

Wizja Maszynowa

Wprowadzenie wykład 1

Adam Szmigielski

aszmigie@pjawst.edu.pl

materiały: *ftp(public) : //aszmigie/WMA*

Semestralny plan wykładów

1. Wizja komputerowa - wprowadzenie,
2. Tworzenie obrazu,
3. Przetwarzanie obrazu - filtrowanie,
4. Przetwarzanie obrazu - wykrywanie cech geometrycznych,
5. Detekcja cech obrazu,
6. Algorytmy detekcja i porównywani cech obrazu,
7. Sieci konwolucyjne,
8. Uczenie głębokie,
9. Architektury splotowych sieci neuronowych,
10. Metody częstotliwościowe w przetwarzaniu obrazu,
11. Stereowizja,
12. Wizja i ruch,
13. Wizualny SLAM - Jednoczesna lokalizacja i mapowanie,
14. Segmentacja obrazu,
15. Sementyczna segmentacja obrazu.

Semestralny plan ćwiczeń

1. Zajęcia wstępne, zapoznanie się z laboratorium,
2. Zajęcia z openCV, modele barw, filtry,
3. Identyfikacja obiektów kolorowych - projekt 1,
4. Śledzenie obiektów kolorowych,
5. Identyfikacja monet i tacy - projekt 2,
6. Opis obiektu z wykorzystaniem deskryptorów (np. algorytm SIFT),
7. Śledzenie obiektu z wykorzystaniem opisu deskryptorów - projekt 3,
8. Sieci konwolucyjne z wykorzystaniem Keras,
9. Nauka sieci konwolucyjnych - tworzenie własnych wzorców,
10. Identyfikacja obiektów z użyciem sieci neuronowych,
11. Implementacja algorytmu Optical Flow (OF w openCV) - projekt 4
12. Implementacja algorytmu Optical Flow (OF w openCV),
13. Implementacja algorytmu segmentacji - projekt 5,
14. Kolokwium,
15. Zajęcia końcowe - oddanie projektów, wystawienie ocen,

Warunki zaliczenia przedmiotu WM

- Z przedmiotu WM są dwie oceny - ocena z ćwiczeń i ocena z egzaminu,
- Przedmiot kończy się egzaminem. Pod koniec semestru odbędzie się kolokwium. Ocena z kolokwium jest oceną z egzaminu zerowego. Nie przewiduje się zwolnień z egzaminu.
- Do zdobycia 100 punktów = 70 p. ćwiczenia + 30 p. kolokwium
 - *Ćwiczenia 70 p.* - W trakcie ćwiczeń studenci dostaną kilka miniprojektów do zrealizowania na ćwiczeniach. Projekt bez straty punktów będzie można oddać na kolejnych zajęciach. Przedłużenie o kolejny tydzień powoduje stratę punktów.
 - *Kolokwium 30 p.* - kolokwium na 14 ćwiczeniach.
- W przypadku uzasadnionej nieobecności na kolokwium (14 tydzień) istnieje możliwość napisania kolokwium poprawkowego na 15 ćwiczeniach.

- Ocenę z ćwiczeń wystawia się w oparciu o poniższą tabelę:

| oceny | punkty - ćwiczenia + wykład |
|-------|-----------------------------|
| 2 | 0- 50 |
| 3 | 50,5 - 60 |
| 3,5 | 60,5 - 70 |
| 4 | 70,5 - 80 |
| 4,5 | 80,5 - 90 |
| 5 | 90,5 - 100 |

Kolokwium - egzamin zerowy

- Na 14 ćwiczeniach jest kolokwium za 30 punktów,
- Kolokwium traktowane jest jak egzamin zerowy,
- Ocenę z kolokwium przepisywana jest na egzamin w oparciu o poniższą tabelę:

| ocena | punkty z wykładu |
|-------|------------------|
| 2 | 0- 15 |
| 3 | 15,5 - 18 |
| 3,5 | 18,5 - 21 |
| 4 | 21,5 - 24 |
| 4,5 | 24,5 - 27 |
| 5 | 27,5 - 30 |

- Warunkiem przepisania oceny z kolokwium na egzamin zerowy jest zaliczenie ćwiczeń.

Oprogramowanie

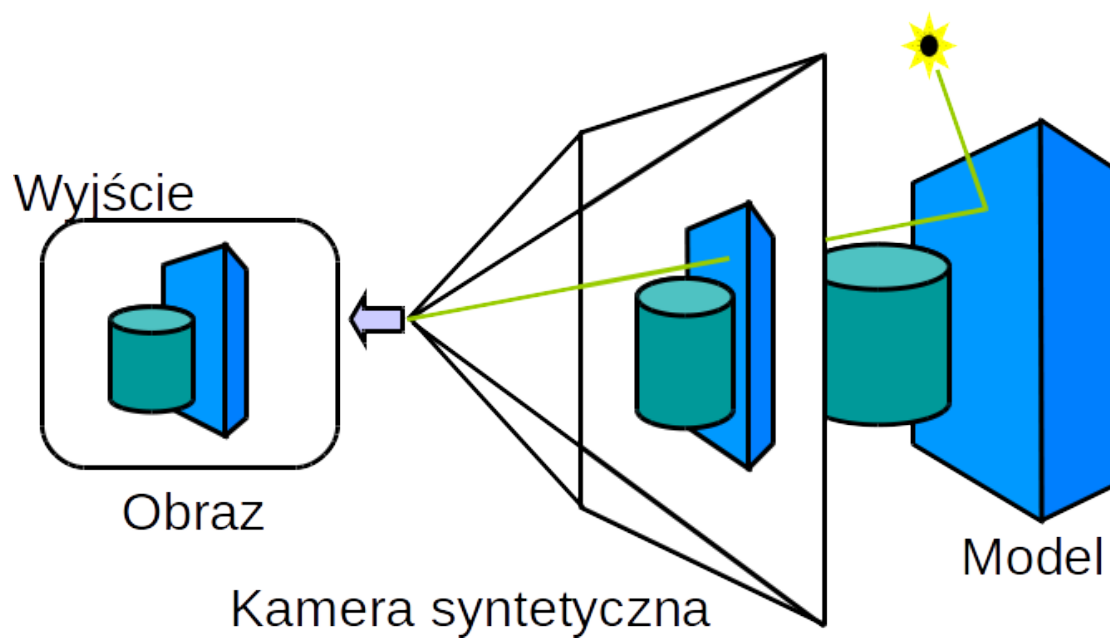
Projekty będą omawiane z wykorzystaniem języka Python 3.

Oprogramowanie

- Python 3
- OpenCV
- Biblioteki numpy, tensorflow, keras etc. - potrzebne środowisko do zarządzania modułami np. PIP

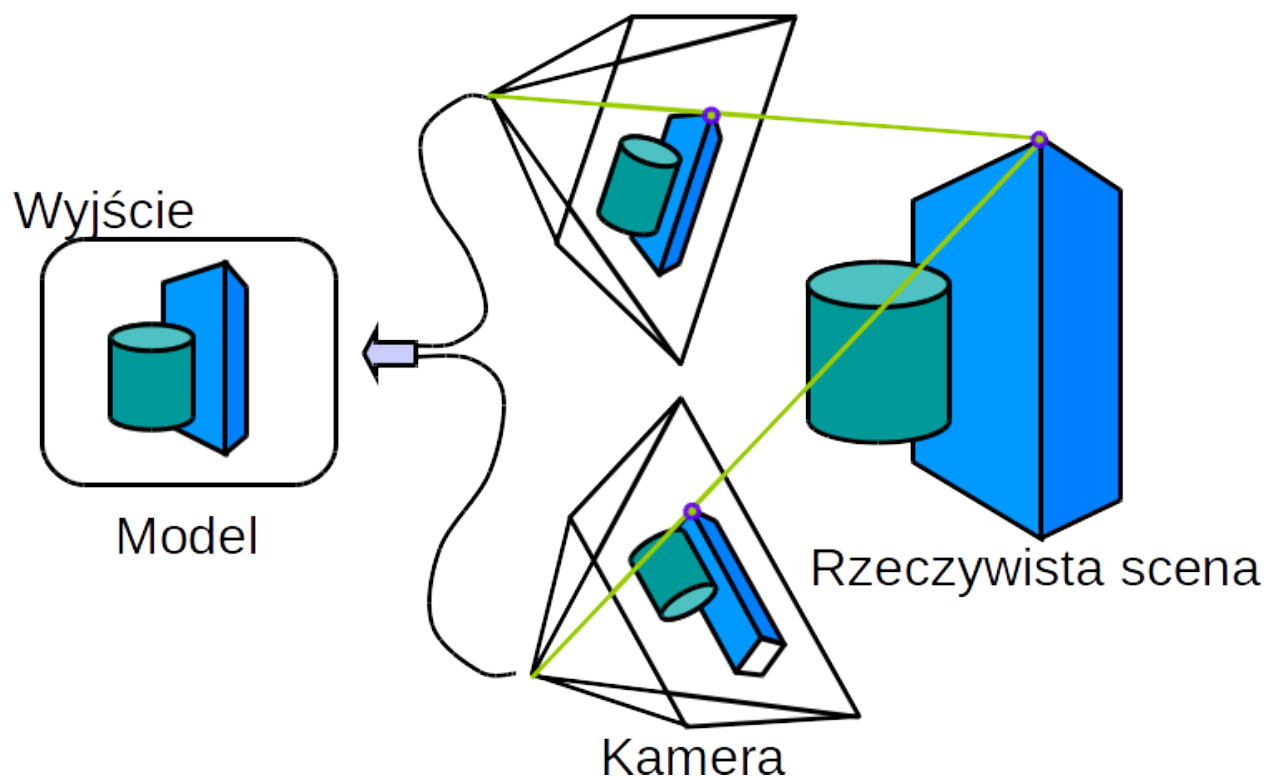
Link do właściwego tutorialu opencv

Grafika komputerowa



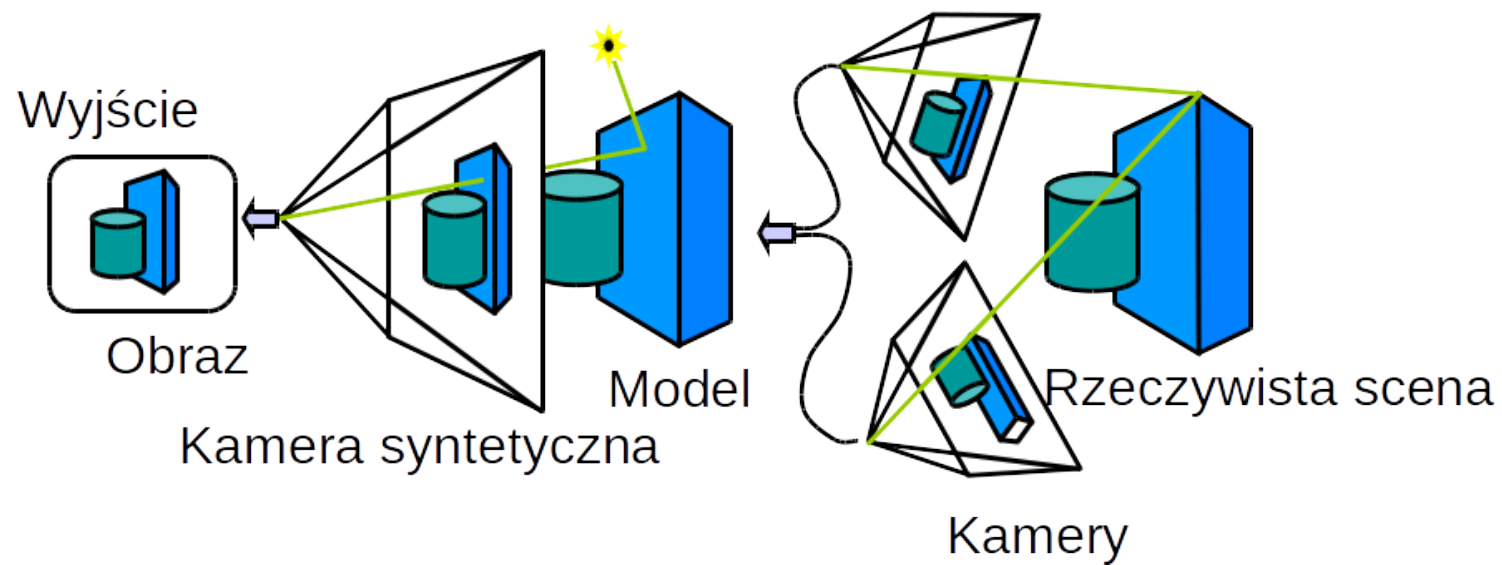
•

Wizja komputerowa



•

Układy kombinowane



- Rzeczywistość wzbogacona,

Wizja komputerowa

- Podstawy
 - Tworzenie obrazu i kalibracja kamery:
 - Funkcje obrazu,
- Rekonstrukcja 3D:
 - Stereowizja,
 - Obraz Mozaiki,
- Wykrywanie i rozpoznawanie obiektów:
 - Grupowanie,
 - Wykrycie,
 - Segmentaiton,
 - Klasyfikacja,
 - Śledzenie,

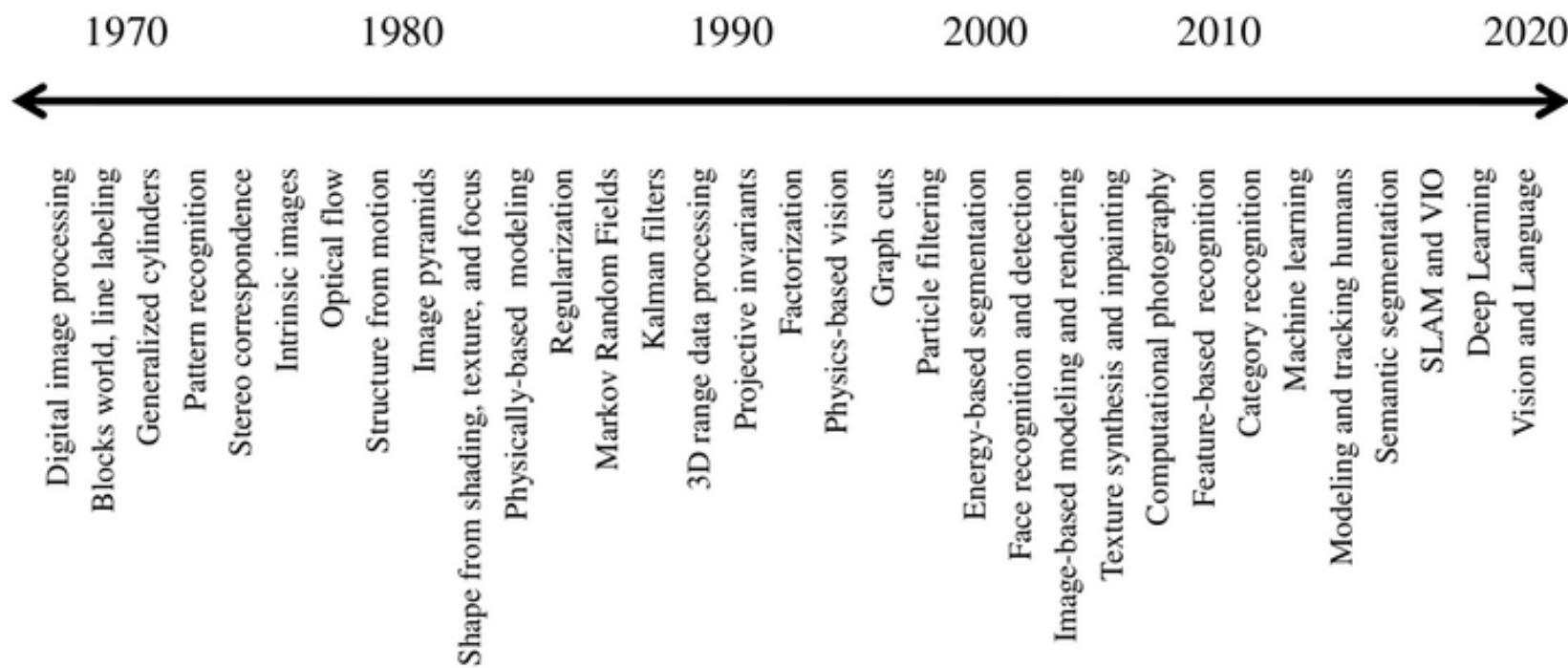
Zastosowanie wizji komputerowej

- Optyczne rozpoznawanie znaków (OCR): czytanie odręczne,
- Kontrola maszyn: szybka kontrola części w celu zapewnienia jakości przy użyciu stereoskopowego widzenia,
- Retail: rozpoznawanie obiektów dla automatycznych pasów kasowych i w pełni zautomatyzowanych sklepów,
- Logistyka magazynowa: autonomiczna dostawa paczek,
- Obrazowanie medyczne: rejestrowanie obrazów przedoperacyjnych i śródoperacyjnych,
- Pojazdy autonomiczne: zdolne do jazdy punkt-punkt między miastami,
- Budowanie modeli 3D (fotogrametria): w pełni zautomatyzowana budowa modeli 3D ze zdjęć z drona
- Motion capture (mocap): używanie markerów oglądanych z wielu

kamer

- Surveillance: monitorowanie intruzów, analizowanie ruchu na autostradach
- Rozpoznawanie odcisków palców i biometria: do automatycznego uwierzytelniania dostępu oraz zastosowań kryminalistycznych,
- Wykrywanie twarzy: dla lepszego ustawienia ostrości aparatu, a także bardziej trafnego wyszukiwania obrazów,
- Uwierzytelnianie wizualne: automatyczne logowanie,

Historia wizji komputerowej



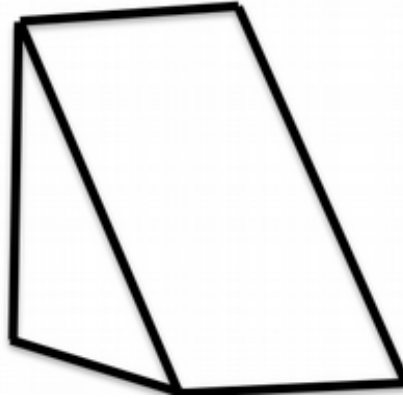
Historia wizji komputerowej

- 1966: Minsky rozpoczyna letni projekt licencjacki z wizji komputerowej,
- Lata 60 (Larry Roberts, 1963) Interpretacja materiału syntetycznego,
- lata 1970 Próby interpretacji obrazu,
- Lata 80 próby stworzenie formalnego opisu - w kierunku geometrii,
- lata 90 Rozpoznawanie twarzy - statystyczny analiza,
- lata 2000 Duże zestawy danych z etykietami rozpoczyna się przetwarzanie video.

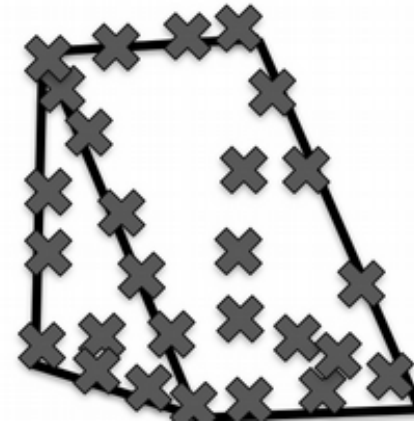
Świat brył (ang. Block world) Larry Roberts, 1963



obraz oryginalny

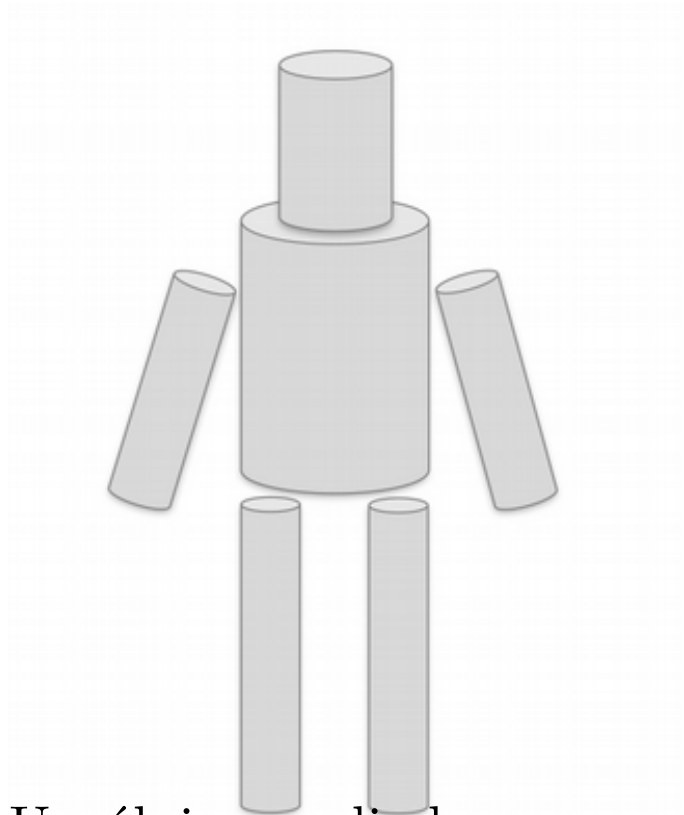


opis różniczkowy

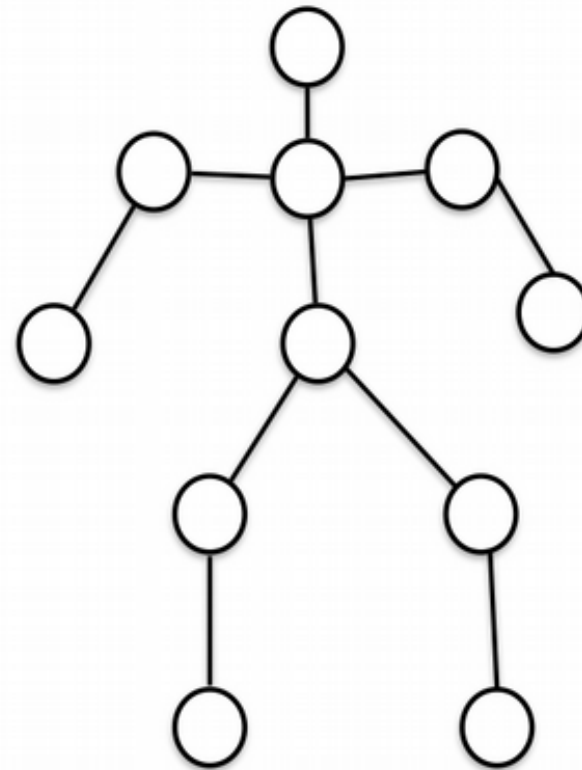


wybrane punkty obrazu

Modele obrazu - interpretacja



Uogólniony cylinder
Brooks Binford, 1979



Struktura obrazkowa
Fischler and Elschlager, 1973

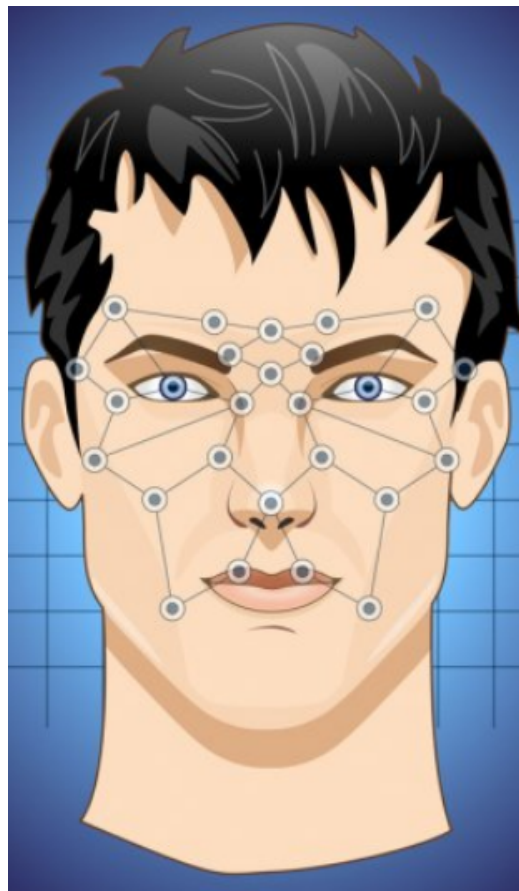
David Lowe, 1987 - filtracja



obraz oryginalny

po filtracji

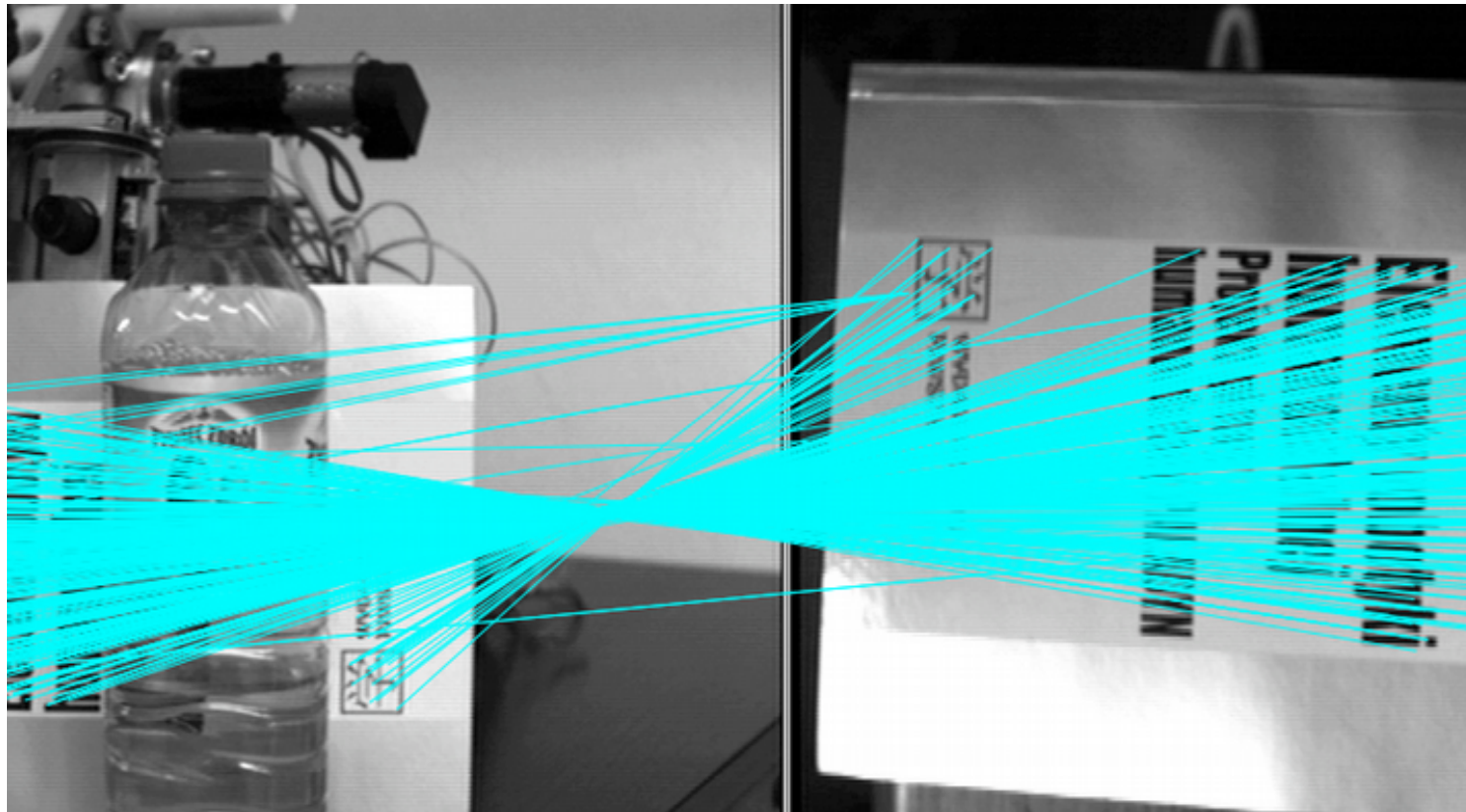
Identyfikacja twarzy - lata 90



Dekada 2000

- Objął podejście oparte na danych i uczeniu się jako podstawowe elementy wizji,
- W tym dziesięcioleciu pojawiły się techniki oparte na cechach (w połączeniu z uczeniem się) do rozpoznawania obiektów,
- Większość aspektów widzenia komputerowego polegała na zastosowaniu zaawansowanych technik uczenia maszynowego do rozwiązywania problemów z widzeniem komputerowym.

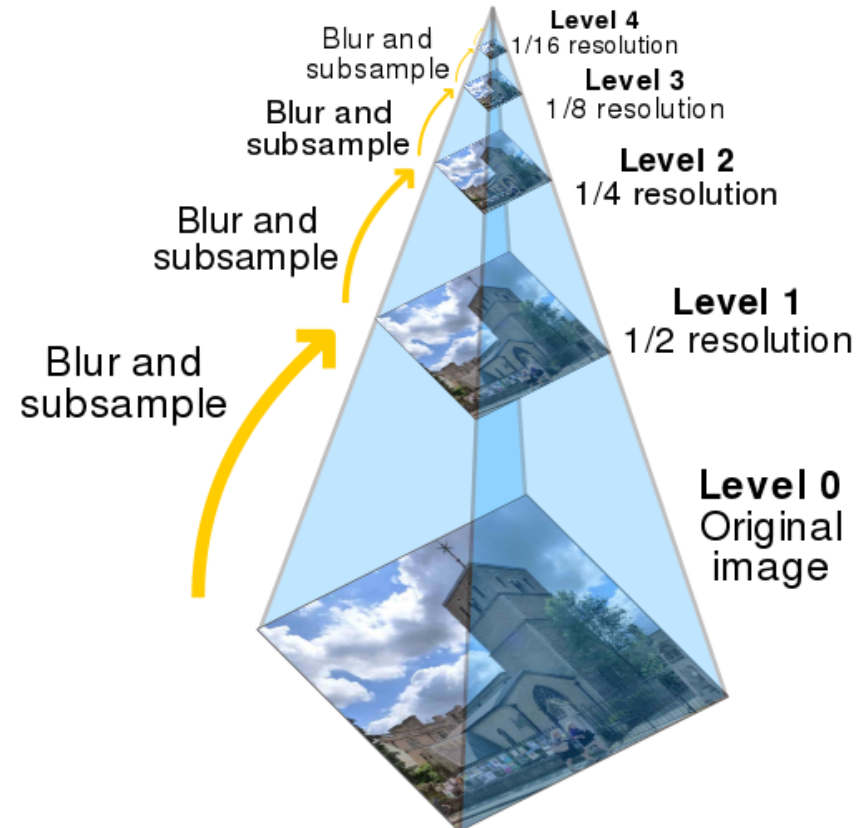
SIFT i identyfikacja wzorców, David Lowe, 1999



Dekada 2010

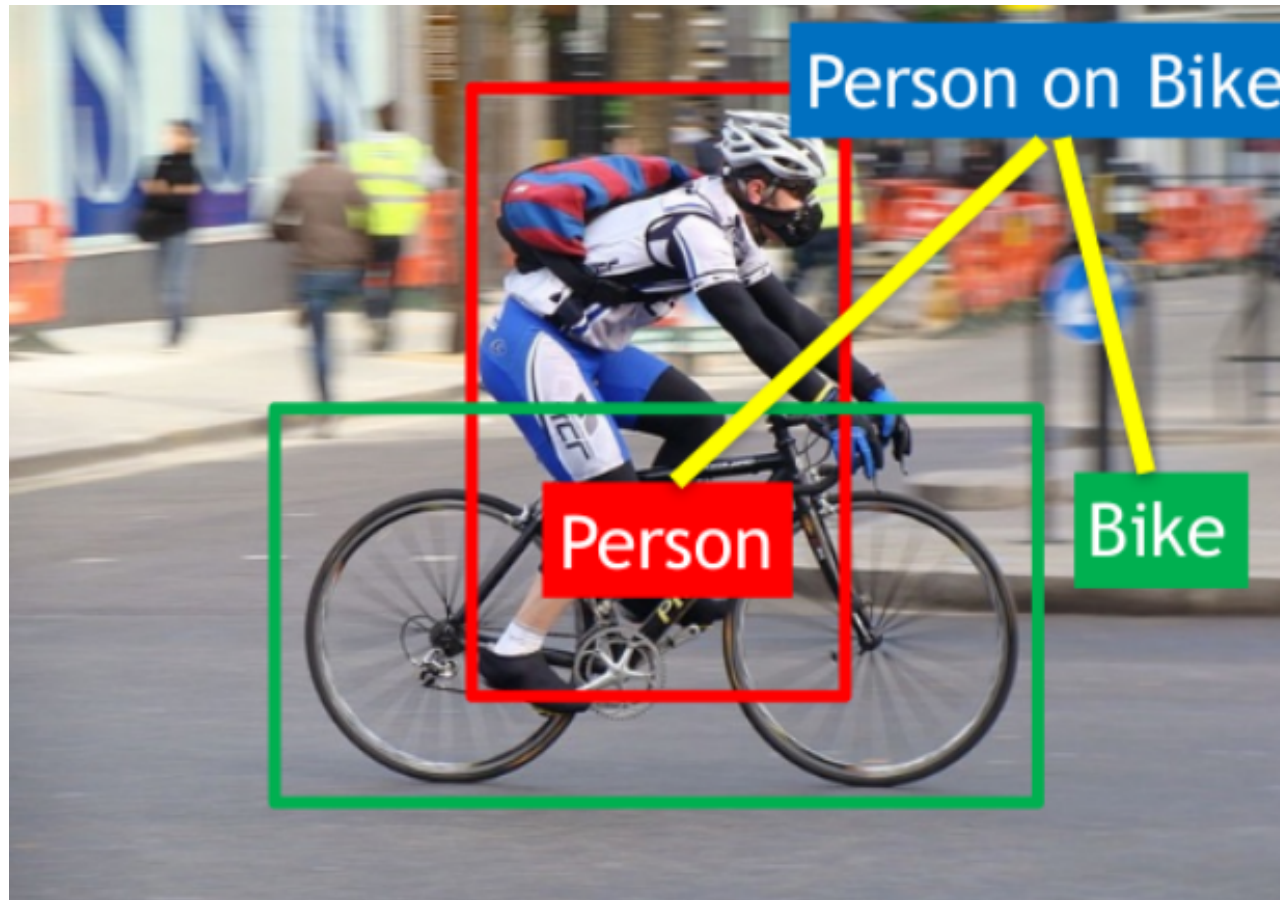
- Tendencja do wykorzystywania dużych, oznaczonych zbiorów danych do opracowywania algorytmów uczenia maszynowego,
- Te zbiory danych dostarczyły nie tylko wiarygodnych metryk do śledzenia postępu algorytmów rozpoznawania i segmentacji semantycznej,
- Specjalistyczne czujniki i sprzęt do zadań wizji komputerowej - kamera głębi Kinect, wydana w 2010 roku, szybko stała się niezbędnym elementem wielu modeli 3D,
- SLAM w czasie rzeczywistym (jednoczesna lokalizacja i mapowanie) i VIO (wizualna odometria inercyjna).

Algorytm piramid

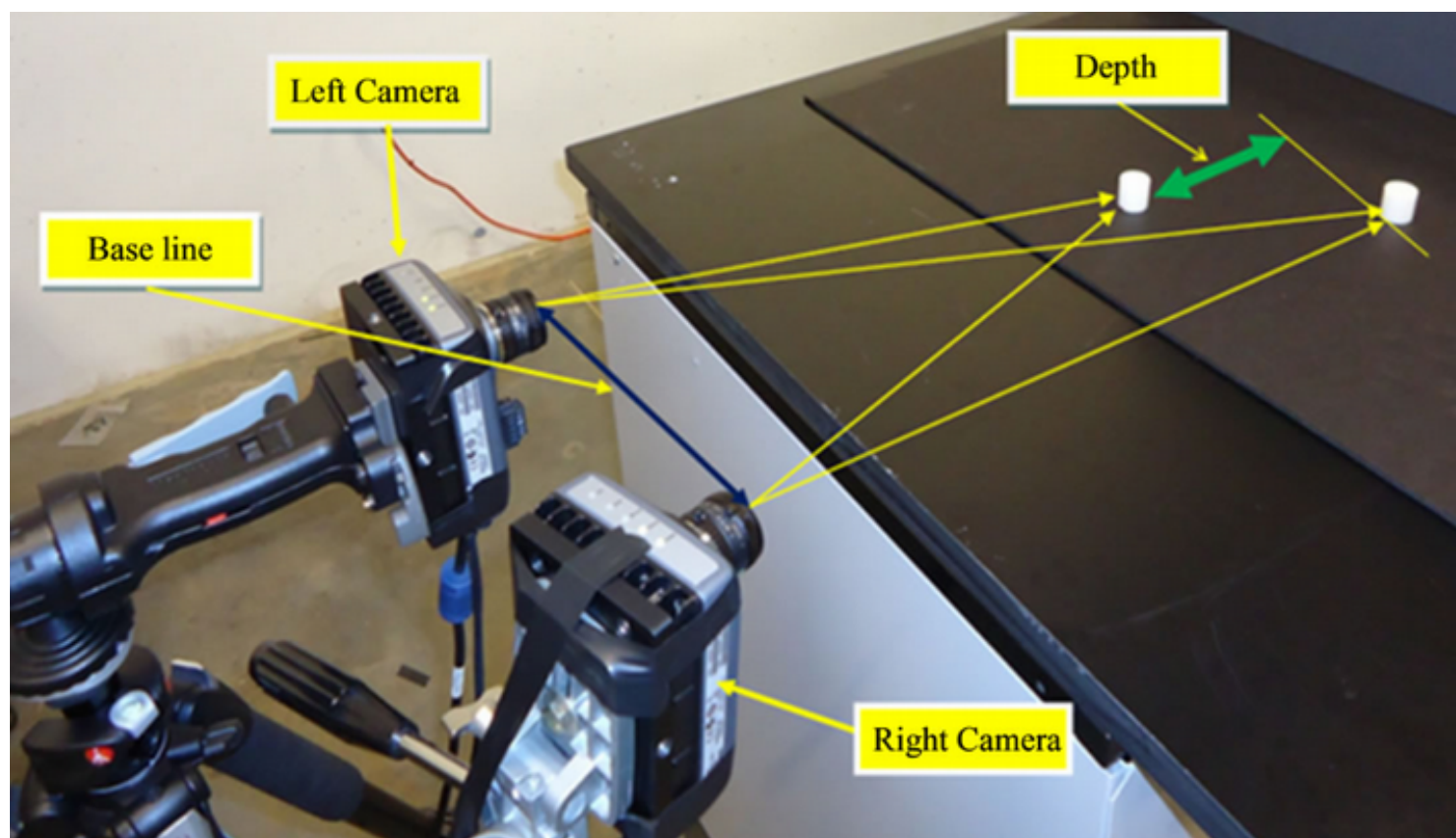


- Algorytm ten ma własność zachowywania niezmienników przekształceń liniowych - translacje, obrót, skalowanie.
- Własność ta daje możliwość opisu wzorca jako zbioru niemienników przekształceń oraz jest powodem powstania nowego rodzaju sieci - sieci konwolucyjnych CNN

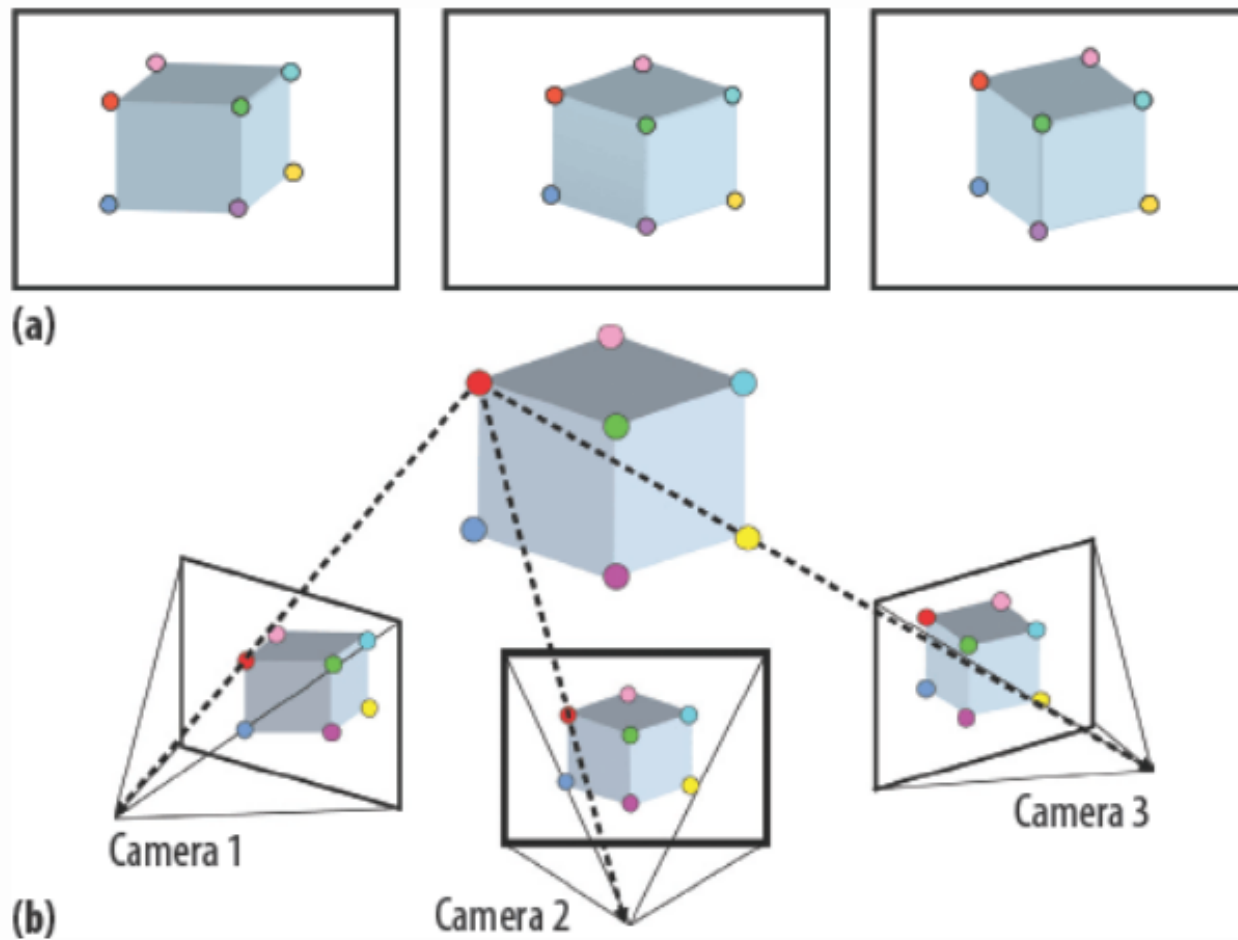
Visual Object Challenge - do chwili obecnej



Stereowizja - wyznaczanie głębi obrazu



Strukturalizacja



Literatura

1. Richard Szeliski: *Computer Vision: Algorithms and Applications*, Springer 2010
2. Richard Szeliski: *Computer Vision: Algorithms and Applications*, 2nd Edition
<http://szeliski.org/Book/>
3. Richard O. Duda, Peter E. Hart, David G. Stork: *Pattern Classification*, 2ed, Wiley, 2001
4. Christopher M. Bishop: *Pattern Recognition and Machine Learning*, Springer, 2006
5. Ian Goodfellow, Yoshua Bengio, Aaron Courville: *Deep Learning Systemy uczące się*, PWN, 2018
6. Sebastian Raschka: *Python Uczenie maszynowe*, Helion, 2018

7. Josh Patterson, Adam Gibson: *Deep Learning. Praktyczne Wprowadzenie*, Helion, 2018
8. Aurelien Geron: *Uczenie maszynowe z użyciem scikit learn i tensorflow*, Helion, 2018
9. Francois-Chollet: *Deep learning praca z jezykiem python i biblioteka keras*, Helion, 2018