodu tworzone modele 3D najczęściej są siatką właśnie trójkątów.

W celu zwizualizowania takiej przestrzennej siatki na płaskim ekranie monitora, należy przekazać karcie graficznej niezbędne informacje. Zostały one wymienione na początku tego rozdziału. Są to współrzędne wierzchołków i wektorów normalnych, a także indeksy do wierzchołków. Ponieważ wielokąty są traktowane jako fragmenty powierzchni zorientowanej, kolejność w jakiej podaje się wierzchołki ma ogromne znaczenie. Można to zrobić zgodnie z ruchem wskazówek zegara lub odwrotnie. Ważne jest jednak to, aby wszystkie zostały podane w ten sam sposób. Dzięki temu jedna ze stron trójkąta, ta niewidoczna nie będzie wyświetlana, co znacząco zmniejszy ilość obliczeń.

Jak łatwo zauważyć każdy z wierzchołków trójkąta należy jednocześnie do kilku innych. Dlatego podczas podawania współrzędnych wierzchołków posługuje się tzw. indeksami, czyli odnośnikami do nich. Jest to jedna z wielu stosowanych optymalizacji.

Kiedy karta graficzna posiada już dane dotyczące wszystkich trójkątów należących do siatki może zacząć obliczenia powierzchni, koloru lub tekstury i przekształceń geometrycznych. Wszystkie te obliczenia składają się na potok przetwarzania.

Potok przetwarzania jest odpowiedzialny za wszelkie niezbędne obliczenia potrzebne do przerysowania modeli 3D na ekranie monitora. Przedstawiony poniżej schemat zawiera następujące po sobie kroki których celem jest zamiana wierzchołków na postać grafiki rastrowej czyli pixeli. W takiej formie grafika ostatecznie jest wyświetlana na ekranie monitora.

Rysunek 2.3.6.1 Potok przetwarzania

### Operacja na wierzchołkach

Wierzchołki oprócz informacji o swoim położeniu posiadają jeszcze kilka innych atrybutów do których należą współrzędne tekstury i wektora normalnego.

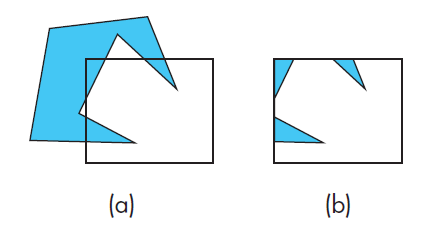
Na początku zostają one poddane 3 podstawowym transformacjom, tj. rotacji, translacji oraz skalowania. Następnie zostają rzutowane równolegle lub perspektywicznie. Dzięki tym operacją możliwe jest obracanie i zmienianie pozycji renderowanego obiektu, a także kamery która go obserwuje.

Następnie dokonywane są obliczenia związane z wyglądem, tj. obliczenie koloru lub współrzędnych tekstury i obliczeń związanych z zaimplementowanym systemem oświetlenia, w którym można uwzględnić zachodzące w rzeczywistości zjawiska odbicia, rozproszenia światła i załamania.

### Przycinanie i budowanie prymitywów

Podczas tego procesu następuje budowanie prymitywów takich jak punkty, linie i trójkąty na podstawie podanych współrzędnych wierzchołków. Następnie dokonuje się obcięcia tych fragmentów które nie znajdują się w polu widzenia kamery lub są ustawione tyłem do kamery.

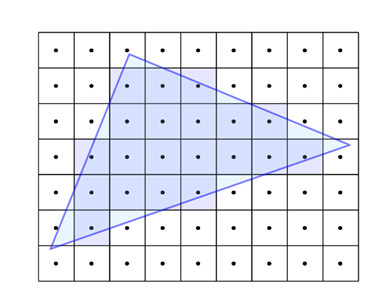
Podobnie jak oko człowieka które posiada kąt widzenia mniejszy niż 90 stopni tak pole widzenia kamery musi zostać ograniczone przez tzw. bryłę obcinającą. Fragmenty które nie leżą w polu widzenia zostają odrzucone i nie są na nich prowadzone dalsze obliczenia.



Rysunek 2.4.2.1 Przycinanie wielokątów.   
(a) przed przycięciem, (b) po operacji przycinania.   
Str. 315 [28]

### Rasteryzacja

Przekształcone prymitywy poddane przycięciu nadal reprezentowane są za pomocą wierzchołków. W tym kroku następuje konwersja wielokątów do postaci pixelowej, czyli fragmentów. Każdy fragment zawiera informacje o kolorze lub teksturze, głębokości itp.



Rysunek 2.4.3.1 Rasteryzacja. Niebieskie linie są krawędziami prymitywu.   
Zamalowane pola oznaczają widoczne fragmenty.

### Operacje na fragmentach

Na tym etapie dokonuje się nakładania tekstury i efektów specjalnych takich jak mgła i rozmycie, usuwania niewidocznych pikseli i innych operacji na pikselach. Na koniec przeprowadza się testy sprawdzające poprawność renderowanej sceny.

# Realizacja

### Schemat działania

### Narzędzia

### Opis aplikacji

### Problemy

### Wyniki badań

# Wnioski

# Bibliografia

|  |  |
| --- | --- |
| [1] | Z. Żelazek i A. Sikora, „Możliwości współczesnej fotogrametrii w aspekcie rekonstrukcji wypadków drogowych,” *Problemy kryminalistyki,* pp. 54-59, 2006. |
| [2] | L. Koźmiński , M. Brzozowska, J. Kościuk i W. Kubisz, „Wykorzystanie możliwości nowoczesnego skanowania 3D w oględzinach miejsca zdarzenia i ch dokumentowania,” *Wiadomości konserwatorskie,* pp. 679-688, 2009. |
| [3] | K. Maksymowicz, M. Kobielarz i T. Jurek, „Skanowanie 3D jako metoda obrazowania złożonych i rozległych relacji przestrzennych dla potrzeb medycyny sądowej i kryminalistyki - ocena przydatności,” *Wiadomości konserwatorskie,* pp. 689-696, 2009. |
| [4] | R. Tokarczyk, Fotogrametryczne pomiary geometrii ciała ludzkiego w zastosowaniu do badania wad postawy, Kraków: AGH Uczelniane Wydawnictwo Naukowo-Dydaktyczne, 2009. |
| [5] | K. Woźniak, A. Moskała, A. Urbanik, P. Kopacz i M. Kłys, „Pośmiertlne badania obrazowe z rekonstrukcją3D: nowa droga rozwoju klasycznej medycyny sądowej?,” *Archiwum medycyny sądowej i kryminologii,* pp. 124-130, 2009. |
| [6] | „Narodowe Centrum Nauki - finansowanie nauki,” [Online]. Available: https://www.ncn.gov.pl/finansowanie-nauki/przyklady-projektow/kubicka. [Data uzyskania dostępu: 23 Luty 2016]. |
| [7] | A. Boroń , A. Rzonca i A. Wróbel, „Metody fotogrametrii cyfrowej i skanowania laserowego w inwentaryzacji zabytków,” *Roczniki Geomatyki,* pp. 129-140, 2007. |
| [8] | E. Bunsch i R. Sitnik, „Kryteria doboru techniki 3D do dokumentacji obiektów dziedzictwa kulturowego,” *Digitalizacja w muzeach,* 02 2014. |
| [9] | A. Wójcicki i S. Herma, „Fotogrametria jako alternatywna metoda modelowania obiektów 3D,” *Modele inżynierii teleinformatyki,* pp. 245-254, 2009. |
| [10] | M. Rychlik, M. Niezgódka, H. Hausa i K. Kotecki, „Opracowanie komputerowej geometrii CAD modelu samolotu do badań flatterowych z zastosowaniem technik reverse engineering,” *Prace instytutu lotnictwa,* pp. 125-134, 2011. |
| [11] | S. Kubica, D. Mielcarek i D. Kurczyk, „Zastosowanie laserowych skanerów 3D do projektowania ergonomicznych stanowisk pracy,” *Pomiary Automatyka Kontrola,* pp. 251-254, 2010. |
| [12] | B. Kraszewski, „Wykorzystanie naziemnego skaningu laserowego do inwentaryzacji pomieszczeń biurowych,” *Archiwum Fotogrametrii, Kartografii i Teledetekcji,* tom 23, pp. 187-196, 2012. |
| [13] | A. Janowski , K. Nagrodzka-Godycka, J. Szluwic i P. Ziółkowski, „Teledetekcyjne badanie konstrukcyjnych elementów żelbetonowych z użyciem naziemnego skaningu laserowego,” *Logistyka,* pp. 856-863, 6 2014. |
| [14] | M. Zygmunt i P. Biłka, „Analiza możliwości zastosowania naziemnego skaningu laserowego w kontroli i ocenie stanu technicznego budowli pietrzących wodę,” *Acta Sci. Pol., Formatio Circumiectus,* pp. 115-124, 1 2014. |
| [15] | Z. Sitek, Fotogrametria inżynierska, Kraków: Wydawnictwa AGH, 1979. |
| [16] | T. Schenk, "Introduction to Photogrammetry," 2005. |
| [17] | C. L. Larsen, "3D Reconstruction of Buildings From Images with Automatic Facade Refinement," 2010. |
| [18] | J. Braux-Zin, "Toward Real-Time Dense 3d Reconstruction usin Stereo Vision," 2011. |
| [19] | A. Dominec, "3D Surface Reconstruction from Video Sequences," 2013. |
| [20] | J. Lankinen, "3D Surface Modelling From Point Clouds," 2009. |
| [21] | G. Bradski and A. Kaebler, Learning OpenCV, O'Relly, 2008. |
| [22] | S. Carlsson, "Point cloud reconstruction from image sequences - Calibrated versus uncalibrated," 2005. |
| [23] | K. M. Tinne Tuytelaars, „Local Invariant Feature Detectors: A Survey,” 2008. |
| [24] | A. Z. Richard Hartley, Multiple View Geometry in Computer Vision, New York: Cambridge University Press, 2004. |
| [25] | H. Huang, S. Wu, M. Gong, D. Cohen-Or, U. Ascher and H. Zhang, "Edge-Aware Point Set Resampling," 2013. |
| [26] | A. Dziubek, „Skaner 3D na bazie strukturalnego oświetlenia,” 2009. |
| [27] | H. Richard, Multiple View Geometry in Computer Vision, Cambridge University Press, 2004. |
| [28] | M. Sonka, V. Hlavac and R. Boyle, Image Processing, Analysis and Machine Vision, Thomson Learning, 2008. |
| [29] | R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2011. |
| [30] | B. Stroustrup, Język C++. Kompedium wiedzy, Helion, 2014. |
| [31] | H. M. Lango i M. Tylden, „Surface Reconstruction and Stereoscopic Video Redndering from Laser Scan Genereated Point Cloud Data,” 2007. |
| [32] | L. S. G. G. Pierre Alliez, „http://doc.cgal.org/latest/Surface\_reconstruction\_points\_3/,” [Online]. |
| [33] | D. S. Edward Angel, Interactive computer graphics, Addison\_wesley, 1012. |
| [34] | „https://en.wikibooks.org/wiki/GLSL\_Programming/Rasterization,” [Online]. |