

Widzenie Komputerowe - Wstęp

Wykład 1.

Magdalena Mazur-Milecka

Katedra Inżynierii Biomedycznej, WETI, PG

25 lutego 2020

Wykład:

- Kolokwium 40 pkt, zalicza 50%
- klucz do kursu na enauczanie.pg.edu.pl:

CV_1920

Laboratorium:

- 5 x laboratorium 3h, w drugiej połowie semestru
- Do zdobycia 5 x 10 pkt. = 50 pkt. + 10 pkt. za test, zalicza 30 pkt.
- Język - Python3, biblioteka OpenCV, scikit-image, TensorFlow, Keras
- Początek lab - 28.04.2020

Wykład:

- Kolokwium 40 pkt, zalicza 50%
- klucz do kursu na enauczanie.pg.edu.pl:

CV_1920

Laboratorium:

- 5 x laboratorium 3h, w drugiej połowie semestru
- Do zdobycia 5 x 10 pkt. = 50 pkt. + 10 pkt. za test, zalicza 30 pkt.
- Język - Python3, biblioteka OpenCV, scikit-image, TensorFlow, Keras
- Początek lab - 28.04.2020

Wykład:

- Kolokwium 40 pkt, zalicza 50%
- klucz do kursu na enauczanie.pg.edu.pl:

CV_1920

Laboratorium:

- 5 x laboratorium 3h, w drugiej połowie semestru
- Do zdobycia 5 x 10 pkt. = 50 pkt. + 10 pkt. za test, zalicza 30 pkt.
- Język - Python3, biblioteka OpenCV, scikit-image, TensorFlow, Keras
- Początek lab - 28.04.2020

Punktacja z całego przedmiotu - 100 pkt. Warunki konieczne zaliczenia

- min. 50 pkt z całego przedmiotu,
- min. 20 pkt z kolokwium + min. 30 pkt. z laboratorium

- Python3
- test zaliczeniowy na stronie kursu tydzień przed rozpoczęciem zajęć (21.04.2020) (w planowanym terminie zajęć - widoczny przez 15')
- Zakres testu:
 - zmienne i typy
 - listy, słowniki i tuple
 - obsługa macierzy
 - pętle i instrukcje warunkowe
 - procedury (funkcje, także `__init__`)
 - programowanie obiektowe

- Programowanie obiektowe, strukturalne oraz funkcjonalne
- Dynamiczny system typów
- Automatyczne zarządzanie pamięcią
- Wieloplatformowość i biblioteki
- Struktura przez wcięcia
- Interaktywna powłoka + IDE (PyCharm, Spyder, IDLE, ActivePython, PyScripter, Jupyter Notebook)

- Widzenie Komputerowe
 - ① Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2011
 - ② Forsyth, D. and Ponce, J. (2003). Computer Vision: A Modern Approach. Prentice Hall, Upper Saddle River, NJ.
- Python
 - ① Mark Lutz, Python. Wprowadzenie. Wydanie IV, Helion 2011
- Open CV
 - ① Gary Bradski, Adrian Kaehler, Learning OpenCV, O'Reilly Media, 2008.

- Transformacja Fouriera w obrazowaniu - Rekonstrukcja i Analiza Obrazów, sem. 6.,
- Operacje morfologiczne - Podstawy Przetwarzania Obrazów, sem. 4.,
- Histogramy - Podstawy Przetwarzania Obrazów,
- Filtracja obrazu - Podstawy Przetwarzania Obrazów,
- Detekcja krawędzi, punktów i obszarów, detektory i deskryptory (Rekonstrukcja i Analiza Obrazu - Harris, SIFT).

- 1 Wstęp. Pojęcia. Historia.
- 2 Segmentacja (CC, Alg. aktywnych konturów (detekcja konturów), snake),
- 3 Segmentacja (alg. działów wodnych, metody AI)
- 4 Analiza tekstury (2D histogram based algorithms)
- 5 Analiza tekstury (Fourier, Gabor)
- 6 Ruch (MHI, optical flow)
- 7 PCA, ICA, optymalizacja, regularyzacja
- 8 Recognition (1. Hough, 2. Detekcja obiektów - detektory, deskryptory),
- 9 Recognition (2. Detekcja obiektów - ekstrakcja, dopasowywanie i śledzenie cech
3. Detekcja twarzy),
- 10 Recognition (4. Rozpoznanie twarzy: Eigenfaces, Active Appearance Models)
- 11 Wstęp do klasyfikacji
- 12 Klasyfikacja (Regresja logistyczna, Bayes),
- 13 Klasyfikacja (SVM, bag of words)
- 14 Convolution Neural Network
- 15 Convolution Neural Network

Człowiek z łatwością:

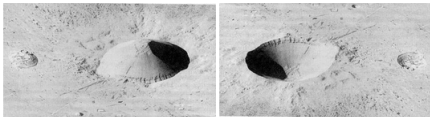
- dostrzega 3-wymiarowe struktury,
- odziera pierwszy plan od tła,
- rozpoznaje nastrój po wyrazie twarzy,
- rozróżnia obiekty na podstawie cech,
- identyfikuje zdarzenia trwające przez określony czas,
- odróżnia obiekty i zdarzenia o subtelnych różnicach,
- bierze pod uwagę inne rozwiązania (statystyka).

Człowiek z łatwością:

- dostrzega 3-wymiarowe struktury,
- odziela pierwszy plan od tła,
- rozpoznaje nastrój po wyrazie twarzy,
- rozróżnia obiekty na podstawie cech,
- identyfikuje zdarzenia trwające przez określony czas,
- odróżnia obiekty i zdarzenia o subtelnych różnicach,
- bierze pod uwagę inne rozwiązania (statystyka).

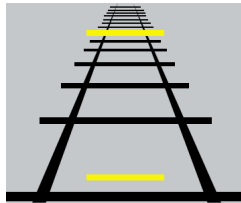
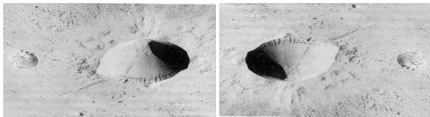
”Miliard lekcji wyniesionych z trylionowych interakcji ze światem, napisanych metodami, które ledwo rozumiemy, dla sieci połączonych ze sobą neuronów, bardziej złożonych niż wszystko, co kiedykolwiek napotkaliśmy.”

Złudzenia optyczne



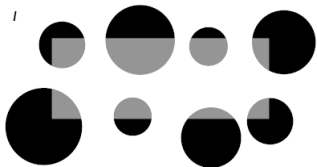
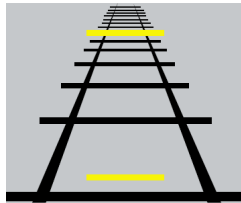
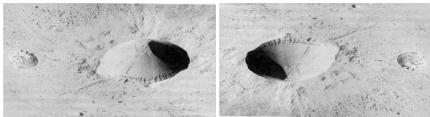
<http://analizaobrazu.x25.pl>

Złudzenia optyczne



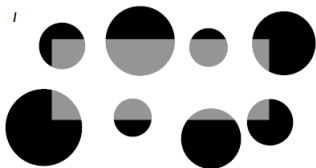
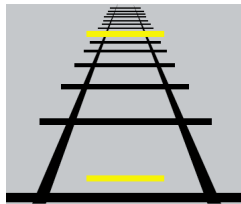
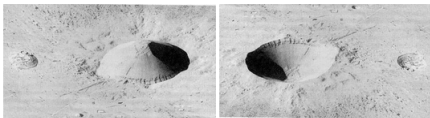
<http://analizaobrazu.x25.pl>

Złudzenia optyczne



<http://analizaobrazu.x25.pl>

Złudzenia optyczne



<http://analizaobrazu.x25.pl>



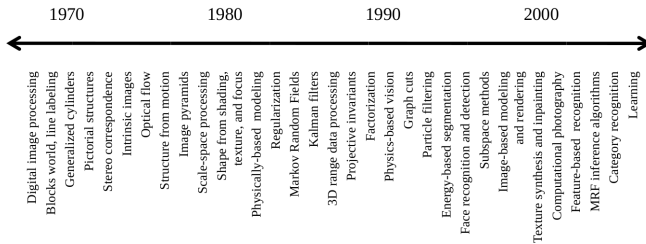
Rysunek: Marvin Minsky, Photo: Marie Cosindas

Widzenie

komputerowe -

algorytmy i systemy pozwalające komputerowi zrozumieć (rozpoznać) otoczenie, z punktu widzenia matematycznego - problem odwrotny.

- złożoność obliczeniowa, duży zbiór uczący,
- człowiek nie do końca wie jak sam to robi,
- nieznaną wpływ pozostałych zmysłów,
- poziom rozumienia obrazu niższy niż 2-letniego dziecka - paradoks Moraveca,
- kreatywność, rozwiązywanie problemów

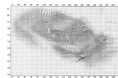


Rysunek: Linia czasu badanych zagadnień widzenia komputerowego ¹

Lata 70.

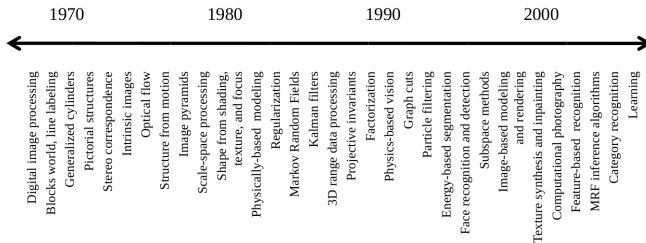


Rysunek: Etykietowanie linii (Nalva 1997)



Rysunek: OF (Nagel and Enkelmann 1986)

¹ Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2011



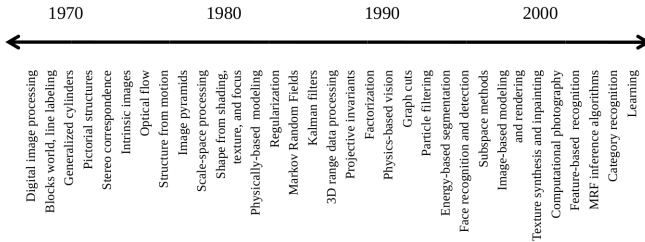
Rysunek: Linia czasu badanych zagadnień widzenia komputerowego ¹

Lata 80.



Rysunek: Canny, J. (1986). A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6):679–698.

¹ Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2011



Rysunek: Linia czasu badanych zagadnień widzenia komputerowego ¹

Lata 90.

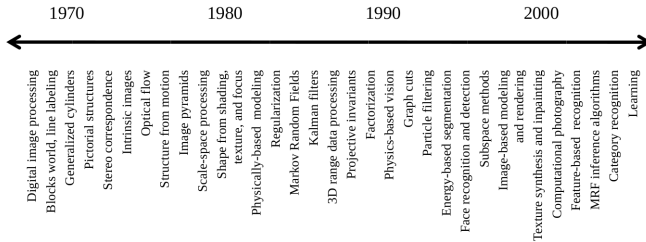


Rysunek: eigenfaces



Rysunek: Active Contours

¹ Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2011



Rysunek: Linia czasu badanych zagadnień widzenia komputerowego ¹

Lata
2000-2010



Rysunek: feature-based recognition

¹ Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2011

Zastosowania Widzenia Komputerowego

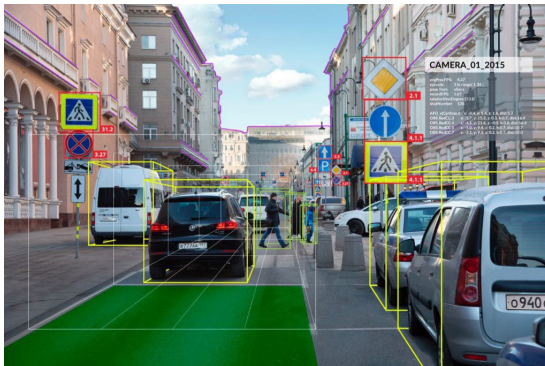
- rozpoznanie pisma, kodów
- kontrola jakości (produkcji w fabrykach, używanego sprzętu - skrzydła samolotu, części samochodu)
- budowa modeli 3D na podstawie obrazów
- obrazowanie medyczne
- biometria
- automatyczne pojazdy
- analiza gestów
- transport (kontrola ruchu) i sprzedaż
- ochrona (kontrola pasażerów na lotniskach, basenach i w dużych skupiskach)
- robotyka
- nawigacja w przestrzeni (drony)
- zabawa :)

Zabawa - blending

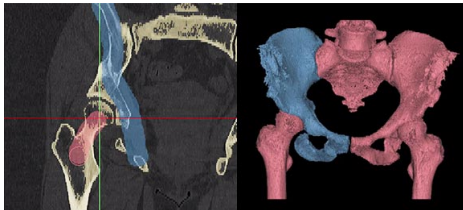


Gomes, J., Darsa, L., Costa, B., and Velho, L. (1999). *Warping and Morphing of Graphical Objects*. Morgan Kaufmann Publishers, San Francisco.

Automatyczne pojazdy

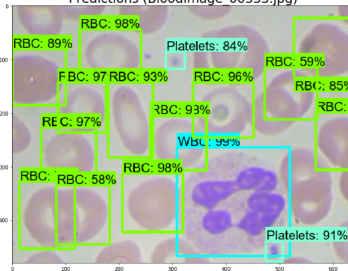
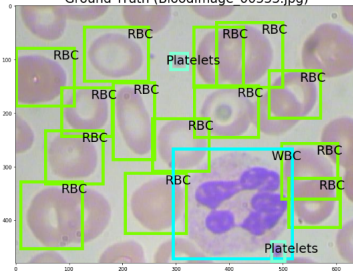


<https://becominghuman.ai/computer-vision-applications-in-self-driving-cars-610561e14118>



Ground Truth (BloodImage_00333.jpg)

Predictions (BloodImage_00333.jpg)



- **Detekcja** (ekstrakcja) - wyodrębnienie z obrazu obszaru, punktu lub krawędzi o określonych cechach



Brown, M., Szeliski, R., and Winder, S. (2005). Multi-image matching using multi-scale oriented patches. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2005), pp. 510–517, San Diego, CA

- Detekcja
- **Śledzenie** - zachowanie ciągłości danych o obiekcie w trakcie trwania sekwencji wideo, brak konieczności powtarzania detekcji w każdej klatce

- Detekcja
- Śledzenie
- **Segmentacja** - podział obrazu na obszary składające się z pikseli o jednakowych właściwościach (kolorze, teksturze)



Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):603-619

- Detekcja
- Śledzenie
- Segmentacja
- **Rozpoznanie** - można wyróżnić: dopasowanie obiektu i rozpoznanie kategorii (klasyfikacja)



Viola, P. A. and Jones, M. J. (2004). Robust real-time face detection. International Journal of Computer Vision, 57(2):137–154.

Shotton, J., Winn, J., Rother, C., and Criminisi, A. (2009). Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling appearance, shape and context. International Journal of Computer Vision, 81(1):2–23.