



**POLITECHNIKA
RZESZOWSKA**
im. IGNACEGO ŁUKASIEWICZA



**WYDZIAŁ
ELEKTROTECHNIKI
I INFORMATYKI**
POLITECHNIKI RZESZOWSKIEJ



Kamery 3D i dane 3D oraz techniki ich pozyskiwania

Struktury danych 3D – mapy głębi i chmury punktów

- ▶ Mapa głębi – obraz monochromatyczny, w którym jasność pikseli określa głębię obiektów.



- ▶ Chmura punktów – reprezentuje zbiór punktów w przestrzeni trójwymiarowej. Jej atrybutami są współrzędne geometryczne x, y, z punktów.

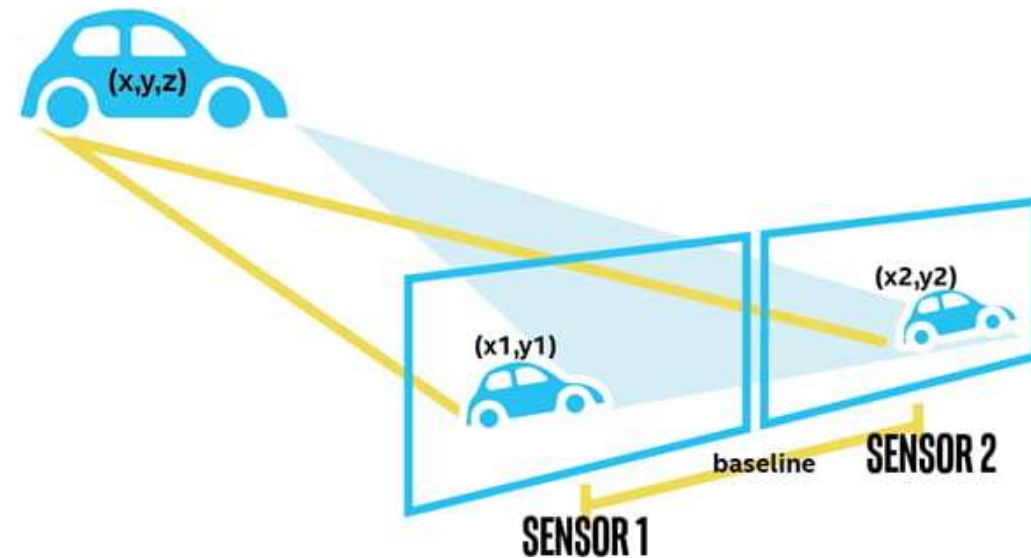


Najważniejsze techniki pozyskiwania danych 3D

1. Stereowizja (ang. stereo vision)
2. Światło strukturalne (ang. structured light)
3. ToF (Time-of-Flight)
4. LIDAR (Light Detection and Ranging)
5. DFD (Depth From Defocus)

Stereowizja (ang. stereo vision)

- ▶ Najstarsza technika, która jest nadal często stosowana.
- ▶ Polega na obserwacji sceny przez układ dwóch kamer kolorowych.
- ▶ Jest wzorowana na widzeniu 3D przez ludzi za pomocą dwojga oczu.
- ▶ Głębina punktów jest obliczana na podstawie dysparycji, czyli różnicy położenia odpowiadających sobie punktów na obrazach z obu kamer.



Źródło: <https://www.intelrealsense.com/beginners-guide-to-depth>

Stereowizja (ang. stereo vision) – wady i zalety

Zalety:

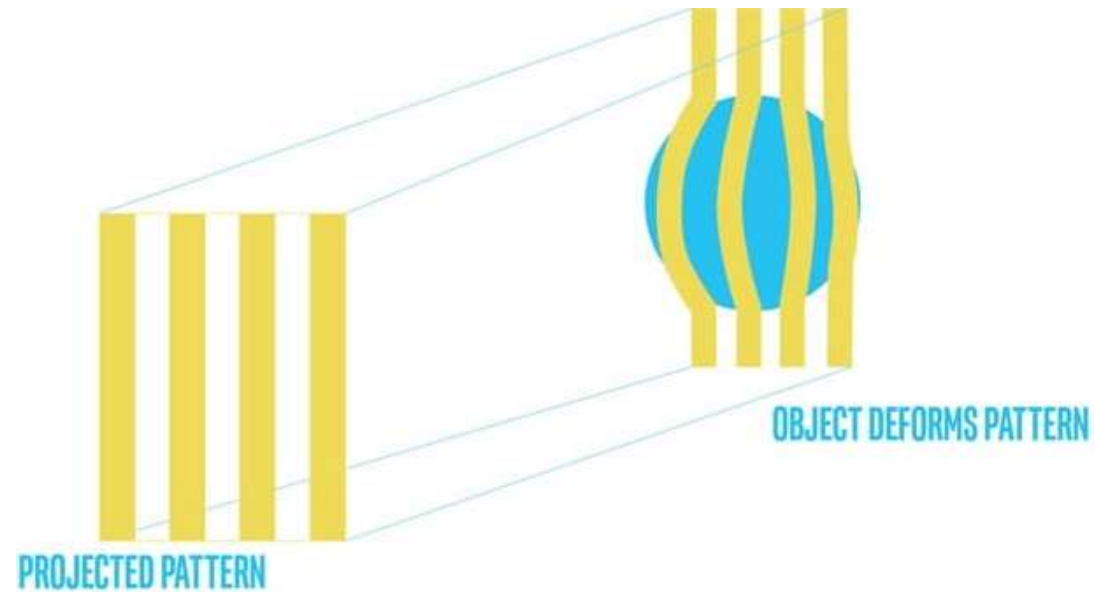
- ▶ Głębina może być dość dokładna, ale to zależy od jakości kamer stereo i algorytmu odnajdywania odpowiadających sobie punktów (tzw. deskryptory lokalne).
- ▶ Możemy uzyskać mapy głębi o wysokiej rozdzielczości, jeśli obie kamery w układzie stereo mają wysoką rozdzielczość.
- ▶ Można ją stosować w pomieszczeniach i w plenerze (ale nie na bardzo duże odległości).

Wady:

- ▶ Działa skutecznie jedynie dla dobrze oteksturowanych i charakterystycznych obiektów oraz przy dobrym oświetleniu.
- ▶ Wymaga dokładnej kalibracji układu kamer stereo.

Światło strukturalne (ang. structured light)

- ▶ Kamera ma wbudowany emiter światła strukturalnego o ustalonym wzorze.
- ▶ Światło jest emitowane w częstotliwości z zakresu podczerwieni, a więc jest niewidoczne dla ludzkiego oka.
- ▶ Głębina punktów jest obliczana na podstawie odkształceń wzorów wyświetlanych na obserwowanych obiektach.



Źródło: <https://www.intelrealsense.com/beginners-guide-to-depth>

Światło strukturalne (ang. structured light) – wady i zalety

Zalety:

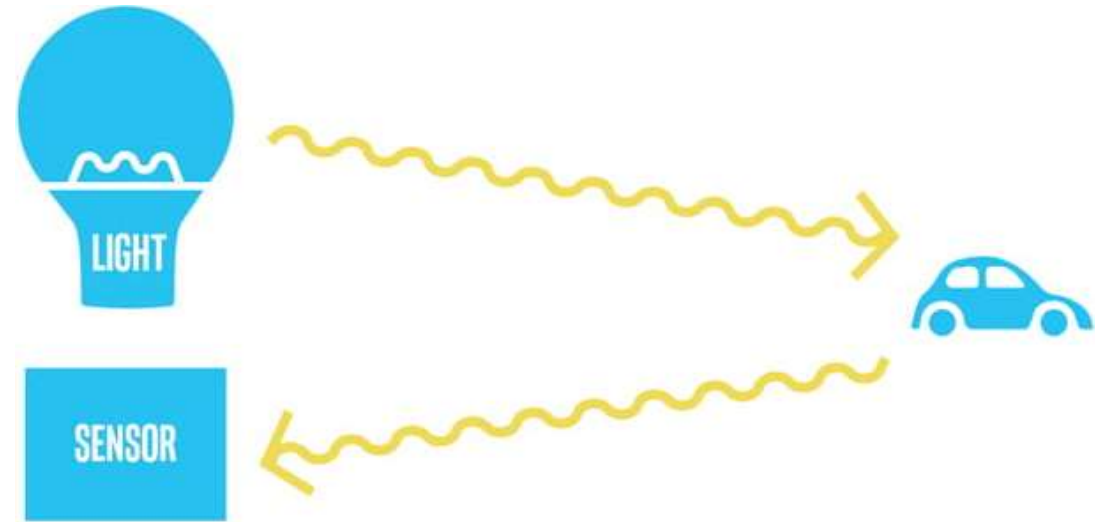
- ▶ Głębina może być dość dokładna, ale to zależy od konstrukcji wzorów światła i algorytmu analizującego ich odkształcenia.
- ▶ Nie zależy od tekstury obserwowanych obiektów.

Wady:

- ▶ Działa skutecznie jedynie dla odpowiednio oświetlonej sceny. Ma problemy szczególnie z mocno naświetloną sceną w plenerze.
- ▶ Dokładność głębi znacznie spada wraz ze wzrostem odległości, ponieważ wyświetlane wzory są słabo widoczne (obejmują mało pikseli obrazu).

ToF (Time-of-Flight)

- ▶ Kamera ma wbudowany emiter światła podczerwonego (niewidocznego dla ludzi), który oświetla obserwowaną scenę.
- ▶ Wyemitowane promienie światła po odbiciu od obiektu wracają do wbudowanego sensora.
- ▶ Kamera mierzy w każdym obserwowanym punkcie czas, który minął od momentu wyemitowania promienia do momentu odebrania go przez sensor.

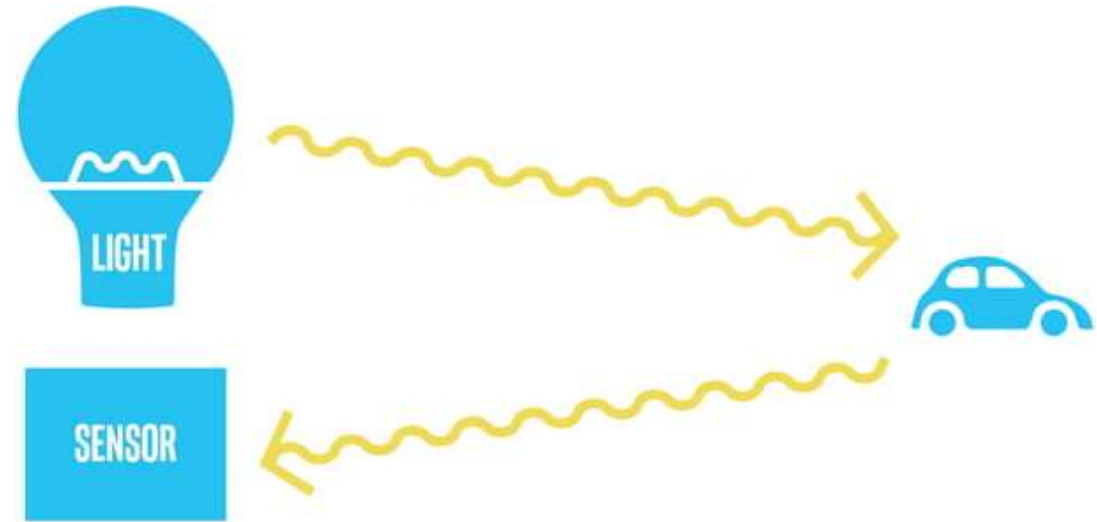


- ▶ Na tej podstawie obliczana jest głębia danego punktu.

Źródło: <https://www.intelrealsense.com/beginners-guide-to-depth>

ToF (Time-of-Flight)

- ▶ Najczęściej obliczanie czasu nie odbywa się za pomocą wbudowanego w kamerę zegara.
- ▶ Mierzenie czasu bezpośrednio jest trudne i kosztowne, ponieważ prędkość światła jest bardzo duża. Wymaga to bardzo precyzyjnych (kosztownych) układów pomiarowych.



- ▶ Zamiast tego czas powrotu promienia do sensora najczęściej aproksymuje się za pomocą pomiaru przesunięcia fazowego fali świetlnej.

Źródło: <https://www.intelrealsense.com/beginners-guide-to-depth>

ToF (Time-of-Flight) – wady i zalety

Zalety:

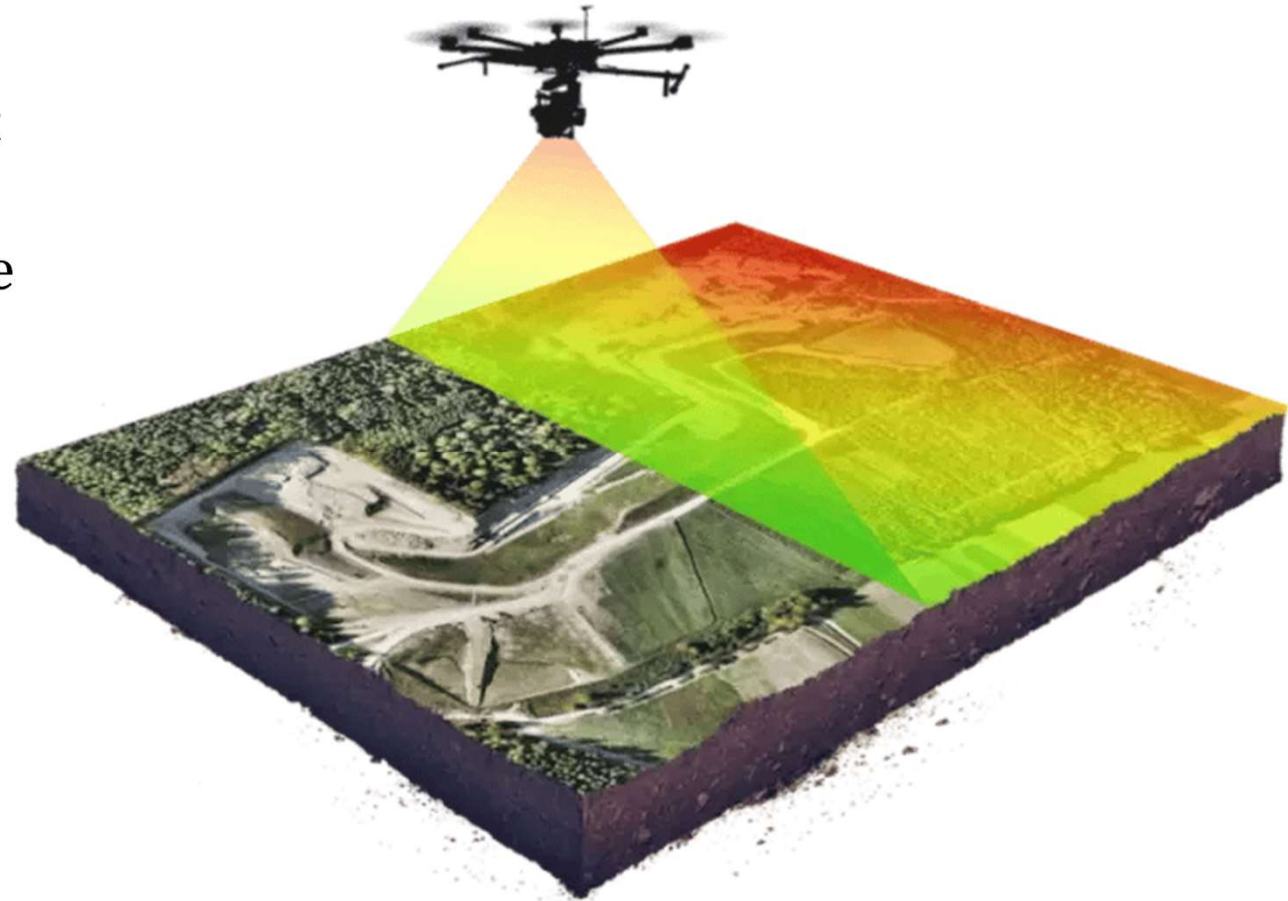
- ▶ Bezpośredni, sprzętowy pomiar głębi, nie wymagający skomplikowanych obliczeń i algorytmów.
- ▶ Mniej zależna od oświetlenia od dwóch poprzednich technik. Może działać nawet przy bardzo słabym oświetleniu.
- ▶ Najlepiej spośród wszystkich technik sprawdza się dla szybko poruszających się obiektów. Metody bazujące na obrazach kolorowych mają z nimi problem, ponieważ takie obiekty się rozmazują. Time-of-Flight nie ma tego problemu.

Wady:

- ▶ Technologia droższa niż dwie poprzednie (jednak kamera Kinect (Xbox One) bazująca na technologii ToF nie była bardzo droga).
- ▶ Nie nadaje się do dużych odległości; nie radzi sobie z mocno oświetlonym i refleksyjnymi powierzchniami; zazwyczaj mapy głębi nie mają dużych rozdzielczości.

LIDAR (Light Detection and Ranging)

- ▶ Zasada działania jest bardzo podobna do ToF. Jednak zamiast standardowych promieni świetlnych stosuje się promienie laserowe.



Źródło: <https://www.flyability.com/lidar-drone>

LIDAR (Light Detection and Ranging) – wady i zalety

Zalety:

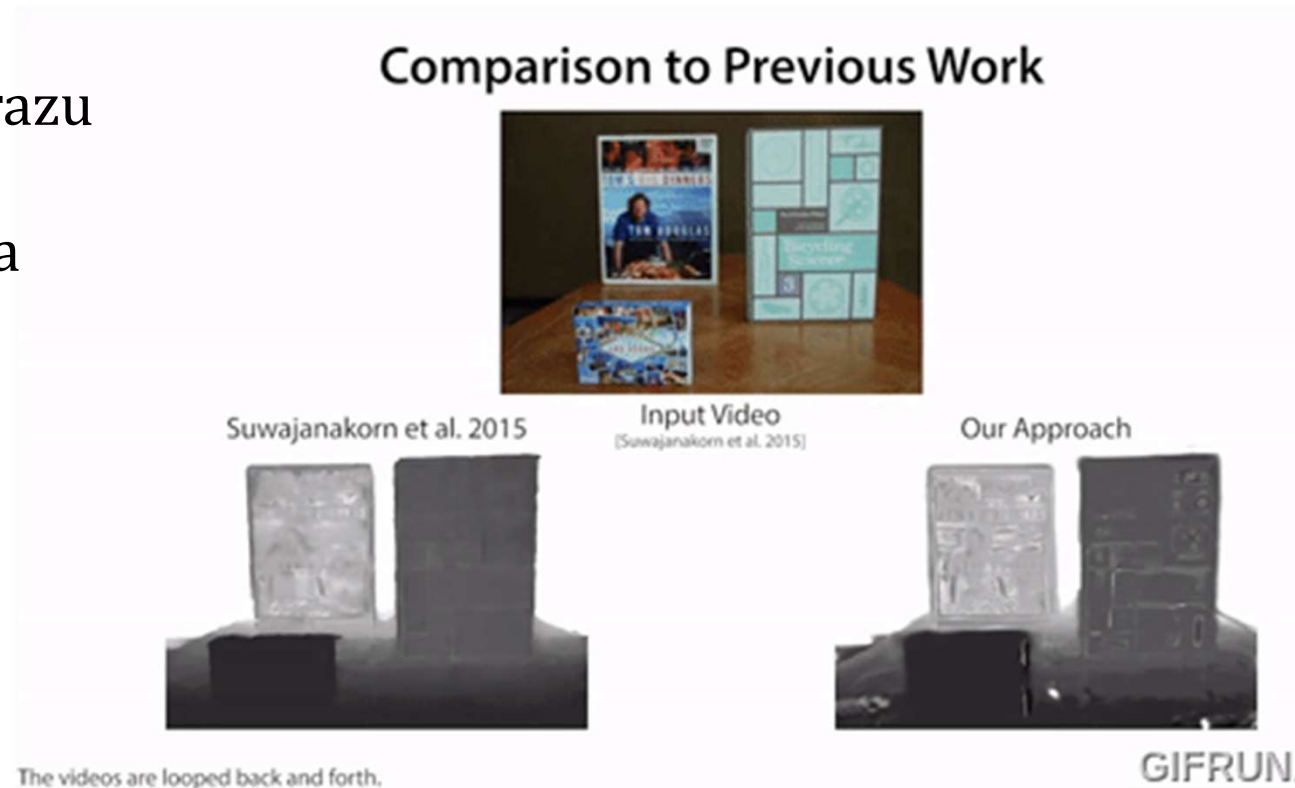
- ▶ Większa precyzja i rozdzielczość niż sensory ToF.
- ▶ Można dokonywać pomiarów na znacznie większych odległościach niż ToF (np. skanowanie terenu przez drona).

Wady:

- ▶ Technologia bardziej kosztowna niż ToF.
- ▶ LIDAR radzi sobie gorzej niż ToF z szybko poruszającymi się obiektami obserwowanymi z małych odległości.

DFD (Depth From Defocus)

- ▶ Algorytm analizuje rozmycie poszczególnych fragmentów obrazu zmieniające się wraz ze zmianą skupienia (ang. focus) kamery na poszczególnych obiektach.



Źródła: <https://doi.org/10.1109/3DV.2016.36>
https://www.youtube.com/watch?v=EYmYnXX_NP0

DFD (Depth From Defocus) – wady i zalety

Zalety:

- ▶ Wymaga tylko pojedynczej kamery kolorowej (nie trzeba drugiej kamery ani dodatkowych emiterów światła).
- ▶ Możemy uzyskać mapy głębi o wysokiej rozdzielczości, ale to zależy od stosowanej kamery.

Wady:

- ▶ Najmniej dokładna głębia. Nie nadaje się do poważnych zastosowań.

Przykładowe kamery pozyskujące dane głębi

Dobre kamery,
niestety już nie
wspierane.

Kinect (Xbox 360)



Kinect (Xbox One)



Azure Kinect DK



Intel RealSense



Luxonis Oak-D Pro



Stereolabs ZED 2i



Nowsze
urządzenia,
wspierane do
dziś.

Przykładowe kamery pozyskujące dane głębi – technologie pozyskiwania głębi

światło strukturalne

Kinect (Xbox 360)



Kinect (Xbox One)



Azure Kinect DK



time-of-flight

stereowizja

Intel RealSense



Luxonis Oak-D Pro



Stereolabs ZED 2i



Kamery głębi w smartfonach

- ▶ Niektóre współczesne smartfony są wyposażone w kamery głębi.
- ▶ Najczęściej są to sensory ToF. Przykładowe modele:
 - ▶ Samsung Galaxy S20+ i późniejsze,
 - ▶ Google Pixel 6 Pro i późniejsze,
 - ▶ Huawei P30 Pro i późniejsze.
- ▶ Smartfony Apple iPhone od modelu 12 Pro mają sensory głębi działające w technologii LIDAR.

Przykładowe biblioteki i funkcje do przetwarzania danych 3D

- ▶ *PCL* (Point Cloud Library) – biblioteka w języku C++ do przetwarzania chmur punktów - <https://pointclouds.org>. Istnieje możliwość korzystania z *PCL* w języku Python – <https://github.com/strawlab/python-pcl>
- ▶ *Open3D* – biblioteka w języku Python do przetwarzania chmur punktów – <https://www.open3d.org>
- ▶ Pakiet *Computer Vision Toolbox* programu Matlab, w tym funkcje z kategorii *Point Cloud Processing* (do przetwarzania chmur punktów):
 - ▶ <https://www.mathworks.com/help/vision>
 - ▶ <https://www.mathworks.com/help/vision/point-cloud-processing.html>
- ▶ *OpenCV* (Open Computer Vision) – biblioteka do przetwarzania obrazów (w tym map głębi) w językach C++ oraz Python – <https://opencv.org>