

# Przetwarzanie i rozpoznawanie obrazów: Wprowadzenie

Krzysztof Krawiec

Przetwarzanie i Rozpoznawanie Obrazów

14 czerwca 2018

## Kwestie organizacyjne

- Nazwa przedmiotu: *Przetwarzanie i rozpoznawanie obrazów*
- Semestr 1 studiów II stopnia, Inteligentne Systemy Wspomagania Decyzji
- Wymiar: 2W+2L
- Zaliczenie:
  - Wykład: egzamin
  - Laboratorium: zaliczenie ćwiczeń i projektów
- Strona WWW przedmiotu: <http://www.cs.put.poznan.pl/kkrawiec/>
- Prowadzący:
  - Krzysztof Krawiec, prof. dr hab. inż., [krawiec@cs.put.poznan.pl](mailto:krawiec@cs.put.poznan.pl)
  - Bartosz Wieloch, dr inż., [bwieloch@cs.put.poznan.pl](mailto:bwieloch@cs.put.poznan.pl)
- Wykład: prezentacja zagadnień teoretycznych i praktycznych przeplatana studiami przypadków.
- Powiązania z innymi przedmiotami: Sztuczna Inteligencja, Uczenie Maszynowe, Komunikacja Człowiek-Komputer

- 1 Wprowadzenie
- 2 Cyfrowa reprezentacja obrazów
- 3 Przetwarzanie obrazu
- 4 Metody ekstrakcji cech
- 5 Detekcja cech
- 6 Segmentacja obrazu
- 7 Głębokie sieci neuronowe
- 8 Elementy widzenia komputerowego i stereoskopia
- 9 Analiza ruchu
- 10 Pozyskiwanie obrazów

## Literatura podstawowa:

- Gonzalez, Wintz, Digital Image Processing. Addison-Wesley 2008
- Domański, M., Obraz cyfrowy. WKŁ 2010.
- Zieliński, T.P., Cyfrowe przetwarzanie sygnałów. WKŁ 2009.
- Kaehler, A., Bradski, G., OpenCV 3: Komputerowe rozpoznawanie obrazu w C++ przy użyciu biblioteki OpenCV, Helion, 2017

## Literatura uzupełniająca:

- Cyganek, B., Komputerowe przetwarzanie obrazów trójwymiarowych. EXIT 2002.
- Stąpor, K., Automatyczna klasyfikacja obiektów. EXIT 2005.
- Mark Owen, Przetwarzanie sygnałów w praktyce. WKŁ 2009.
- Choraś, R. Komputerowa wizja. Metody interpretacji i identyfikacji obiektów. EXIT, 2006.
- Watkins, C.D., Sadun, A., Marenka, S.: Nowoczesne metody przetwarzania obrazu. WNT 1995.
- Wojciechowski, K., Rozpoznawanie obrazów, Politechnika Śląska, Gliwice 1997.

- Pavlidis, T., Grafika i przetwarzanie obrazów. Algorytmy, (Seria: Biblioteka Inżynierii Oprogramowania), WNT, Warszawa 1987.
- Tadeusiewicz, R., Systemy wizyjne robotów przemysłowych, WNT, Warszawa 1992.
- Tadeusiewicz, R., Flasiński, M., Rozpoznawanie obrazów, (Seria: Współczesna Nauka i Technika. Informatyka) PWN, Warszawa 1991.
- Ostrowski, M. (red.), Informacja obrazowa. WNT, Warszawa, 1992.
- Nałęcz, M., Problemy biocybernetyki i inżynierii biomedycznej, WKiŁ, Warszawa 1990.
- Nieniewski, M., Morfologia matematyczna w przetwarzaniu obrazów, Akademicka Oficyna Wydawnicza, Warszawa 1998.
- Zieliński, K.W., Strzelecki, M.: Komputerowa analiza obrazu biomedycznego. Wstęp do morfometrii i patologii ilościowej, Wydawnictwo Naukowe PWN, 2002.
- Duda, R., Hart, P.: Pattern Classification and Scene Analysis, Wiley, New York, 1973.

- Biblioteki OpenCV, SciKit-Image
- ImageJ / Fiji
- ImageMagick
- R
- Matlab [?]
- Tensorflow / Keras
- Jupyter notebooks



# Wprowadzenie



- Fukunaga (1972) „Pattern recognition consists of two parts, feature selection and classifier design”
- Duda & Hart (1973) „Pattern recognition, a field concerned with machine recognition of meaningful regularities in noisy or complex environments”
- Gonzalez and Thomason (1978) „Pattern recognition can be defined as the categorization of input data into identifiable classes via the extraction of significant features or attributes of the data from a background of irrelevant detail”
- Bezdek (1981) „Pattern recognition is a search for structure in data”
- Schalkoff (1992) „Pattern recognition is the science that concerns the description or classification (recognition) of measurements”

## Uwarunkowania historyczne:

- Problemy terminologiczne: w tradycji PR “pattern” – obraz
- W ramach tego wykładu: Rozważamy obrazy w sensie wizyjnym - wykluczamy np. przebiegi czasowe
- ‘Rozwód’ Sztucznej Inteligencji i Rozpoznawania obrazów w latach 60-tych

Przecięcie:

- sztucznej inteligencji (ang. Artificial Intelligence),
- inteligencji obliczeniowej (ang. Computational Intelligence),
- przetwarzania sygnałów (ang. Signal Processing).

Inne powiązane dziedziny:

- uczenie maszynowe (ang. Machine Learning)
- psychologia poznawcza (ang. Cognitive Science)
- grafika komputerowa (ang. Computer Graphics)

- przetwarzanie sygnałów (ang. Signal Processing): obejmuje podejścia i metody modelujące obraz jako dwu- lub więcej-wymiarowy obraz w dziedzinie przestrzeni,
- widzenie komputerowe (ang. Computer Vision): zorientowana na analizę scen i obiektów 3D
- widzenie maszynowe (ang. Machine Vision): dotyczy głównie zastosowań przemysłowych, głównie kontrola jakości (QC)
- rozumienie obrazów (ang. Image Understanding): koncentruje się na interpretacji *całych* scen

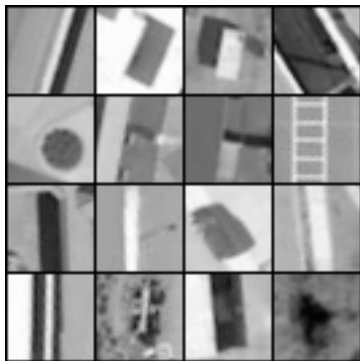
- Ponad 80% informacji wpływa do mózgu drogą wzrokową
- Multimedia, Internet, “cywilizacja medialna”
- Coraz szersze spektrum zastosowań
- Rosnąca dostępność
  - sprzętu (koszt kamery w telefonach komórkowych rzędu kilku dolarów),
  - bibliotek i APIs (np. API MS Kinect),



- niepełność/niepewność informacji wynikająca z:
  - procesu akwizycji obrazu (szumy, zniekształcenia optyczne, zniekształcenia chromatyczne, etc.)
  - natury procesu akwizycji obrazu (mapowanie 3D->2D, oświetlenie sceny)
- konieczność uwzględnienia wiedzy:
  - dziedzinowej (ang. *domain knowledge*)
  - zdroworozsądkowej (ang. *common sense knowledge*) zwłaszcza w uczeniu z informacji obrazowej.
- znaczna ilość przetwarzanej informacji => złożoność obliczeniowa/pamięciowa procesów związanych z analizą/rozpoznawaniem

# Przykład: rola wiedzy (kontekstu)





Despite this significant level of activity, the field [of image recognition and interpretation] remains a challenge. In particular, solutions of problems in image analysis are characterized by task-specific formulations, thus limiting the capability for advancement using the time-tried method of building a generalized body of results based on preceding accomplishments. **For the foreseeable future, the design of image analysis systems will continue to require a mixture of art and science.** (Gonzalez & Woods, str.657)

With no effort we scan a scene by directing our gaze at specific objects, discerning them individually despite the background of other objects, contours, shadows, and changes in illumination. The process is partially intentional, partially automatic, and **entirely amazing: ~~no machine can accomplish this, but the simplest insect can.~~** A single glance captures megabytes of data; we reduce this flood by singling out specific objects for attention (Hemmen, Domany, Cowan, 2001)





"construction worker in orange safety vest is working on road."



"two young girls are playing with lego toy."

... though not always



"a woman holding a teddy bear in front of a mirror."



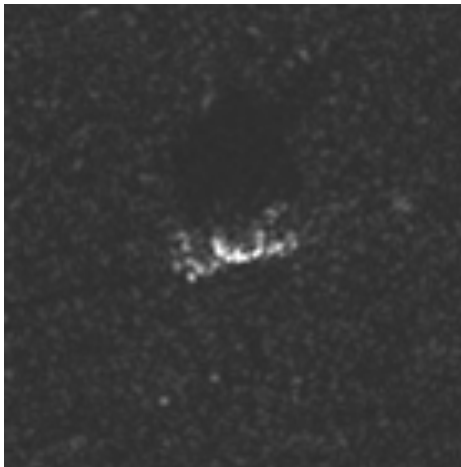
"a horse is standing in the middle of a road."

<http://cs.stanford.edu/people/karpathy/deepimagesent/>

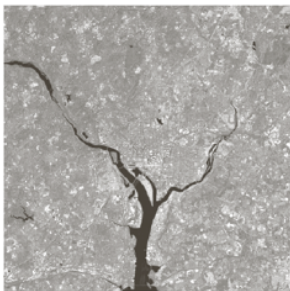
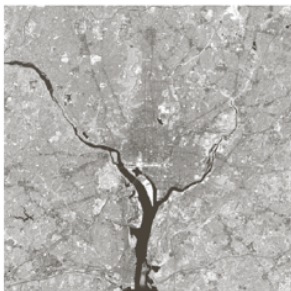
Obrazy mogą być pozyskiwane z wykorzystaniem różnych:

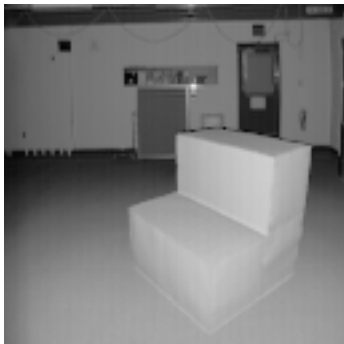
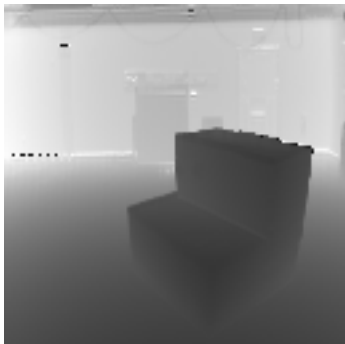
- urządzeń i technologii,
- modalności (podpasmo spektrum EM, i nie tylko),
- sposobach interakcji promieniowania z obrazowanym obiektem.





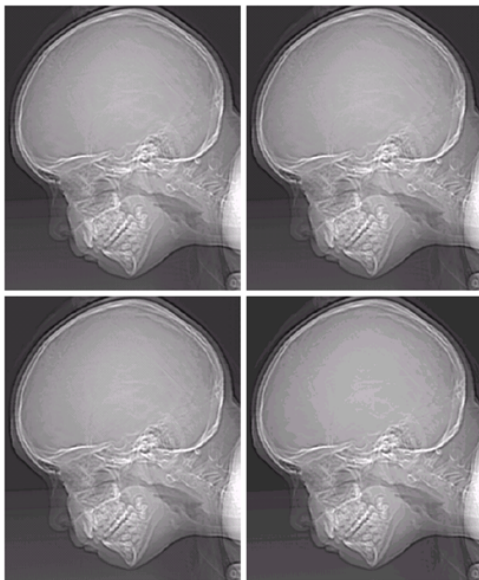








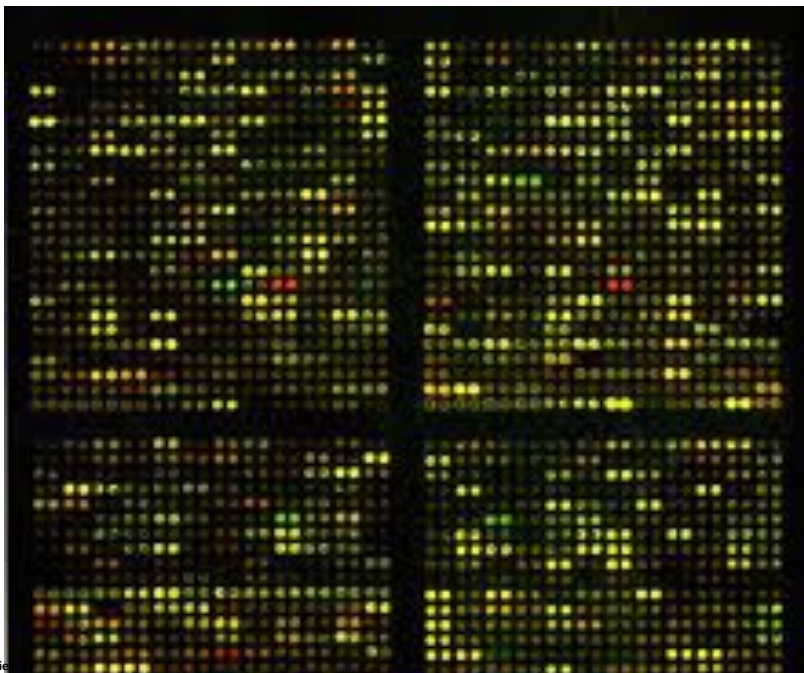


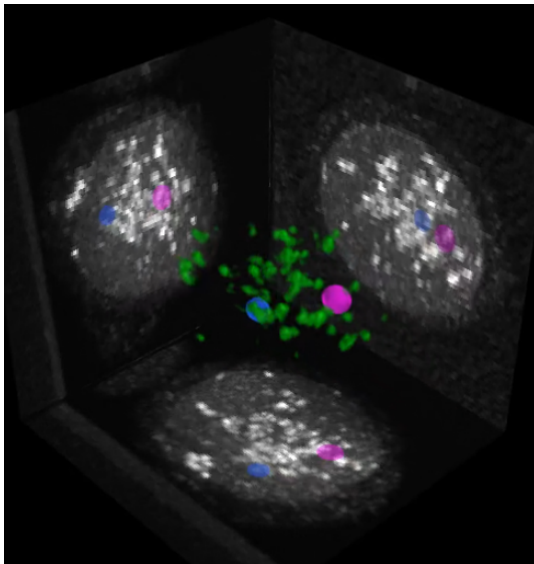


a b  
c d

**FIGURE 2.21**

(a)  $452 \times 374$ ,  
256-level image.  
(b)–(d) Image  
displayed in 128,  
64, and 32  
gray levels, while  
keeping the  
spatial resolution  
constant.





file:///img/anim\_Nokodazol.avi

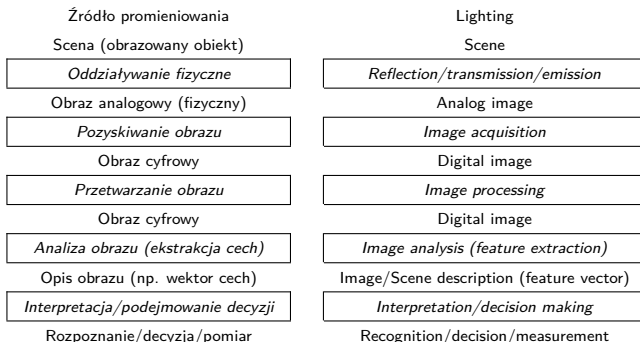
<https://pictures.lytro.com/>

## Typowe zadania przetwarzania i analizy obrazów:

- Polepszanie jakości obrazu:
  - zorientowane na percepcję człowieka (ang. image enhancement),
  - zorientowane na możliwe wierne odtworzenie obrazu fizycznego (ang. image restoration)
- Rozpoznawanie (ang. recognition)
- Identyfikacja (ang. identification)
- Lokalizacja (ang. localization)
- Estymacja położenia (ang. pose estimation),
- Analiza obrazu (ang. image analysis) - ekstrakcja innej informacji z obrazu (w tym cechy o nieintuicyjnej interpretacji)
- Zadania typowe dla sekwencji wideo: detekcja ruchu (ang. motion detection), śledzenie obiektów (ang. object tracking),
- Interpretacja scen (ang. scene interpretation)

## Kryteria oceny systemów RO:

- złożoność obliczeniowa i pamięciowa,
- jakość obrazu (wizualizacja)
- trafność (rozpoznawania),
- czułość, specyficzność (vide krzywa ROC)
- dokładność (np. lokalizacji),
- niezawodność,
- odporność na dane (obrazy) zaszumione/zniekształcone/uszkodzone,
- niezmienniczość T,S,R (+3D).

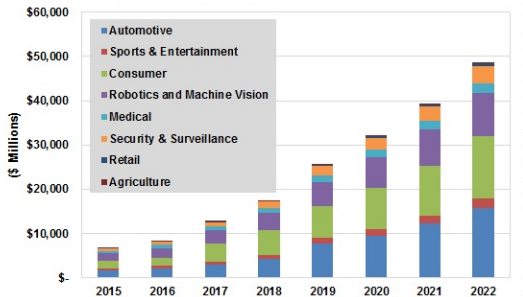




- Rozpoznawanie znaków: maszynowych i ręcznie pisanych,
- Kartografia, geologia, meteorologia: analiza zdjęć satelitarnych i lotniczych
- Medycyna, np. badania przesiewowe, analiza obrazów mikroskopowych,
- Archeologia: odtwarzanie zniszczonych eksponatów,
- Astronomia: odsiewanie artefaktów i automatyczna pre-klasyfikacja ciał niebieskich,
- Fizyka: polepszanie jakości obrazów eksperymentów w fizyce plazmowej i mikroskopii elektronicznej,
- Bioidentyfikacja na podstawie: odcisków palców, obrazu tęczówki oka, siatkówki oka, kształtu dłoni, sposobu chodzenia, odcisku ust, i innych cech biometrycznych
- Bezpieczeństwo: wspomaganie monitorowania obiektów (CCTV), kontrola bagażu, monitoring sterowanie ruchem miejskim,
- Przemysł: kontrola jakości produktów,
- Robotyka: roboty przemysłowe, roboty autonomiczne,
- Pojazdy samobieżne
- Rozrywkowe: np. Sony AIBO dog robot, MS Kinect
- Wojskowe
- ...



Computer Vision Revenue by Application Market, World Markets: 2015-2022



Source: Tractica


[RESEARCH](#) ▾ [CONSULTING](#) ▾ [RESOURCES](#) ▾ [ABOUT](#) ▾ [NEWSROOM](#)

June 20, 2016

**Computer Vision Hardware and Software Market to Reach \$48.6 Billion by 2022**

- Realizacje programistyczne (software)
  - Obecnie dominujące
  - Elastyczne, łatwo reprogramowalne, rozszerzalne, aktualizowalne ...
- Realizacje sprzętowe (hardware)
  - Np. FPGA (Field-Programmable Gate Array)
- Nowe architektury obliczeniowe
  - W szczególności karty graficzne (GPU), także w trybie General-Purpose Computing on GPUs (GPGPU)
  - Niezbędne w przypadku metod uczenia głębokiego (deep learning)

