

Pojęcie obrazu i jego akwizycja

WYKŁAD 1

Dla studiów niestacjonarnych 2021/2022
Semestr letni

Dr hab. Anna Korzyńska, prof. IBIB PAN

Co to jest obraz?

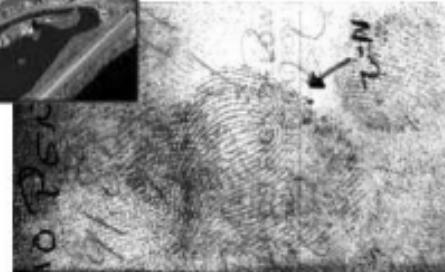
Obraz to wynik obserwacji świata przedstawiony na ograniczonej płaszczyźnie

(najczęściej prostokątnej)



CEL
tworzenia
obrazu

Artystyczny i użytkowy
(informacyjny), pamiątkowy



Poznawczy i użytkowy
(klasyfikacyjny), archiwizujący

Obraz to ...

Dwuwymiarowa funkcja mówiąca o wartości pewnej mierzalnej wielkości $f(x, y)$, (najczęściej wartość to określa intensywność światła/luminancję lub intensywność kolorów podstawowych) w miejscu o współrzędnych x, y na ograniczonej, dwuwymiarowej, powierzchni.



Obraz niesie informację o odwzorowywanej rzeczywistości lub o wizji autora, umieszczoną na ograniczonej, dwuwymiarowej przestrzeni (2D)

Obraz analogowy – przestrzeń jest spójna;
obraz cyfrowy – przestrzeń jest dyskretna



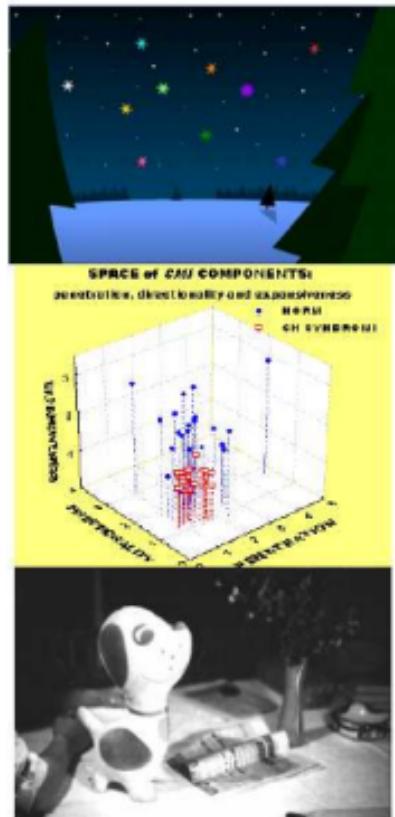
Obraz analogowy i cyfrowy

Obraz jako odwzorowanie rzeczywistości na ograniczonej dwuwymiarowej przestrzeni; może być odwzorowany na przestrzeni ciągłej (płótno, papier fotograficzny itp.) lub na przestrzeni dyskretnej (ekran monitora komputerowego, telewizora cyfrowego, itp.)



Obrazy w komputerze

- Narysowane narzędziami grafiki komputerowej (cel informacyjny lub artystyczny)
- Wygenerowane przez oprogramowanie jak wizualizacji informacji będącej rezultatem obliczeń lub zbierania informacji (cel naukowo-informacyjny)
- Pozyskane z rzeczywistości dzięki urządzeniom akwizycji obrazów: kamery, skanery, aparaty cyfrowe, medyczne urządzenia diagnostyczne, itp.. (cel poznawczy, archiwizacyjny, użytkowy np. klasyfikacja, detekcja, itp., lub cel artystyczny)



Jak powstał obraz cyfrowy



31 czerwca 1964 o 13:09 UT (9:09 AM EDT)

Vidicon B

Rozmiar w pikselach:
1150x1150

Obraz cyfrowy to informacja podwójnie dyskretna:

- odwzorowuje 3D na **dyskretną** i ograniczoną przestrzeń 2D



Księżyc

- informacja o intensywności (skwantowana)

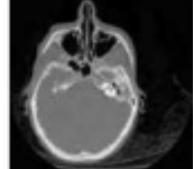
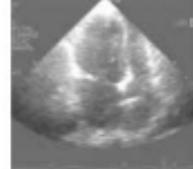
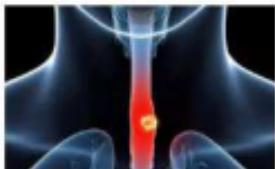


Pierwsze kolorowe zdjęcie z łazika Perseverance wysłane na Ziemię. Zostało wykonane przez jedną z kamer Hazcam. Źródło: NASA/JPL-Caltech.

Obraz cyfrowy w naukach technicznych i przyrodniczych to:

Zwarty, jednorodny i przestrzennie uporządkowany zbiór sygnałów:

- związanych z cechą/cechami pomiarowymi, na bazie których tworzymy obraz (natężenie fali elektromagnetycznej, akustycznej, wielkości nie falowe np. czas relaksacji)
- dostosowanych do materialnego nośnika obrazu (papieru, kliszy, dyskietki, pamięci dyskowej itp.)
- niosących informację o odwzorowywanej rzeczywistości



Podstawowe definicje wielkości opisujących obraz

Obraz to dyskretna dwuwymiarowa funkcja $f(x, y)$ określona na ograniczonym fragmencie płaszczyzny, której wartości f to **intensywność** (jasność, kolor) w tym punkcie (x, y) .

- Dla obrazów szaroodcieniowych wartość f to luminancja
- Dla obrazów kolorowych wartość f to wektor o trzech (lub więcej) składowych, określający np.: kolor w wybranej przestrzeni koloru $f=\{f_1, f_2, f_3\}$
- Dla obrazów wielomodalnych i multispektranych wartość f to wielowymiarowy wektor określający różne dane pomiarowe.

$f \in [L_{min}, L_{max}]$ - skala szarości/intensywności składowej pojedynczego kanału obrazu

$L_{min} = 0$, minimalna intensywność odpowiada czerni

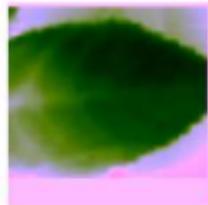
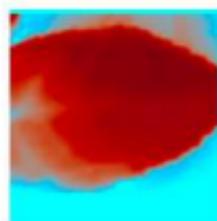
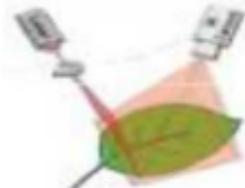
L_{max} = maksymalna intensywność odpowiada bieli

M - liczba poziomów szarości $M = L_{max} - L_{min} + 1$ $M = 2^k$

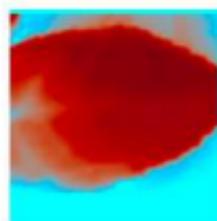
Obrazy konstruowane na podstawie wielu sygnałów

Obrazy multymodalne

Intensywność
odbicia światła
białego



Intensywność
odbicia światła
podczerwonego



Intensywność
rozproszenia fali
akustycznej



Obrazy konstruowane na podstawie wielu sygnałów

Obrazy multispektralny



Newada

Skanner multispektralny OPS z 8 zakresami spektralnymi od 0,52 do 2,40 um oraz radar obrazujący wykorzystujący długość fali 23,5 cm (1275 MHz) stosujący polaryzację poziomą (HH) kąt padania środka wybieranego pasa to 35 stopni, szerokość obrazowanego pasa wynosi 75

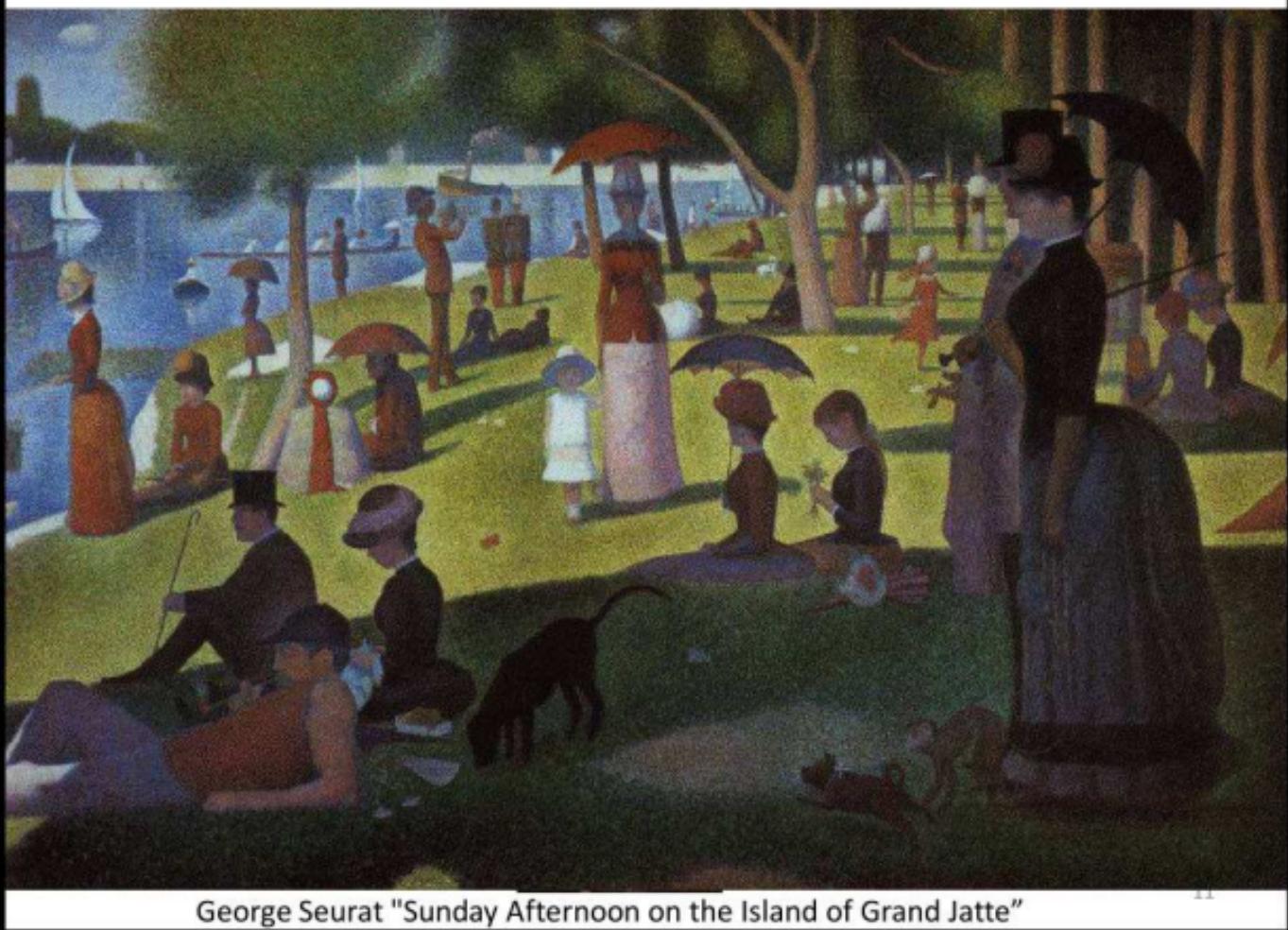
Rozdzielcość zasięgu i rozdzielcość azymutalna to 25 m.

JERS -1 (Japanese Earth Resources Satellite)



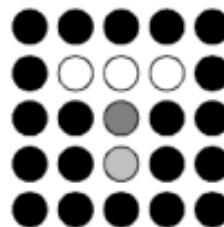
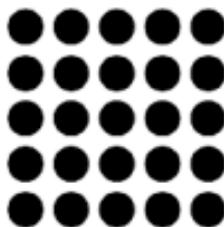
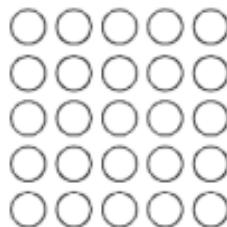
Galaktyki a520485

Zdjęcie przedstawia gromadę galaktyk „Abel 520” wykonane przez nałożenie składowych jej obrazów wykonanych w różnych długościach fal. Kosmiczny teleskop Chandra obserwował obiekt w zakresie promieniowania X (kolor czerwony); Teleskopy Kanadyjski, Francuski, wyspy Hawaii oraz teleskop Subaru w zakresie widzialnym (kolor żółty oraz pomarańczowy); Materię zabarwiono na kolor niebieski, z której to większości stanowi ciemna materia.



George Seurat "Sunday Afternoon on the Island of Grand Jatte"

Model obrazu cyfrowego



- $N_1=N_2=5$
- $M=4$ $L_{min}=0$ czerń; $L_{max}=3$ biel,
1 jasna szarość, 2 ciemna szarość

3 3 3 3 3

0 0 0 0 0

0 0 0 0 0

3 3 3 3 3

0 0 0 0 0

0 3 3 3 0

3 3 3 3 3

0 0 0 0 0

0 0 2 0 0

3 3 3 3 3

0 0 0 0 0

0 0 1 0 0

3 3 3 3 3

0 0 0 0 0

0 0 0 0 0

Obraz cyfrowy i piksel

Obraz cyfrowy niesie podwójną informację o:

- lokalizacji w przestrzeni obrazu wartość intensywności cechy/cech pomiarowych oraz o samej wartości tej cech
- czyli o wartości poziomu szarości, koloru w danym punkcie $f(i, j)=n$ gdzie n jest skalarem lub wektorem i $n \in \{0, 1, \dots L_{\max}\}$

Piksel to podstawowy element obrazu (*ang. Picture element*)



0,0						0,0						i					
							15	15	0	0	2						
j						j	13	13	15	0	0						
							0	0	7	14	14						

$N_1=5$
 $N_2=3$
 $L_{\min}=0$
 $L_{\max}=15$
 $M=16$

Rozdzielcość obrazu

- **Rozdzielcość przestrzenna** - określa stopień rozróżnialności detali; tym lepsza, im większa wartość N /na jednostkę długości. Jest wyrażana w jednostkach zwanych punktami na cal (*ang. Dot per inch*)
- **Rozdzielcość poziomów szarości** – określa ilość rozróżnianych poziomów szarości lub kolorów; tym lepsza, im większa wartość M . Jest dobierana tak, aby była potągą liczby 2.
- Inne rozdzielcości: **całkowita**, interpolowana, urządzeń prezentacji obrazu jak monitor, drukarka, ploter, urządzeń akwizycji.

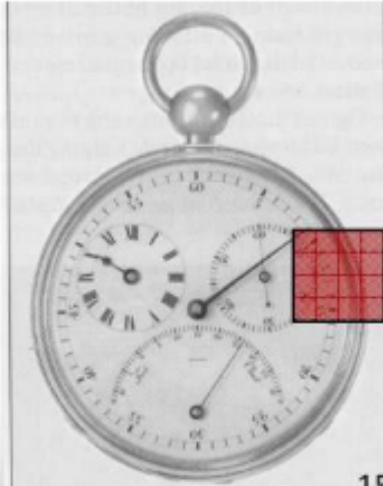
Rozdzielcość przestrzenna

Ilość szczegółów zapisanych w obrazie na jednostkę długości (najczęściej na cal)

Granica rozdzielczości to rozmiar szczegółów i obiektów, które są widoczne na obrazku.

O rozdzielczości mówimy na poziomie:

- Akwizycji (punkty na cal),
- Wyświetlania na ekranie (linie na cal, punkty),
- Drukowania (punkty na cal)



150dpi



72dpi
15

Wymiar obrazu w pikslach

Jeśli akwizycja była
przeprowadzona na matrycy
1200x1600

To jego rozdzielcość akwizycji
wynosi:

$$1200:4,3\text{ cala} = 280 \text{ ppi}$$

$$1960:7\text{ cali} = 280 \text{ ppi}$$



Rozdzielcość poziomów szarości



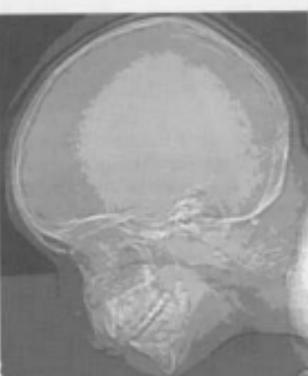
256



128



16



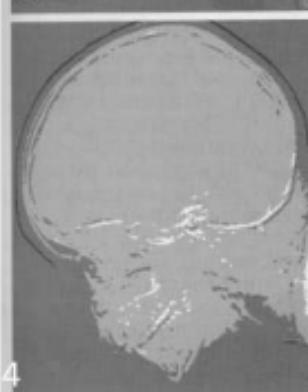
8



64



32



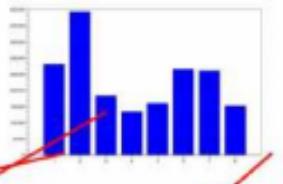
4



2

Histogram - rozkład ilości wystąpień pikseli o zadanych poziomach jasności w obrazie

obraz z 8 poziomami intensywności

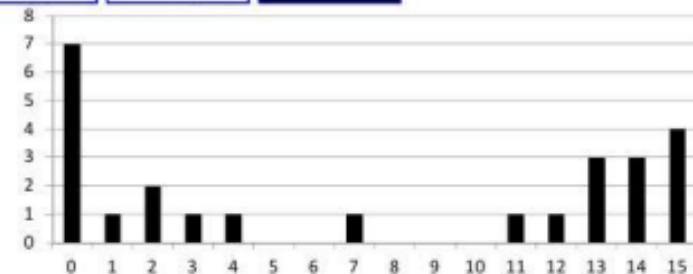


$M=8$,
 $L_{min}=0$,
 $L_{max}=7$



15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

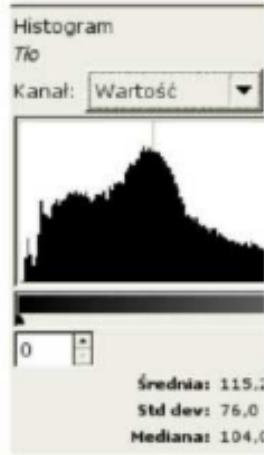
$M=16$,
 $L_{min}=0$,
 $L_{max}=15$



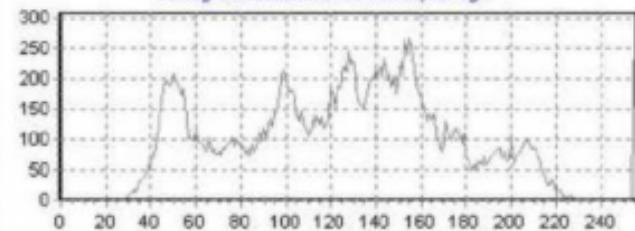
Histogram definicja

Histogram to wykres słupkowy przedstawiający ilość pikseli o każdej potencjalnej wartości występującej w obrazie.

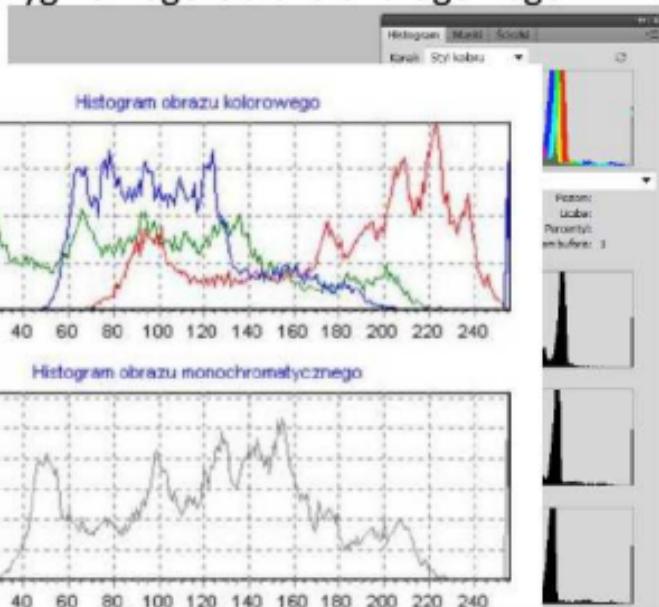
- Statystyka odzwierciedlająca rozkład jasności punktów w obrazie.
- Pewna estymata rozkładu jasności oryginalnego obrazu analogowego i rzeczywistości .



Histogram obrazu monochromatycznego



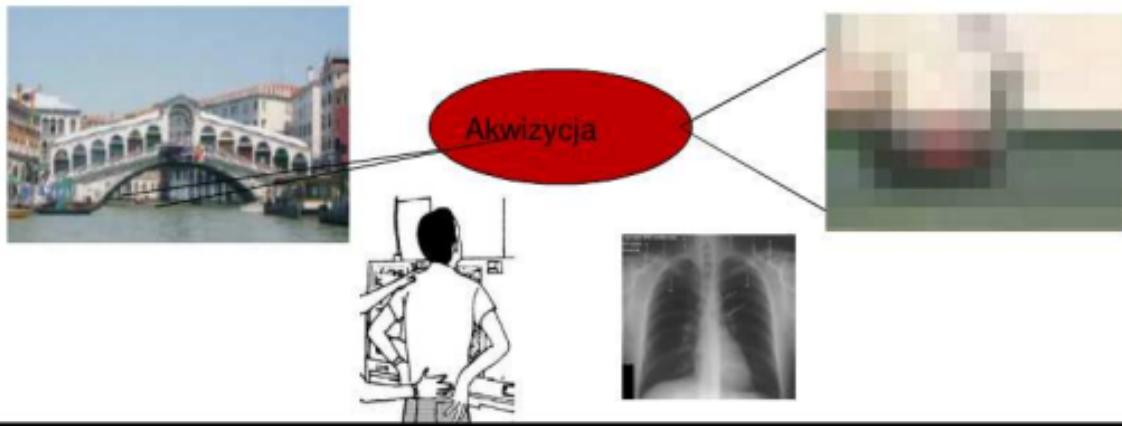
Pixel: Liczba: Percentyl: em buffer: 3



Akwizycja obrazu

Akwizycja (pozyskiwanie) obrazu cyfrowego

Akwizycja obrazu - przetworzenie **informacji o obiekcie fizycznym** do postaci zbioru danych dyskretnych (**obrazu cyfrowego**).



Akwizycja obrazu cyfrowego

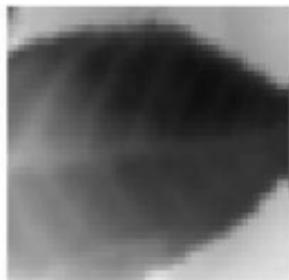
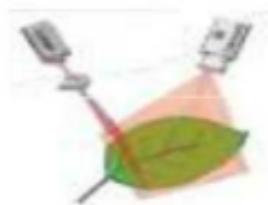
Akwizycja (pozyskiwanie, zbieranie) obrazu - przetworzenie informacji o fizycznym obiekcie lub scenie do postaci zbioru danych dyskretnych ($f(x,y)$) obraz cyfrowy) nadających się do zapisania w pamięci komputera, a następnie do wyświetlenia, drukowania i dalszego przetwarzania za pomocą odpowiedniego oprogramowania.

Elementy procesu akwizycji:

1. Oświetlenie obrazu.
2. Formowanie obrazu (optyczne).
3. Detekcja obrazu.
4. Formowanie wyjściowego sygnału z urządzenia (kamera, skaner)

Elementy procesu akwizycji

1. Formowanie obrazu/sygnału (optyczne, akustyczne, inne)
2. Detekcja sygnału przez czujniki analogowo-cyfrowe.
3. Formowanie kodu obrazu (kamera, skaner, MRI, USG).



Najważniejsze elementy procesu akwizycji (pozyskania)

- **Formowanie obrazu**
 - **Dyskretyzacja obrazu**, czyli dyskretyzacja funkcji $f(x,y)$ na dwóch poziomach:
 - przestrzenna (próbkowanie dziedziny funkcji)
 - amplitudowa (kwantyzacja wartości funkcji)
- Proces wiąże się z nieuniknioną i nieodwracalną utratą informacji o wartości cech pomiarowej poza węzłami siatki, i o dokładnej wartości w punktach siatki
 - Przekształcenie danych surowych w dane do zapisu obrazu
- **Kodowanie obrazu** zgodne z wybranym formatem zapisu (RAW, JPEG, TIFF)

Matryce CCD jako elementy pozyskania obrazu w aparacie fotograficznym cyfrowym

Matryca CCD (charge coupled device) ma płaską periodyczną strukturę (dyskretną), której podstawowe elementy mają $10-5\mu\text{m}$ długości i szerokości, która zamienia światło na prąd (sygnał analogowy) o napięciu proporcjonalnym do jasności. Prąd ten jest próbkowany przez konwerter A/D (framegrabber).



Nagroda Nobla za sensor CCD

Fizyka 2009

Willard S. Boyle (Kanada/USA) i George E. Smith (USA)
za "wynalezienie półprzewodnikowego obwodu
obrazującego - sensora CCD"

1969 - rozpoczęcie prac

Pierwszy układ wczytujący miał zaledwie osiem pikseli ulożonych w jednym rzędzie

1973 - 100 x 100 pikseli

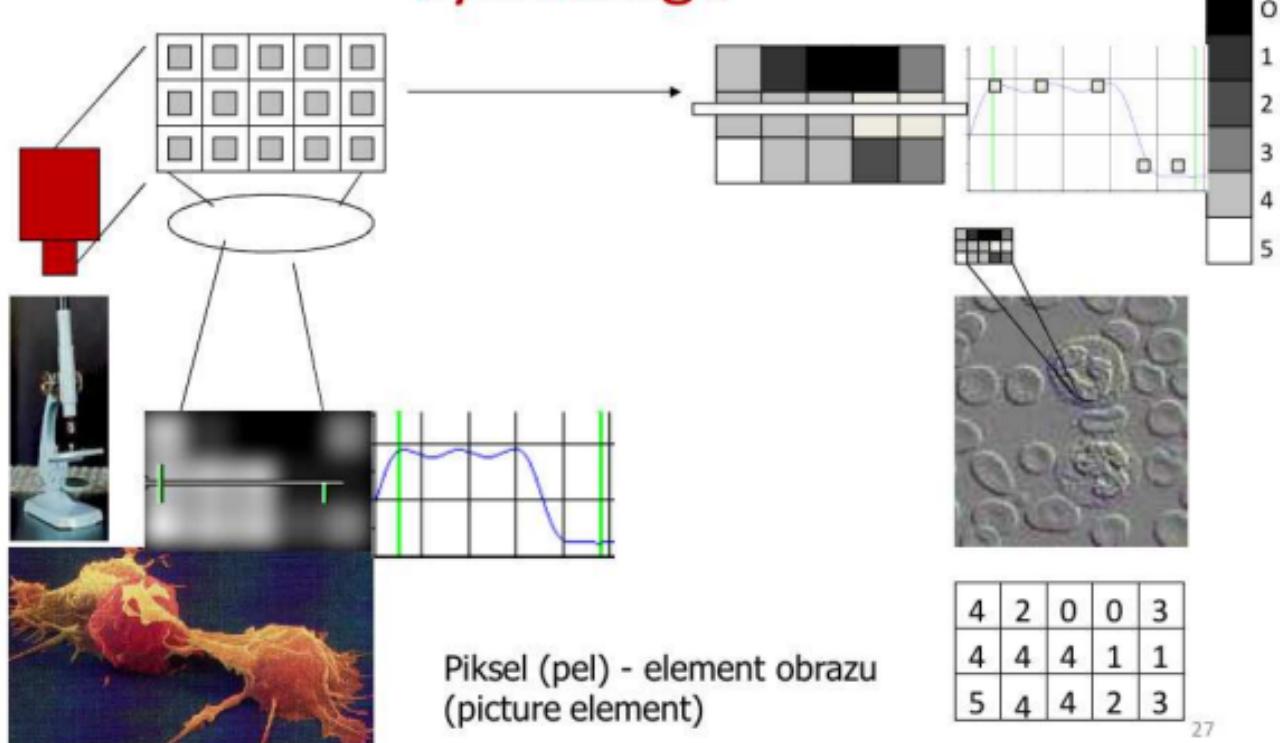


Pierwszy aparat fotograficzny wykorzystujący matrycę światłoczułą zamiast błony nazywał się MAVICA -1981

CCD sygnał zapisywany jako NTSC

570 linii i 490 pixels

Proces pozyskiwania obrazu cyfrowego



Najważniejsze elementy formowanie sygnału w procesie akwizycji

Dyskretyzacja obrazu to dyskretyzacja funkcji $f(x,y)$ na dwóch poziomach:

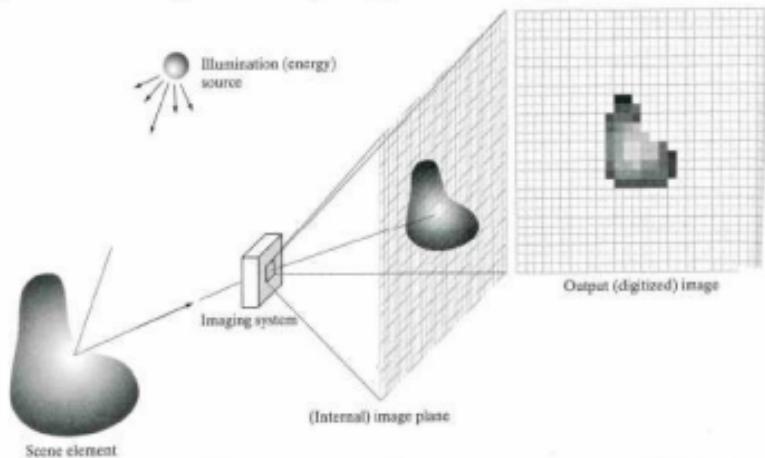
- przestrzenna (próbkowanie dziedziny funkcji)
- amplitudowa (kwantyzacja wartości funkcji)

Obraz jest więc strukturą **podwójnie dyskretną**

Próbkowanie sygnału

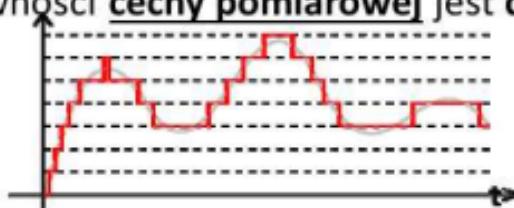
Obraz cyfrowy to funkcja $f(x, y)$ podwójnie dyskretna:

- odwzorowuje 3D na **dyskretną** i ograniczoną przestrzeń 2D



Kwantyzacja sygnału

- informacja o intensywności cechy pomiarowej jest dyskretna (skwantowana)



Kwantyzacja sygnału

- To proces polegający na przypisaniu wartościom analogowym do najbliższych poziomów reprezentacji cyfrowej (np.: 8-bitowego kodu)
- Niedokładności wynikające z zaokrągleń stanowią szum kwantyzacji

- W dyskretnych punktach przestrzeni dokonujemy pomiaru sygnału na podstawie, którego tworzymy obraz

- Wyszukanie wartości minimalnej i maksymalnej lub odczytanie z parametrów urządzeń detekcji sygnału

- Unormowanie sygnału przez przeliczenie wartości pomiarowych na zakres $<0,1>$

- Wybór zakresu bitów kodu, czyli ilości przedziałów kwantowania i przeliczenie unormowanego sygnału na kody

- Zaokrąglenie do wartości naturalnych

0	5	6	8.5
2.1	3.3	5.2	9.0
-8	-2	1.2	3
0.5	0.75	0.8	0.92
0.6	0.66	0.76	0.95
0.1	0.4	0.56	0.65
127.5	191.1	204.0	235.8
154.2	169.5	193.8	242.2
25.5	102.0	142.8	165.7
127	191	204	235
154	169	193	242
25	102	142	165

Zapis obrazu cyfrowego w pamięci komputera



InBmp	Histograms	Grid																	
186	186	187	187	188	188	188	189	189	190	190	190	190	190	190	190	190	190	190	
186	186	186	187	187	188	188	188	189	189	189	189	189	189	189	189	189	189	189	189
185	186	186	186	187	187	187	188	188	188	188	188	188	188	188	188	188	188	188	188
185	185	186	186	186	187	187	187	187	187	187	187	187	187	187	187	187	187	187	187
185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185
184	185	185	185	185	185	185	186	186	187	187	187	187	187	187	187	187	187	187	187
184	184	185	185	185	185	186	186	186	187	187	187	187	187	187	187	187	187	187	187
184	184	184	185	185	185	185	186	186	186	186	186	186	186	186	186	186	186	186	186
183	184	184	184	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185	185
183	183	184	184	184	184	185	185	185	185	185	185	185	185	185	185	185	185	185	185
183	183	183	184	184	184	184	184	184	184	184	184	184	184	184	184	184	184	184	184
182	183	183	183	183	183	183	183	183	183	183	183	183	183	183	183	183	183	183	183
182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182
182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182
181	182	182	182	182	182	183	183	183	183	183	183	183	183	183	183	183	183	183	183
181	181	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182
181	181	181	181	181	181	181	181	181	181	181	181	181	181	181	181	181	181	181	181

Charakterystyczne cechy procesu akwizycji i obrazu

Inherentne cechy procesu akwizycji

- Nakładanie szumu
(dla aparatu cyfrowego – szumu kwantyzacji, czyli odpowiednika ziarna dla tradycyjnej fotografii)
- Nakładanie zniekształceń
(dla aparatu cyfrowego - zniekształcenia obiektywów szerokokątnych, dystorsje soczewek, gorące i zimne piksle w matrycy)

Urządzenia akwizycji charakteryzuje się przez podanie informacji o poziomie szumu i zakłóceń



Szum

Szum (ang. random noise) to przypadkowe, niestacjonarne zakłócenia wartości sygnału:

- w czasie;
- w przestrzeni.

Nakładanie szumu na sygnał to inherentna cecha urządzeń elektronicznych, które albo transmitują albo odbierają sygnał w postaci prądu elektrycznego

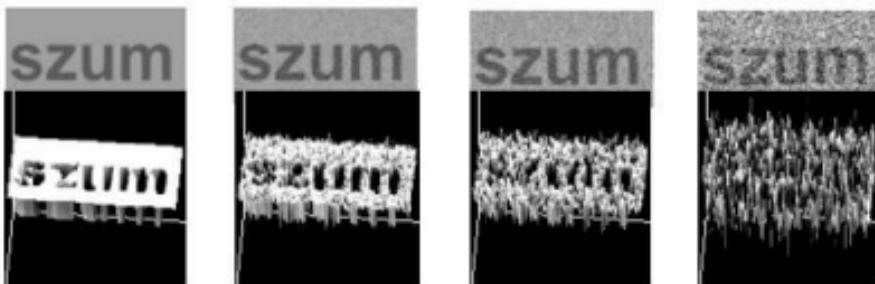
Szum powstaje na skutek

- Nałożenia termicznych ruchów elektronów w materii na mierzony sygnał, który jest zamieniany na prąd (np.: w przetwornikach zamieniających światło na prąd zgodnie ze zjawiskiem fotoelektrycznym)
- Niestabilności źródła formowania sygnału, np.: promieniowania X, fali akustycznej czy radiowej;

Zrozumienie zjawiska i jego przyczyn zwykle prowadzi do unikania zaszumienia, a nie do jego likwidacji czy choćby redukcji

Miara szumu

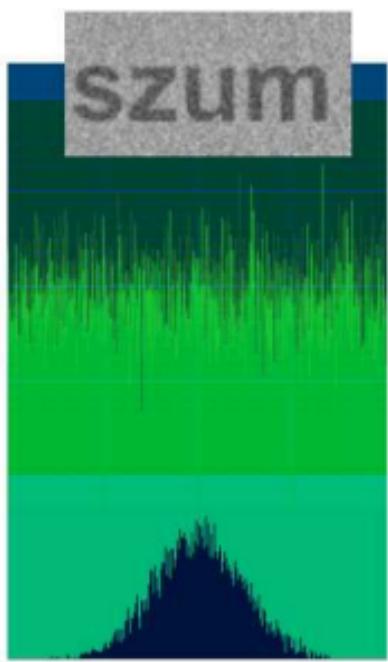
Stosunek sygnału do szumu (ang. signal to noise ratio – SNR), jednostka decybeli dB.



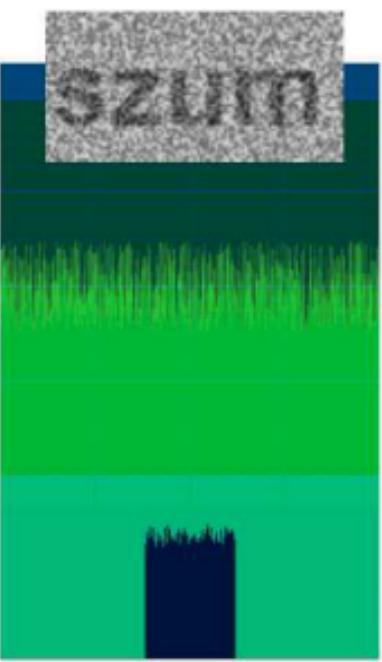
Brak „odczuwania” szumu oznacza, że jego stosunek do sygnału jest taki, że nasze sensory (wzrokowe, słuchowe) odbierają sygnał a pomijają szum.

Charakterystyka szumu w obrazie

szum



gaussian noise



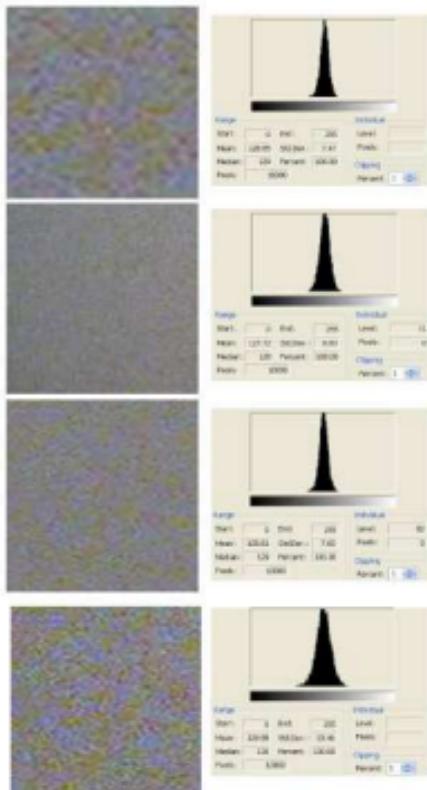
uniform noise



salt and pepper noise

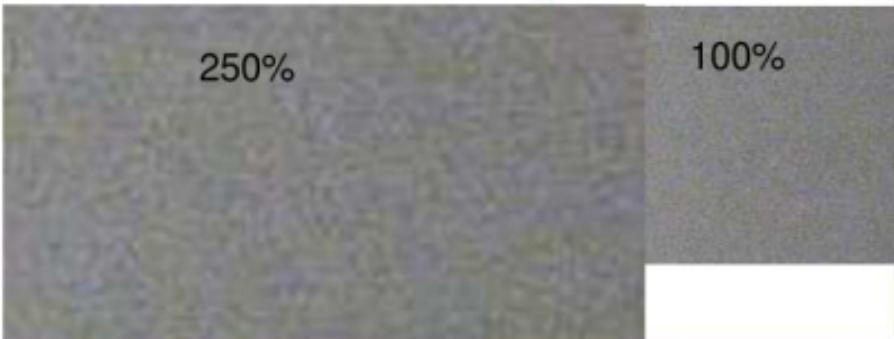
Typy szumu w obrazie

- Szum niskoczęstotliwościowy
(ang. *coarser texture* - chropowaty
gruboziarnisty)
- Szum wysokoczęstotliwościowy
(ang. *fine-grained* - drobny)
- Szum o małej amplitudzie
(ang. *smoother texture* - gładki)
- Szum o dużej amplitudzie
(ang. *rougther texture* - szorstki)

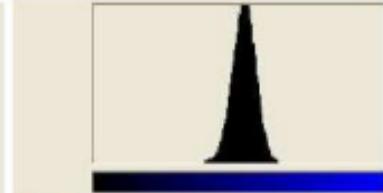
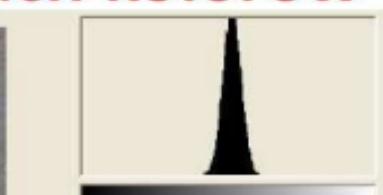


Szum w obrazie: w luminancji i w kanałach kolorów

250%



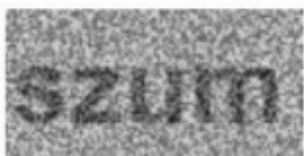
100%

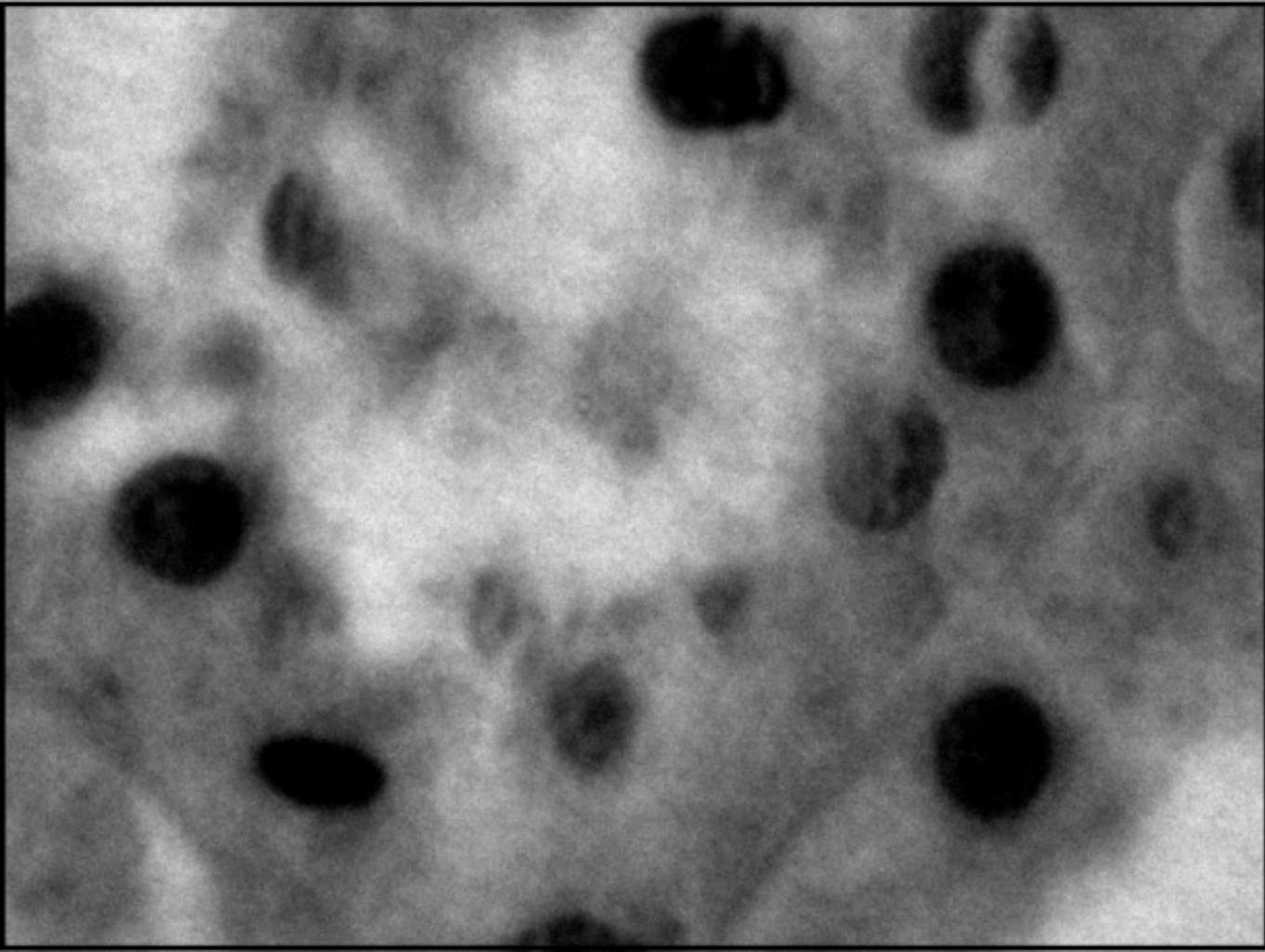


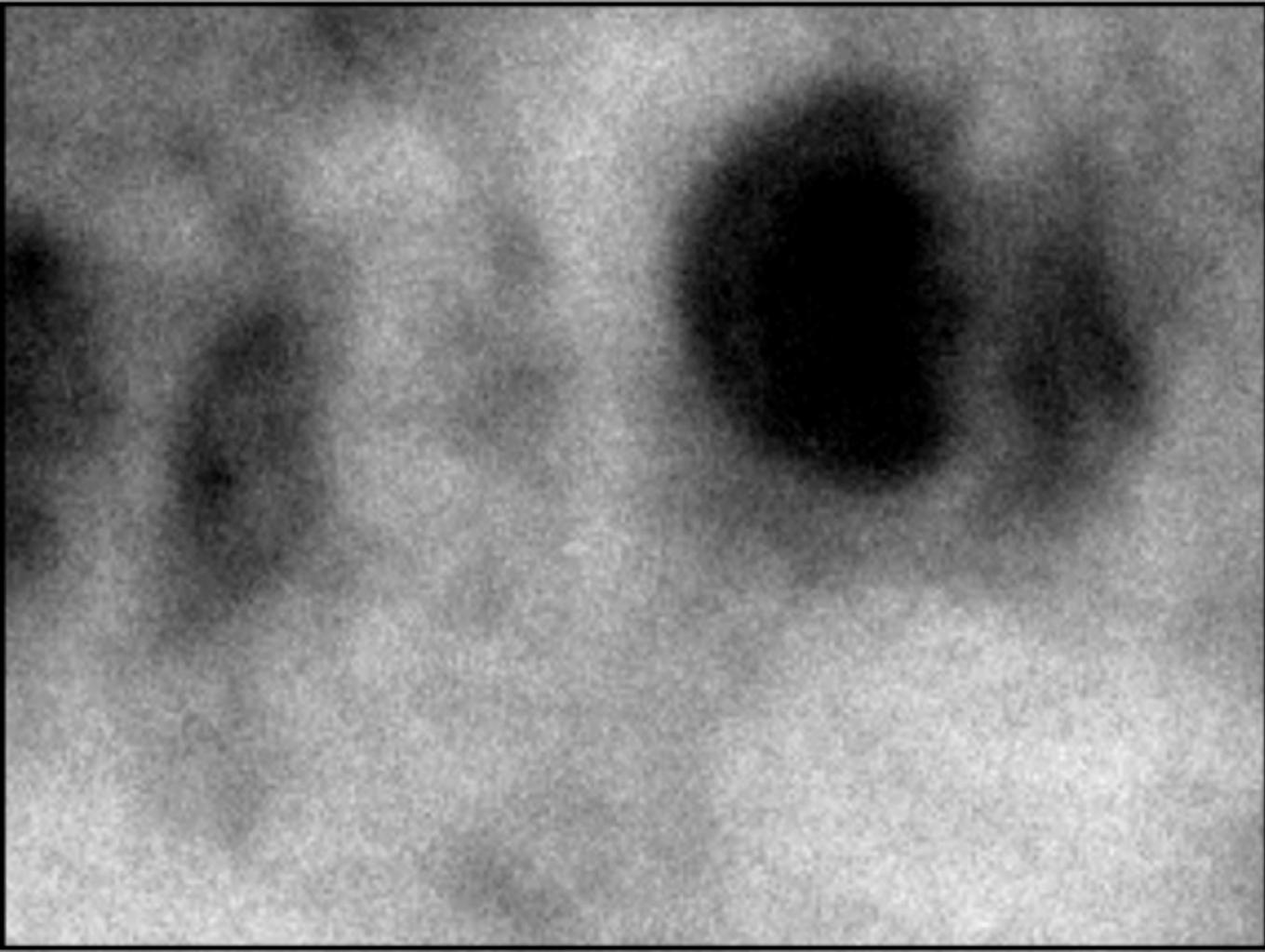
Na jakich obrazach występuje dużo szumu?

- Rentgenowskich, w mikroskopach fluorescencyjnych, ze znacznikami emitującymi wąski zakres fal elektromagnetycznych – powód: mało fotonów (charakterystyka gaussowska)
- Mikroskop elektronowy skanujący – powód: niestabilność źródła światła - długie czasy zbierania (charakterystyka niegaussowska)
- Astronomicznych - zbieranych przy pomocy kamery integrującej sygnał

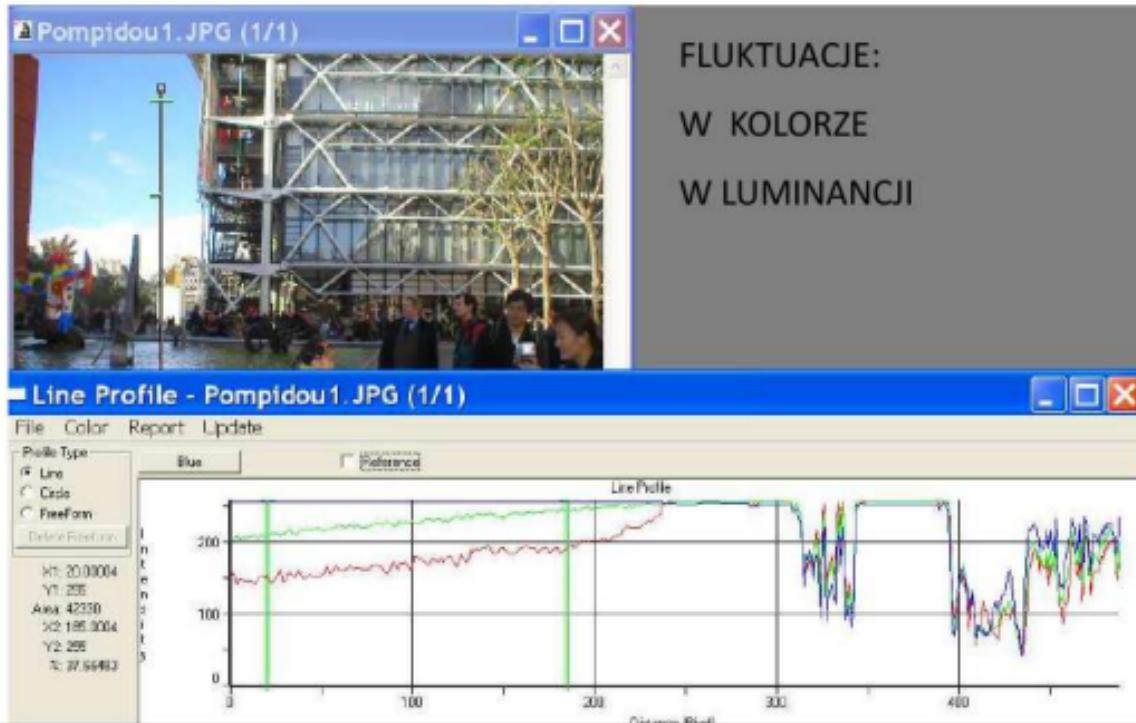
szum







Szum nakładany na obraz przez cyfrowy aparat fotograficzny



Szum nakładany na obraz powstający w cyfrowym aparacie

wynika z:

- Parametrów technicznych matrycy fotoczułej (wielkość sensora)
- Niestabilności światła lub z jego niewielkiej ilości (nocą)
- Zjawisk towarzyszących zamianie światła na prąd (ang. *banding noise, truncated, ..*)
- Nałożenia termicznych ruchów elektronów na mierzony sygnał



Szum periodyczny
ang. periodic noise

Szum w obrazach cyfrowych zależny od:

- typu (egzemplarza) aparatu cyfrowego
- od ustawień przy wykonywaniu zdjęcia
(czułość ISO, czas naświetlenia)
- jasności fotografowanych obiektów i
obszarów (w cieniach szumu jest więcej niż
w tonach jasnych)
- temperatury otoczenia - im wyższa tym
silniej widać
(ang. Fixed pattern noise)

Rozłożenie szumu na obrazie



250%

250%

100%

100%

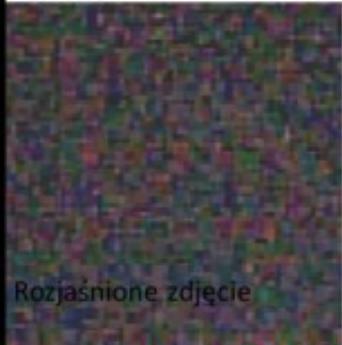
100%

100%

250%

250%

250%



Rozjaśnione zdjęcie

UWAGA! Odwrotnie w klasycznej fotografii negatywowej

Redukcja szumu

- Na poziomie akwizycji:
 - Wybór maksymalnie długiego czasu naświetlania
 - Wybór niskich czułości ISO
 - Chłodzenie matrycy
 - Wykonywanie zdjęć wielokrotnych (braketing)
 - Wykonywanie zdjęć lekko prześwietlonych (HighKey)
- Na poziomie przetwarzania (obróbki) obrazów:
 - Uśrednianie zdjęć wielokrotnych
 - Filtrowanie
- **Uwaga:** Nie redukujemy szumu przez
 - Rozjaśnianie obrazu w miejscach niedoświetlenia
 - Rozjaśnianie cieni w obrazie

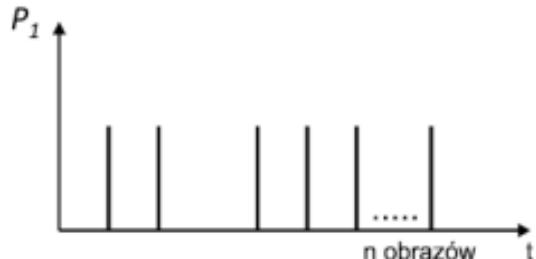
Uśrednianie czasowe (dotyczy obrazów statycznych)

$N \times N$ – liczba pikseli w obrazie.

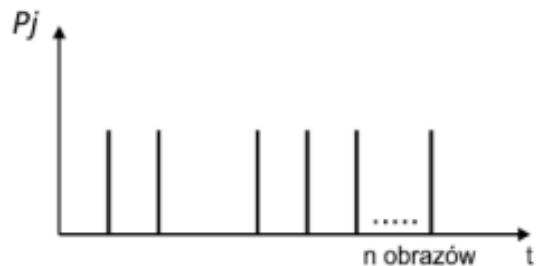
Np. dla $N=4$ liczba pikseli w obrazie wynosi
 $N \times N = 16$.

15	14	10	0
12	11	5	0
11	4	3	2
3	2	0	1

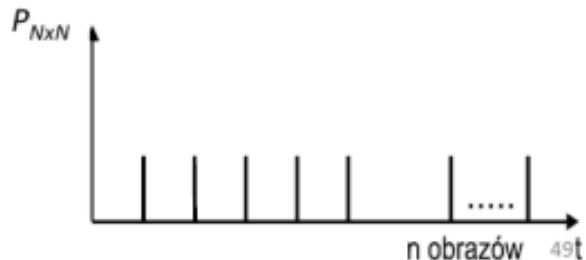
P_1 - wartość pierwszego piksela obrazu



P_j = wartość j-tego piksela obrazu



P_{NxN} - wartość ostatniego piksela



$$P_{js} = \frac{\sum_{i=1}^n P_{ji}}{n}$$

P_{ji} - wartość j-tego piksela obrazu w i -tej chwili czasowej

P_{js} - wartość średnia j-tego piksela obrazu w n chwilach czasowych

n - liczba pojawięń się obrazu $i = 1, n$

NxN – liczba pikseli w obrazie $j = 1, N \times N$

$$[\quad + \quad + \dots + \quad] / n = \quad$$
$$q(i, j) = \frac{1}{n} \sum_{l=1}^n p_l(i, j)$$

Stosowane do redukcji szumu (obrazy astronomiczne)



Uśrednianie przestrzenne (obrazy statyczne oraz zmienne w czasie)

15	13	15
14	0	15
12	12	14

⇒

15	13	15
14	12	15
12	12	14

$$P_{\text{sr}} = \frac{\sum_i^n P_i}{n}$$

$n = 9$ - otoczenie 8-spójne

	13	
14	0	15
	12	

n - liczba pikseli otoczenia
(wraz z pikselem
przetwarzanym)

$n = 5$ - otoczenie 4-spójne

MATERIAŁ DO WYKŁADU I ĆWICZEŃ LABORATORYJNYCH:

R. Tadeusiewicz, P. Korohoda: **Komputerowa analiza i przetwarzanie obrazów**, Kraków 1997. <http://winntbg.bg.agh.edu.pl/skrypty2/0098/> (także plik ..\POBD\2019-2020\Materialy\TadKoroh.pdf na UBIKu)

Wykład

Materiał podstawowy:

M. Doros, Przetwarzanie obrazów, Skrypt WSISiZ

Ćwiczenia Laboratoryjne

Materiał podstawowy:

M.Doros, A. Korzyńska, M.Przytulska, H.Goszczyńska: „Przetwarzanie Obrazów, ćwiczenia laboratoryjne”, Skrypt WSISiZ

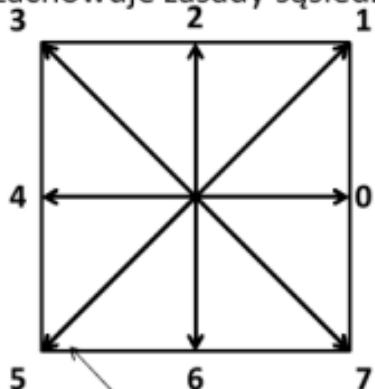
Rzadziej stosowane siatki

- Siatka prostokątna
- Siatka sześciokątna (heksagonalna)
- Siatka trójkątna

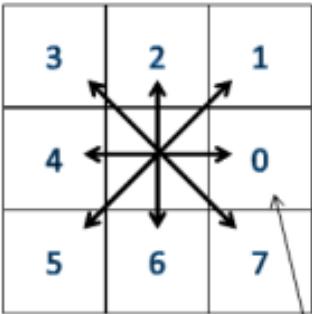
Rodzaje sąsiedztwa: np. 8-spójne, 4-spójne

Dualizm oczko - węzeł (siatka prostokątna)

zachowuje zasady sąsiedztwa np. ośmiospójnego.

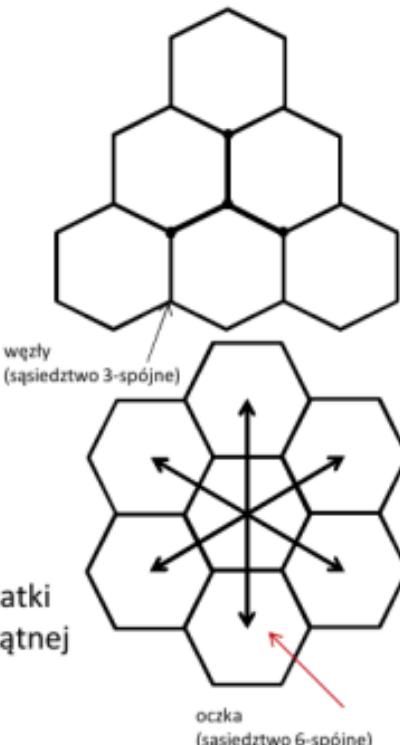


węzły siatki prostokątnej



oczka siatki prostokątnej

Piksel może być skojarzony z węzłem
lub z oczkiem siatki



Sąsiedztwo

- **Siatka dyskretna** (*discrete net*) - wzorzec według którego dokonywana jest *dyskretyzacja przestrzenna* obrazu; linie, oczka, węzły
- **Siatka prostokątna** – najczęściej stosowana: oczko siatki jest kwadratem
- **Piksel** – podstawowy element obrazu; odniesienie do oczka lub węzła siatki

Paradoks spójności

0	1	0	0
0	1	0	0
0	0	1	0
0	0	0	1

1 - obiekt spójny

2 - tło: spójne(?)
niespójne(?)

Przeciwdziałanie:

przypisanie różnych rodzajów
sąsiedztwa pikselom obiektu i tła

Akwizycja obrazu na przykładzie aparatu fotograficznego; operacje na obrazach (I)

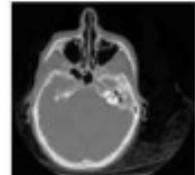
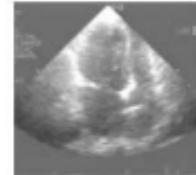
WYKŁAD 2
Dla studiów niestacjonarnych
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

Obraz w naukach technicznych i przyrodniczych to:

Zwarty, jednorodny i przestrzennie uporządkowany zbiór sygnałów:

- związanych z cechą/cechami pomiarowymi, na bazie których tworzymy obraz (natężenie fali elektromagnetycznej, akustycznej, wielkości nie falowe np. czas relaksacji)
- dostosowanych do materialnego nośnika obrazu (papieru, kliszy, dyskietki, pamięci dyskowej itp.)
- niosących informację o odwzorowywanej rzeczywistości



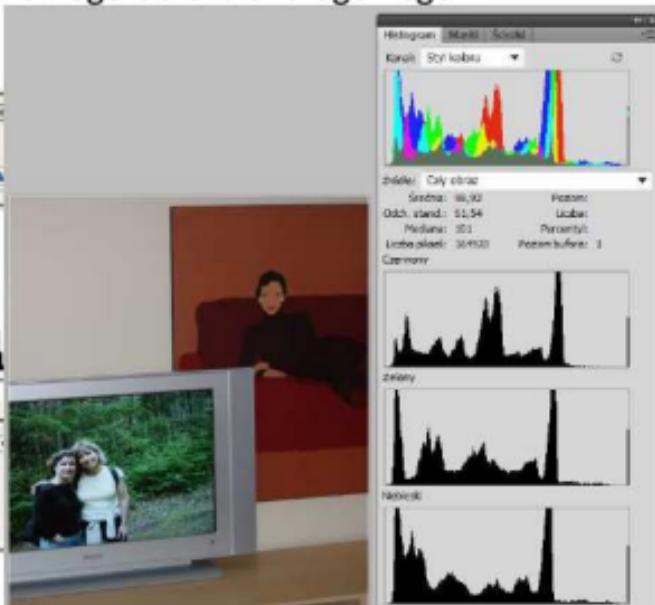
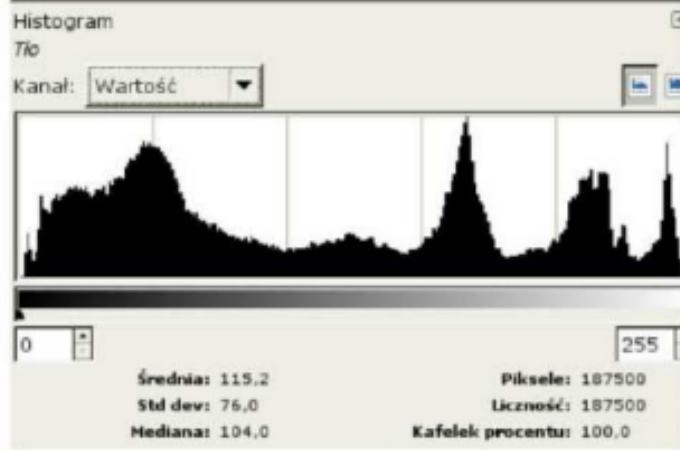
Rozdzielcość obrazu

- **Rozdzielcość przestrzenna** - określa stopień rozróżnialności detali; tym lepsza, im większa wartość N/na jednostkę długości. Jest wyrażana w jednostkach zwanych punktami na cal (*ang. Dot per inch*)
- **Rozdzielcość poziomów szarości** – określa ilość rozróżnianych poziomów szarości lub kolorów; tym lepsza, im większa wartość M. Jest dobrana tak, aby wartość M była potągą liczby 2 – co ułatwia kodowanie.
- Inne rozdzielcości: **całkowita**, interpolowana, urządzeń prezentacji obrazu jak monitor, drukarka, ploter, urządzeń akwizycji.

Histogram definicja

Histogram to wykres słupkowy przedstawiający ilość pikseli o każdej potencjalnej wartości poziomu szarości lub intensywności koloru występującej w obrazie.

- Statystyka odzwierciedlająca rozkład jasności punktów w obrazie.
- Pewna estymata rozkładu jasności oryginalnego obrazu analogowego i rzeczywistości .



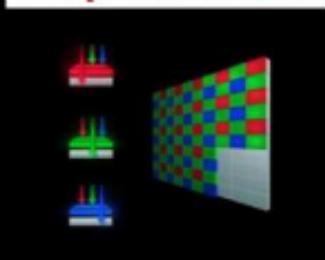
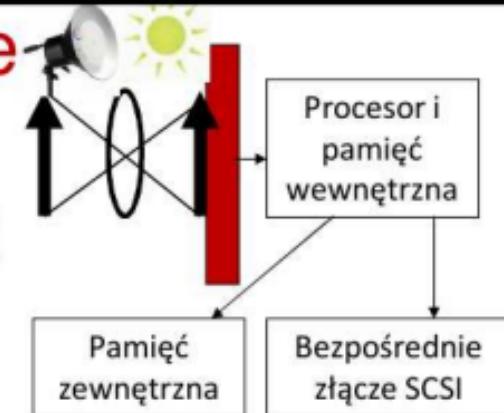
Akwizycja obrazu cyfrowego

Akwizycja (pozyskiwanie, zbieranie) obrazu - przetworzenie informacji o fizycznym obiekcie lub scenie do postaci zbioru danych dyskretnych ($f(x,y)$) obraz cyfrowy) nadających się do zapisania w pamięci komputera, a następnie do wyświetlenia, drukowania i dalszego przetwarzania za pomocą odpowiedniego oprogramowania.

Elementy procesu akwizycji:

1. Oświetlenie obrazu.
2. Formowanie obrazu (optyczne).
3. Detekcja obrazu.
4. Formowanie wyjściowego sygnału z urządzenia (kamera, skaner)

Sprzętowo –programowe formowanie obrazu w aparacie fotograficznym

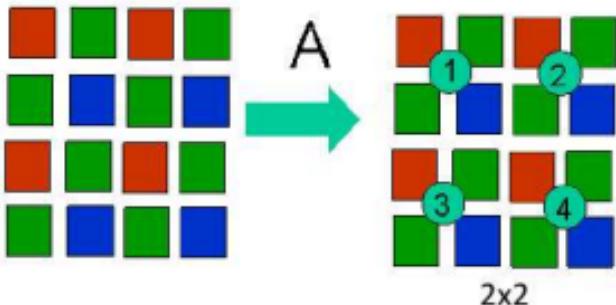


$2^{12} = 4096$ poziomów intensywności dla każdego kanału

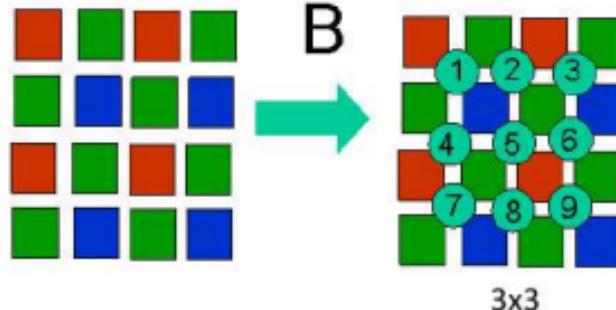
Droga światła i demozaikowanie



DSLR Digital Camera



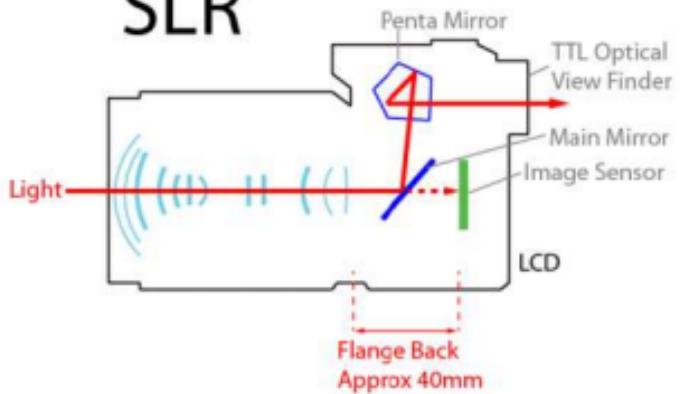
2x2



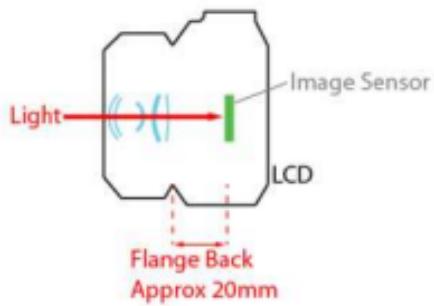
3x3

- | | |
|-------------------------------|-----------------------------------|
| 1. Viewfinder (eyepiece lens) | 9. Electronics |
| 2. Pentaprism | 10. Autofocus system |
| 3. Focusing screen | 11. Reflex and relay mirror |
| 4. Condenser lens | 12. Focusing elements |
| 5. Color and infrared filter | 13. Aperture |
| 6. Digital sensor | 14. Zoom elements |
| 7. Shutter | 15. Front line gathering elements |
| 8. Display | 16. Batteries |

SLR



Mirrorless



Najważniejsze elementy formowanie sygnału w procesie akwizycji

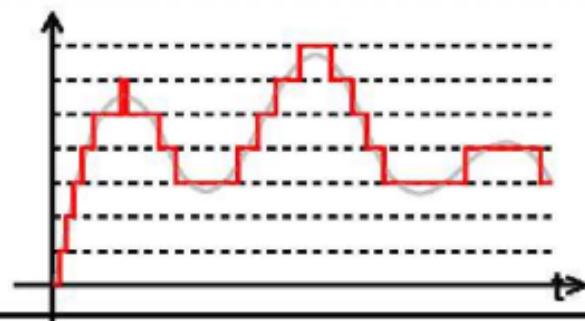
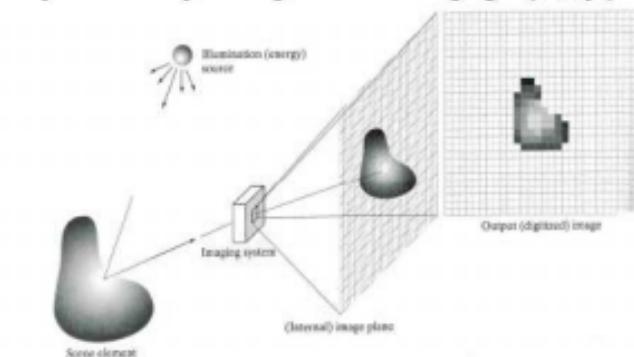
Dyskretyzacja obrazu to dyskretyzacja funkcji $f(x,y)$ na dwóch poziomach:

– przestrzenna

(próbkowanie dziedziny funkcji)

– amplitudowa

(kwantyzacja wartości funkcji)



Inherentne cechy procesu akwizycji

- Nakładanie szumu i delikatnego rozmycia krawędzi (dla aparatu cyfrowego – szumu kwantyzacji, czyli odpowiednika ziarna dla tradycyjnej fotografii oraz s-owaty kształt linii profilu przechodzącej przez krawędzie)
- Nakładanie zniekształceń (dla aparatu cyfrowego - zniekształcenia obiektywów szerokokątnych, dystorsje soczewek, gorące i zimne piskle w matrycy, halo na około obiektów fotografowanych na tle nieba)

Urządzenia akwizycji charakteryzuje się przez podanie informacji o poziomie szumu i rozmycia krawędzi oraz zakłóceń radiomatrycznych, geometrycznych i chromatycznych

Zniekształcenia obrazu

w procesie akwizycji:

- Zniekształcenia **radiometryczne** to zniekształcenia odczytu wartości cechy, na podstawie której tworzymy obraz;
 - nierównomiernością oświetlenia,
 - błędami konwersji oświetlenie – sygnał elektryczny (tzn. błędami detekcji, uszkodzeniami przetworników)
- Zniekształcenia **geometryczne** to zniekształcenia odczytu położenia odczytanych wartości względem siebie

dla obrazów kolorowych i wielospektralnych

- Zniekształcenia (aberracje) **chromatyczne** to zniekształcenia wynikające z charakterystyki współczynnika odbicia lub przejścia przez ośrodek fali elektromagnetyczne o różnej długości

Celem wstępnego przetwarzania obrazu jest redukcja zniekształceń obrazu powstających w procesie akwizycji

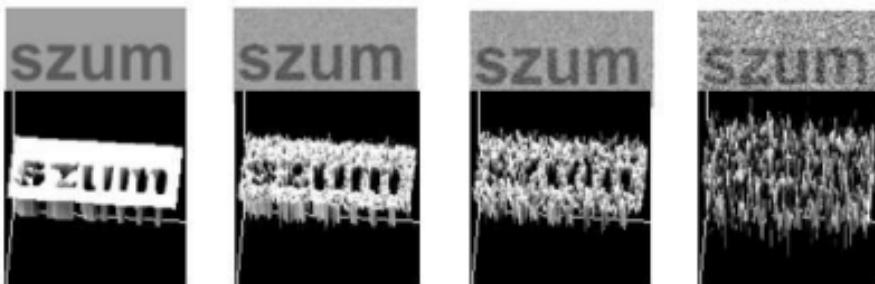
Szum powstaje na skutek

- Nałożenia termicznych ruchów elektronów w materii na mierzony sygnał, który jest zamieniany na prąd (np.: w przetwornikach zamieniających światło na prąd zgodnie ze zjawiskiem fotoelektrycznym)
- Niestabilności źródła formowania sygnału, np.: promieniowania X, fali akustycznej czy radiowej;

Zrozumienie zjawiska i jego przyczyn zwykle prowadzi do unikania zaszumienia, a nie do jego likwidacji czy choćby redukcji

Miara szumu

Stosunek sygnału do szumu (ang. signal to noise ratio – SNR), jednostka decybeli dB.



Brak „odczuwania” szumu oznacza, że jego stosunek do sygnału jest taki, że nasze sensory (wzrokowe, słuchowe) odbierają sygnał a pomijają szum.

Szum nakładany na obraz powstający w cyfrowym aparacie

wynika z:

- Parametrów technicznych matrycy fotoczułej (wielkość sensora)
- Niestabilności światła lub z jego niewielkiej ilości (nocą)
- Zjawisk towarzyszących zamianie światła na prąd (ang. *banding noise, truncated, ..*)
- Nałożenia termicznych ruchów elektronów na mierzony sygnał

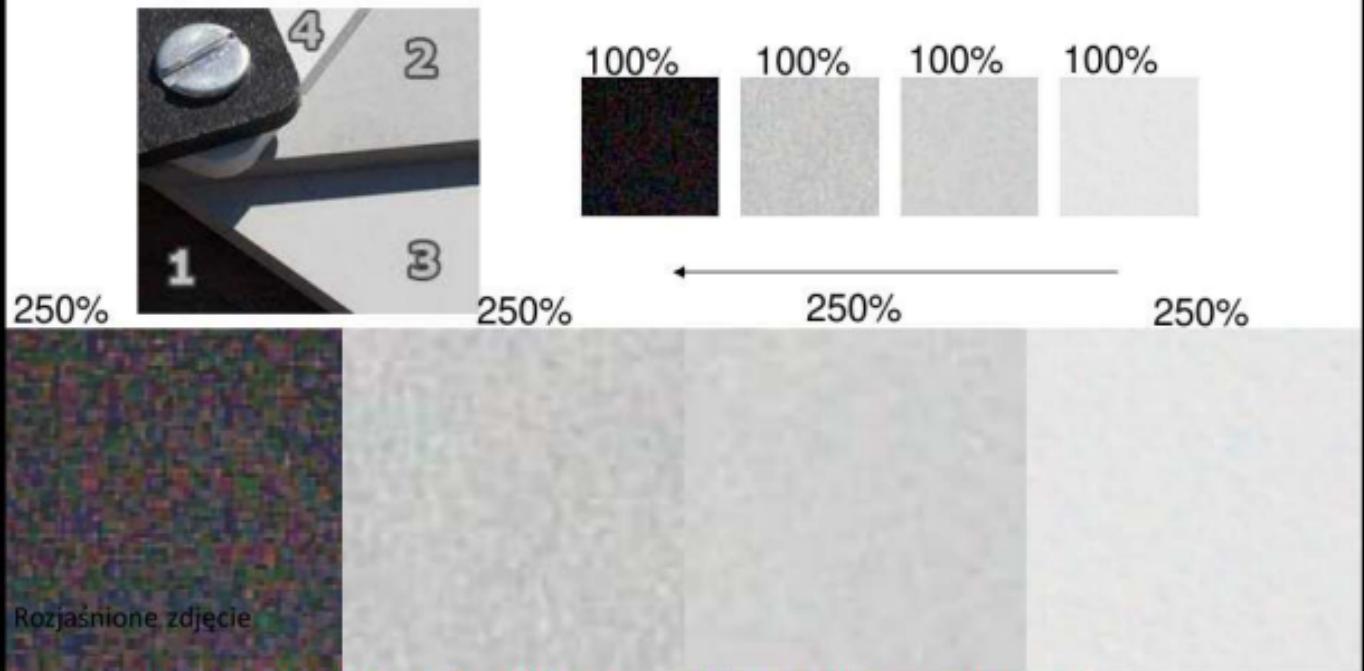


Szum periodyczny
ang. periodic noise

Szum w obrazach cyfrowych zależny od:

- typu (egzemplarza) aparatu cyfrowego
- od ustawień przy wykonywaniu zdjęcia
(czułość ISO, czas naświetlenia)
- jasności fotografowanych obiektów i
obszarów (w cieniach szumu jest więcej niż
w tonach jasnych)
- temperatury otoczenia - im wyższa tym
silniej widać
(ang. Fixed pattern noise)

Rozłożenie szumu na obrazie



UWAGA! Odwrotnie w klasycznej fotografii negatywowej

Redukcja szumu

- Na poziomie akwizycji:
 - Wybór maksymalnie długiego czasu naświetlania
 - Wybór niskich czułości ISO
 - Chłodzenie matrycy
 - Wykonywanie zdjęć wielokrotnych (braketing)
 - Wykonywanie zdjęć lekko prześwietlonych (HighKey)
- Na poziomie przetwarzania (obróbki) obrazów:
 - Uśrednianie zdjęć wielokrotnych
 - Filtrowanie
- **Uwaga:** Nie redukujemy szumu przez
 - Rozjaśnianie obrazu w miejscach niedoświetlenia
 - Rozjaśnianie cieni w obrazie

Uśrednianie czasowe (dotyczy obrazów statycznych)

$N \times N$ – liczba pikseli w obrazie.

Np. dla $N=4$ liczba pikseli w obrazie wynosi $N \times N = 16$.

15	14	10	0
12	11	5	0
11	4	3	2
3	2	0	1

$$P_{j's} = \frac{\sum_{i=1}^n P_{ji}}{n}$$

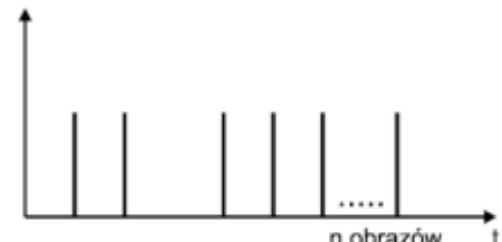
P_{ji} – wartość j-tego piksela obrazu w i-tej chwili czasowej

$P_{j's}$ – wartość średnia j-tego piksela obrazu w n chwilach czasowych

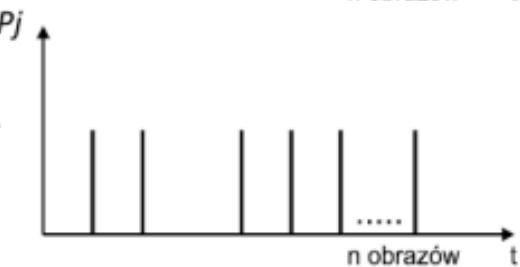
n – liczba pojawiń się obrazu $i = 1, n$

$N \times N$ – liczba pikseli w obrazie $j = 1, N \times N$

P_1 – wartość pierwszego piksela obrazu

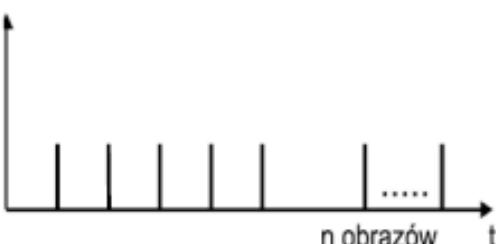


P_j = wartość j-tego piksela obrazu



P_{NxN}

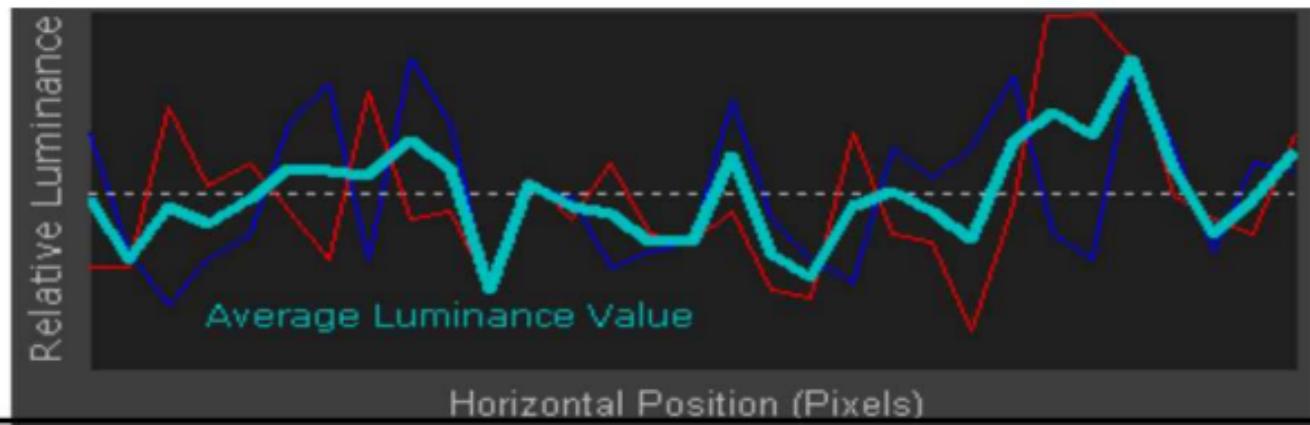
P_{NxN} – wartość ostatniego piksela



$$[\quad + \quad + \dots + \quad] / n = \quad$$
$$q(i, j) = \frac{1}{n} \sum_{l=1}^n p_l(i, j)$$

Stosowane do redukcji szumu (obrazy astronomiczne)

$$\text{[Noisy Image] + [Noisy Image] = [Smoothed Image]}$$



Uśrednianie przestrzenne (obrazy statyczne oraz zmienne w czasie)

15	13	15
14	0	15
12	12	14

⇒

15	13	15
14	12	15
12	12	14

$$P_{\text{sr}} = \frac{\sum_i^n P_i}{n}$$

$n = 9$ - otoczenie 8-spójne

	13	
14	0	15
	12	

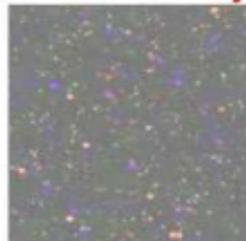
n - liczba pikseli otoczenia
(wraz z pikselem
przetwarzanym)

$n = 5$ - otoczenie 4-spójne

Zniekształcenia radiometryczne

Zniekształcenia radiometryczne

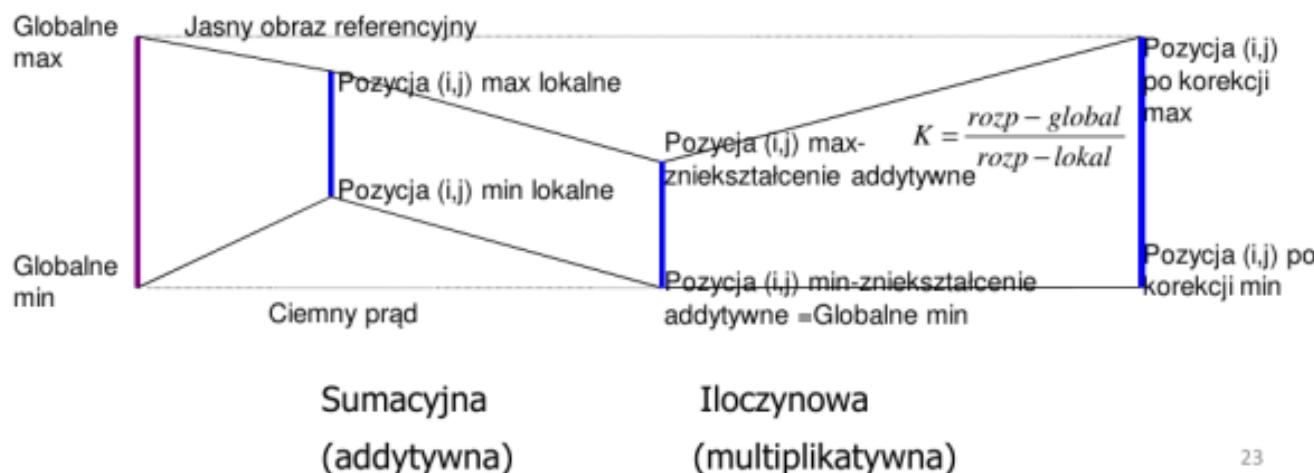
(ang. fixed pattern noise)



- mają charakter **stacjonarny**,
- są **nieprawidłowym** odwzorowaniem jasności światła, padającego na element światłoczuły, na odpowiadającą mu wartość zapisaną w tablicy obrazu (powstaje podczas zamiany kwantu światła na prąd albo podczas kodowania),
- nasilają się przy bardzo **długim czasie ekspozycji**
- nasilają się w miarę **podnoszenia temperatury otoczenia**.

Korekcja radiometryczna

Korekcja zapewnia jednakowe odwzorowanie jasności bez względu na to w jakim miejscu pola widzenia jasność jest rejestrowana.



Proces korekcji radiometrycznej

Korekcja sumacyjna jednorodnego jasnego obrazu odniesienia $[P_{od}(x,y)]$:

$[KORA(x,y)]$

0	0	2	0
0	1	1	0
1	1	0	0
1	2	0	1

$[P_{od}(x,y)]$

10	11	12	12
12	14	11	11
13	10	12	10
10	11	11	10

$[P(x,y)]$

8	8	2	1
8	9	2	3
4	3	3	1
3	2	1	1

$$P_{KORA}(x, y) = P_{od}(x, y) - KORA(x, y) \text{ dla } x=1, \dots, N, y=1, \dots, N,$$

$KORA(x,y)$ - wartość (poziom jasności) piksela obrazu przy zasłoniętym obiektywie (*ciemny obraz odniesienia*)

$P_{od}(x,y)$ - wartość piksela jednorodnego jasnego obrazu odniesienia

$P_{KORA}(x,y)$ - wartość piksela jednorodnego jasnego obrazu odniesienia po korekcji sumacyjnej

Korekcja iloczynowa

$$P_{KORM}(x,y) = [P(x,y) - KORA(x,y)] \cdot KORM(x,y)$$

$P(x,y)$ - wartość piksela obrazu wejściowego

$KORM(x,y)$ – wartość współczynnika korekcji dla piksela o współrzędnych (x,y) obliczona według wzoru:

$$KORM(x,y) = P_{KORA\ max} / P_{KORA}(x,y)$$

$P_{KORA\max}$ - maksymalna wartość piksela w obrazie $[P_{KORA}(x,y)]$

$P_{KORM}(x,y)$ - wartość piksela obrazu wynikowego (po korekcji radiometrycznej)

Przykład

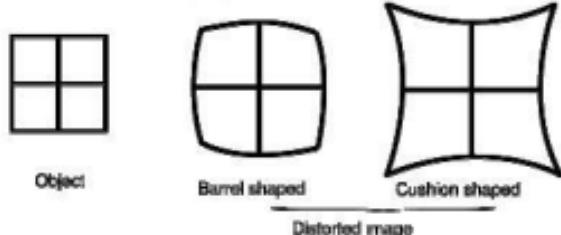


Zniesztalconia geometryczne

Zniekształcenia geometryczne

Zniekształcenia geometryczne (*dystorsje*) są spowodowane:

- nieliniowością układów optycznych wchodzących w skład toru optycznego np.: mikroskopu



- nierównoległością płaszczyzn obrazu i matrycy światłoczułej np.: krzywizna Ziemi w zdjęciach satelitarnych, zmienna wysokość powierzchni ziemi w zdjęciach do sporządzania map, skaningowy mikroskop elektronowy,
- obrotem płaszczyzn obrazu i matrycy światłoczułej np.: skróty perspektywy
- zmianami skali

lub połączeniem powyżej wymienionych.



Konieczność usuwania zniekształceń geometrycznych:

- Pomiary odległości i pól powierzchni,
- Dopasowywanie obiektów na obrazach,
- Dopasowywanie obrazów na potrzeby rekonstrukcji obrazu z fragmentów,
- Tworzenie map, planów na podstawie zdjęć,
- Nakładania obrysów mapy na zdjęcia satelitarne, pogodowe.
- Nakładanie map rozkładu potencjału elektrycznego na powierzchnie 3D.

Realizacja korekcji zniekształceń geometrycznych

- Aproksymacja transformacji wielomianem
- Przekształcenia rozciągające
- Przekształcenia afiniczne

Aproxymacja transformacji wielomianem pierwszego stopnia

$$u = ax + by + c$$

$$v = dx + ey + f$$

x, y - obraz nieznieksztalcony,

u, v - obraz znieksztalcony,

punkty kontrolne (niewspółliniowe) $(x_1, y_1), (x_2, y_2), (x_3, y_3),$

$(u_1, v_1), (u_2, v_2), (u_3, v_3)$,

Aproksymacja transformacji wielomianem pierwszego stopnia

transformacja globalna wyrażona równaniami

$$u = ax + by + c$$

$$v = dx + ey + f$$

punkty kontrolne (niewspółliniowe):

x, y - obraz niezniekształcony; u, v - obraz zniekształcony,

Korekcja obrotu,
przesunięć i skrótu
perspektywy



Formularz

PODAJ WSPÓŁRZĘDNE PUNKTÓW KONTROLNYCH:

Obraz niezniekształcony		Obraz zniekształcony	
x_1	y_1	u_1	v_1
1	1	11	10
50	90	101	30
90	50	106	18

OK



Przykład:

Dane:

$$(x_1, y_1) = (8, 10)$$

$$(x_2, y_2) = (6, 4)$$

$$(x_3, y_3) = (10, 2)$$

$$(u_1, v_1) = (14, 6)$$

$$(u_2, v_2) = (11, 7)$$

$$(u_3, v_3) = (10, 5)$$

Znaleźć:

obraz skorygowany.

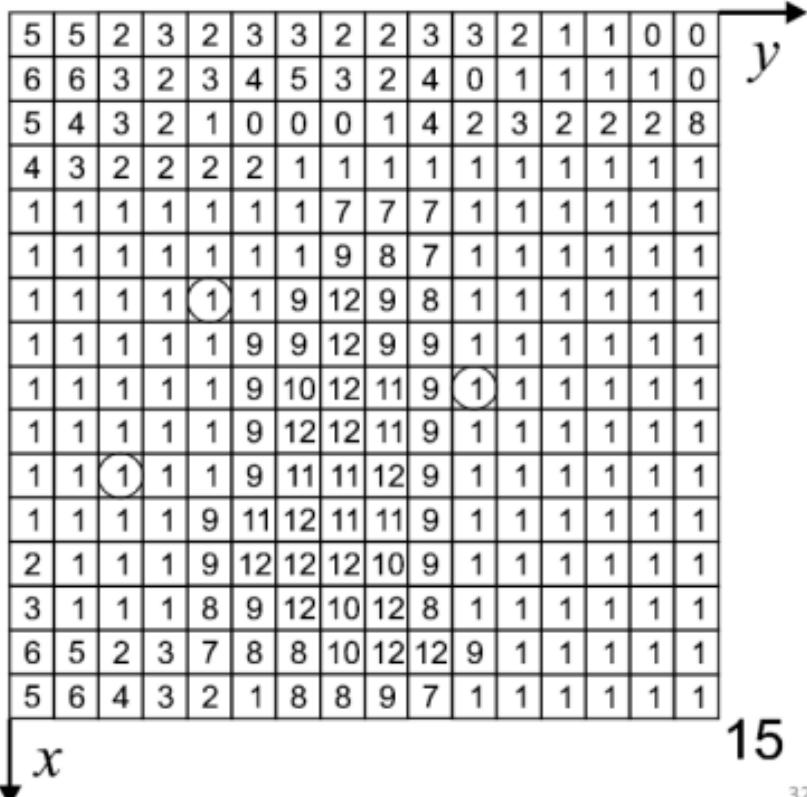
(znalezienie a, b, c, d, e, f,

rozwiązuając

6 równań liniowych

1-go stopnia).

Obraz znieksztalcony



15

Siatka afiniczna

Transformacja lokalna

Przekształcenia afiniczne płaszczyzny i przestrzeni w siebie obejmują m.in. **izometrie** (np. przesunięcie równoległe, obrót, symetrię osiową, symetrię płaszczyznową, symetrię z obrotem, symetrię z poślizgiem), **jednokładności** i **powinowactwa osiowe**.

Niezmiennikami afinycznymi są: równoległość prostych i skośne położenie prostych



Zniekształcenia chromatyczne

Zniekształcenia chromatyczne

- Powstają wtedy, gdy tworzymy obraz wykorzystując więcej niż jeden zakres fali elektromagnetycznej (obrazy wielokanałowe np.: trójkanałowe, multispektralne)
- oraz wtedy, gdy wykorzystujemy różne cechy pomiarowe do tworzenia pojedynczego obrazu (obrazy multimodalne)

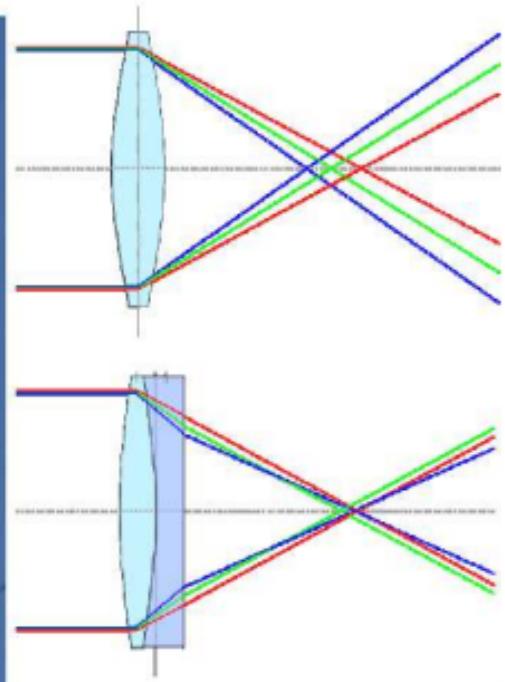
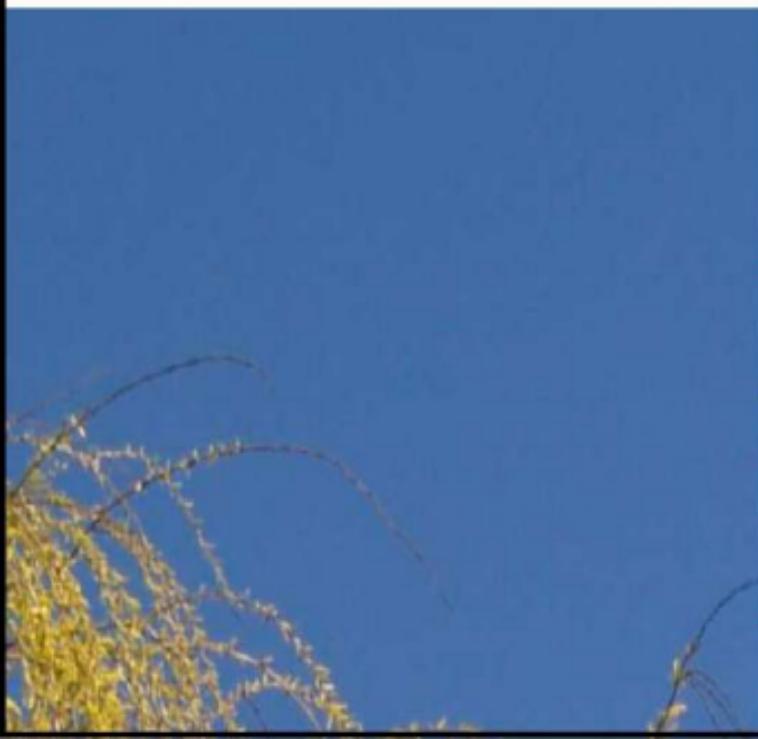
Fale elektromagnetyczne o różnej długości inaczej załamują się na soczewkach

Wczesne (1909) przykłady fotografii kolorowej składanej na szkle obfitują w efekty rozdzielenia barwy na kanały

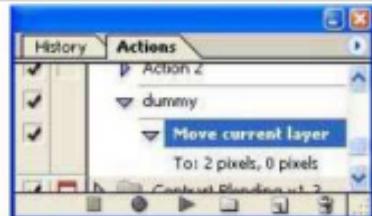
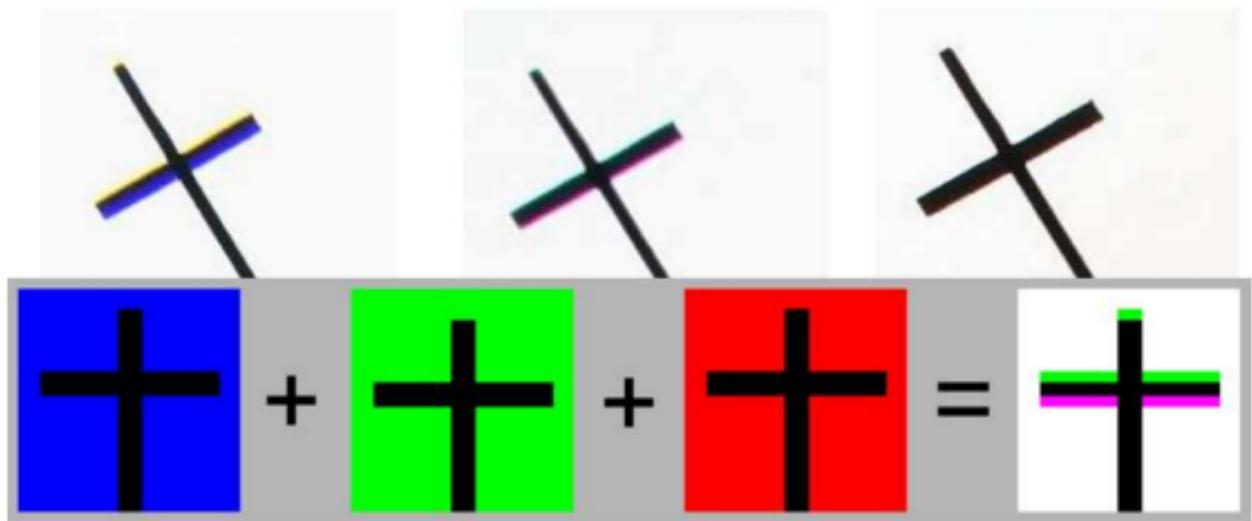


Zdjęcia w kolorze z 1909
Siergiej Prokudin-Gorski,
pionier barwnej fotografii

Korekcja optyczna zniekształceń chromatycznych



Korekcja programowa zniekształceń chromatycznych



Narzędzia do przetwarzania obrazów

Operacje na obrazach

Metody w przetwarzaniu obrazów ze względu na sposób liczenia:

- Globalne (operacje na wszystkich pikselach obrazu)
- Lokalne (operacje na ROI; ang. region of interest)
 - Punktowe;
 - Małym otoczeniu

Ze względu na **typy matematycznych (w tym arytmetycznych, logicznych , statystycznych) operacji**, które są wykonywane na wartościach intensywności.

Funkcje operacji na obrazach będą omawiane przy każdej operacji oddzielnie.

Operacje na obrazach

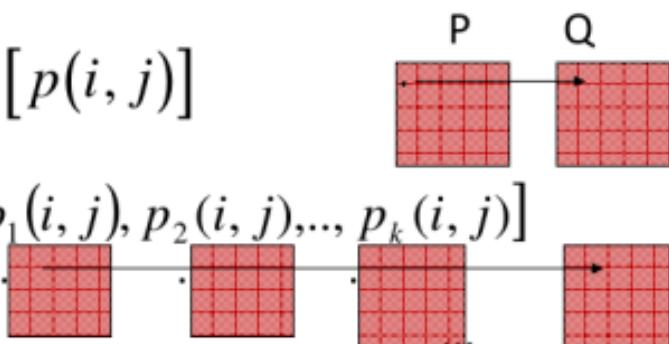
- Operacje punktowe (jednopunktowe):

Jednoargumentowe

$$[q(i, j)] = f[p(i, j)]$$

Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$

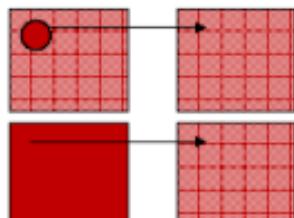


- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

- Operacje globalne transformaty

$$[q(i, j)] = f[P]$$



Operacje punktowe (lokalne, jednopunktowe)

Podział:

Jednoargumentowe

Wieloargumentowe

Operacje punktowe

Proste:

- operacje jednoargumentowe – transformacje jasności
- operacje arytmetyczne: jedno-, dwu- i wieloargumentowe
- manipulowanie histogramem

Zaawansowane:

- metody progowania
- wyrównywanie histogramu
- klasyfikacja punktów obrazu

Operacje jednopunktowe jednoargumentowe

Są to operacje, w których na wartość zadanego piksela obrazu wynikowego o współrzędnych (i, j) ma wpływ wartość **tylko jednego piksela** obrazu pierwotnego o współrzędnych (i, j) :

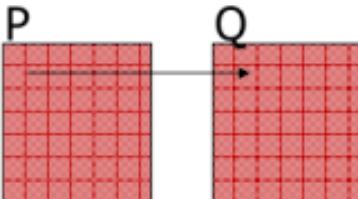
$$[q(i, j)] = f[p(i, j)] \quad f - \text{operator (liniowy lub nielinowy)}$$

$[p(i, j)]$ - obraz pierwotny

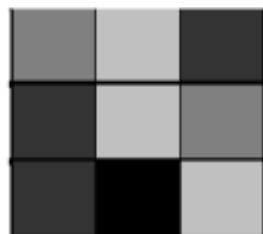
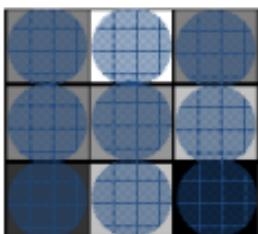
$[q(i, j)]$ - obraz wynikowy

$p(i, j), q(i, j)$ - wartości piksela o współrzędnych (i, j)
obrazu pierwotnego i wynikowego

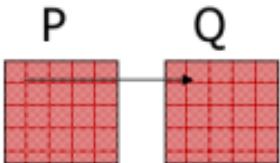
Operacje punktowe jednoargumentowe



Wartość piksla obrazu wyjściowego zależy od wartości piksla o takich samych współrzędnych na obrazie wejściowym (nie uwzględniamy sąsiedztwa)



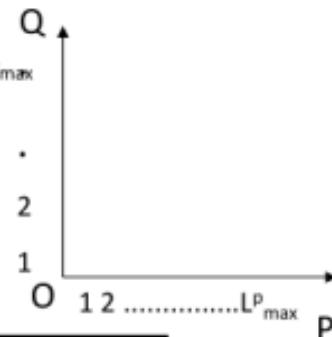
Wykonywanie po wierszach, „piksel po pikselu”



Operacje punktowe

Definiowane przez:

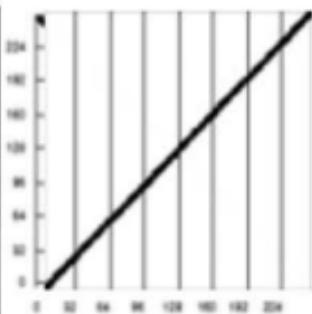
- Definicję funkcji; z jawnie postawionymi warunkami logicznymi np.: dla UOP (**Uniwersalnego Operatora Punktowego**);
- Wykres funkcji we współrzędnych OXY; na osi OX są **wszystkie potencjalne** wartości poziomów szarości obrazu pierwotnego P, a na OY obrazu po przekształceniu Q;
- Tablica przekodowań (LUT – *Look Up Table*).



0	1	2	L_P^{\max}	P
					Q

Tożsamość

Identyczność;



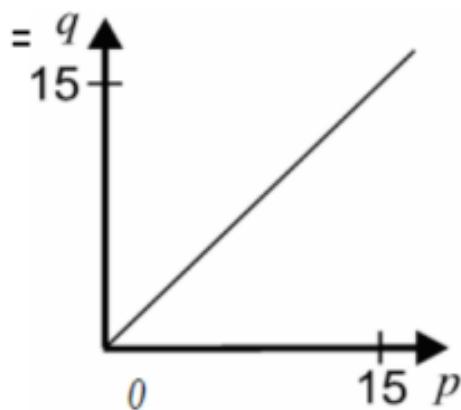
Kopiowanie obrazu

Rodzaje operacji jednopunktowych jednoargumentowych

Operacja identyczności

$$q(i,j) = p(i,j) \text{ dla } L_{\min} \leq p \leq L_{\max}$$

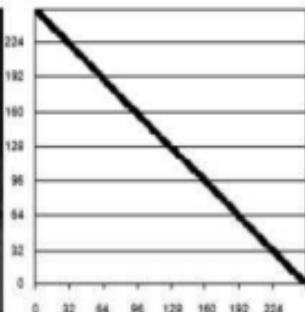
Dla $L_{\min} = 0$, $L_{\max} = 15$ (czyli $M = 16$): [q]



15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

Negacja

Negatyw obrazu



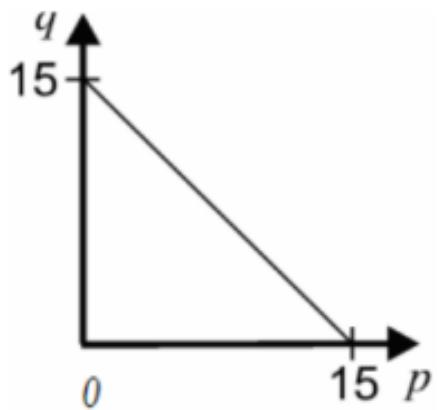
Do prezentacji informacji zawartej w ciemnych tonach (cieniach) jeśli jasne tony są nieistotne

15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

Operacja odwrotności (negacji)

$$q(i,j) = L_{max} - p(i,j) \text{ dla } L_{min} \leq p \leq L_{max}$$

Dla $L_{min} = 0$, $L_{max} = 15$ (czyli $M=16$): $q(i,j)$
 $= 15 - p(i,j)$



[q]

0	0	15	15	13
2	2	0	15	15
15	15	8	1	1
15	14	13	12	11
0	1	2	3	4

Stosowanie negatywów

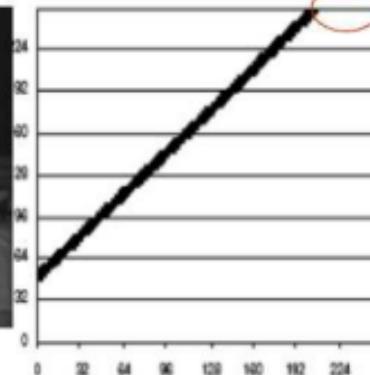
Zdjęcia Rtg



Dodawanie liczby do obrazu

Rozjaśnianie obrazu

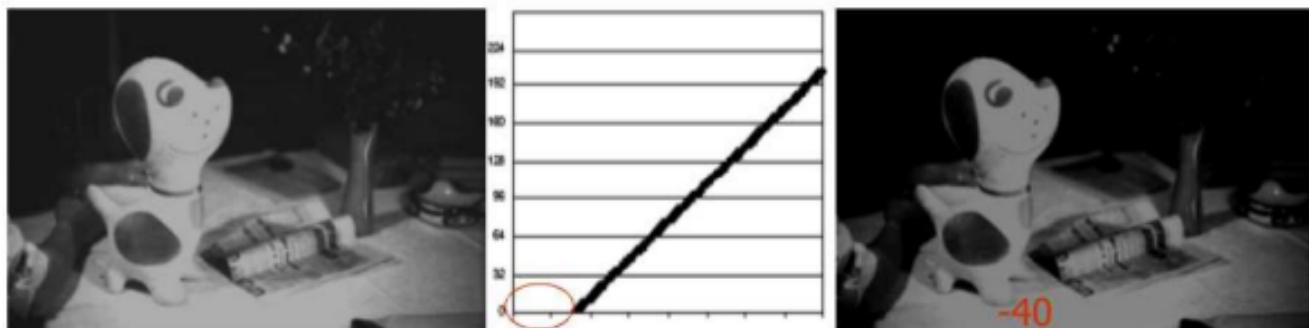
Wykroczenie poza zakres dopuszczalnych wartości



Do prezentacji informacji zawartej w ciemnych tonach (cieniach), jeśli jasne tony są nieistotne

Odejmowanie liczby od obrazu

Przyciemnianie obrazu



Wykroczenie poza zakres dopuszczalnych wartości

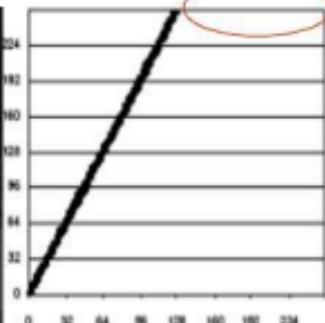
-40

Do prezentacji informacji zawartej w jasnych tonach, jeśli ciemne tony są nieistotne

Mnożenie obrazu przez liczbę

Rozjaśnienie obrazu

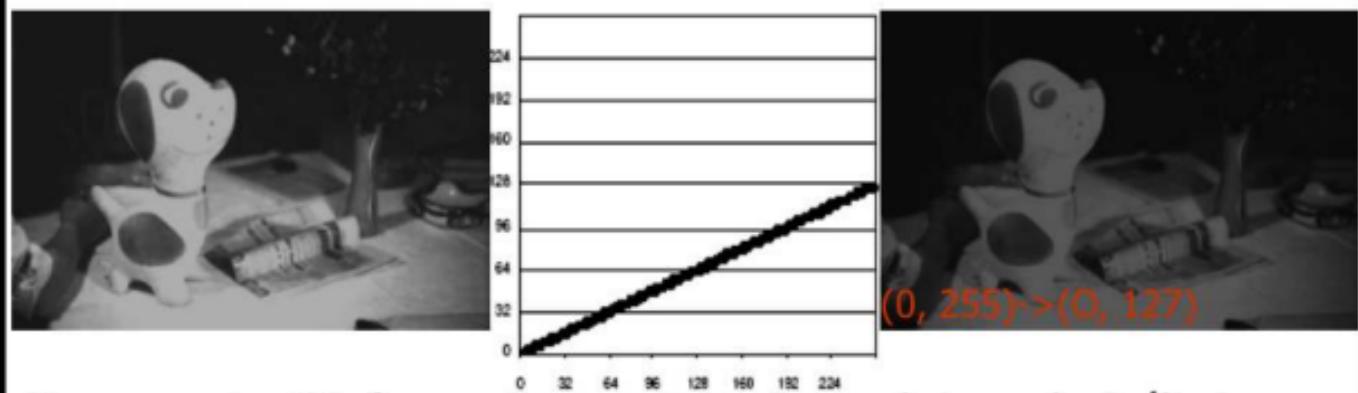
Wykroczenie poza zakres dopuszczalnych wartości



Do prezentacji informacji zawartej w ciemnych i średnich tonach (cieniach), jeśli jasne tony są nieistotne

Dzielenie obrazu przez liczbę

Przyjemnienie obrazu i skalowanie obrazów o podwyższonym zakresie poziomów szarości do obrazów o mniejszym zakresie tonalnym



Do prezentacji informacji zawartej w jasnych tonach, jeśli ciemne tony są nieistotne, przy zauważaniu zakresu poziomów szarości

Manipulowanie histogramem

Zapewnianie odpowiedniego zakresu poziomów szarości przez:

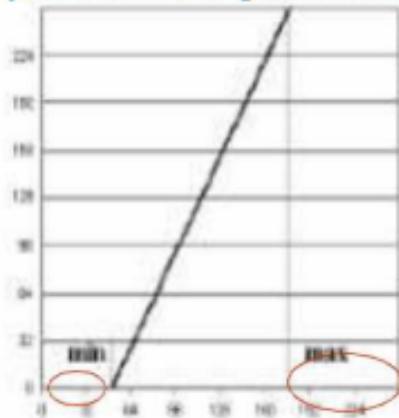
- Rozciąganie (na pełny zakres poziomów jasności – normalizacja)
- Zawężanie
- Modulowanie

Zapewnia dostosowanie do formy prezentacji, do przekazu informacji lub emocji, do planowanych kolejnych operacji na obrazie,

Liniowe rozciąganie histogramu

Optymalne wykorzystanie zakresu poziomów jasności:

- Zagospodarowanie całego zakresu dostępnych poziomów szarości
- Wykorzystanie skrajnych zakresów do prezentacji zakresów średnich (mało liczne, nieistotne)

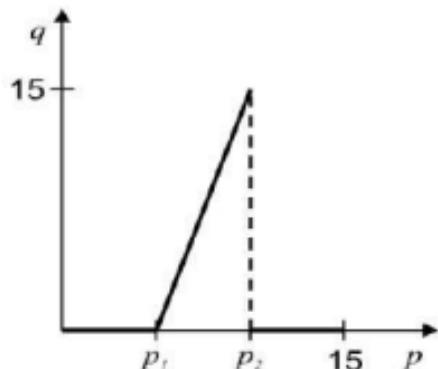


$$q(i, j) = \begin{cases} L_{\min} & \text{dla } p(i, j) < \min \\ \frac{(p(i, j) - \min) * L_{\max}}{\max - \min} & \text{dla } \min \leq p(i, j) \leq \max \\ L_{\max} & \text{dla } p(i, j) > \max \end{cases}$$

Operacja rozciągania

Ogólnie:

$$q = \begin{cases} (p - p_1) \frac{L_{\max}}{(p_2 - p_1)} & \text{dla } p_1 < p \leq p_2 \\ 0 & \text{dla } p \leq p_1, p > p_2 \end{cases}$$



Dla $L_{\min} = 0$, $L_{\max} = 15$ (czyli $M=16$)

$$q = \begin{cases} (p - p_1) \frac{15}{(p_2 - p_1)} & \text{dla } p_1 < p \leq p_2 \\ 0 & \text{dla } p \leq p_1, p > p_2 \end{cases}$$

Wyznaczyć obraz [q] i jego histogram dla $p_1=4$, $p_2=8$

Contrast Enhancement

X

U -



O, 255)



Eyedropper Sampling

- Set input values
 Set output values

Channel

Gray Channel

 Auto-adjust

Options...

Histogram Display Clipping

5 %

 Automatically

17

17

Input Value Clipping



0

25

Output Range Compression

5 %

1,00

Histogram Display Clipping

Gamma Adjustment



Preview

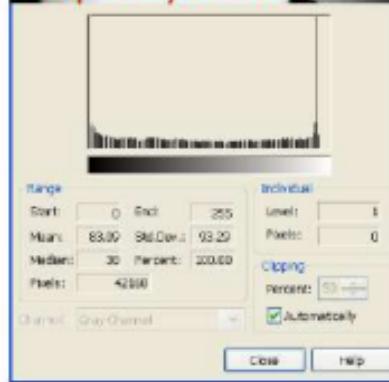
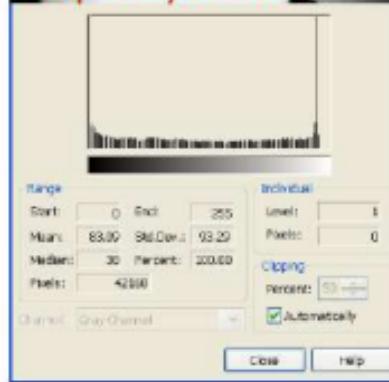
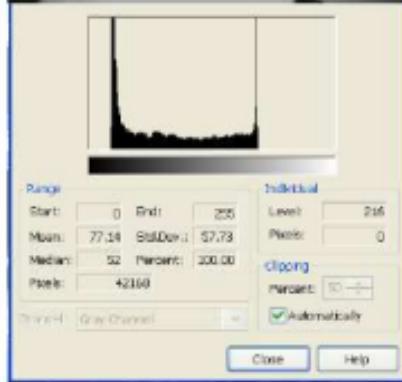
Reset

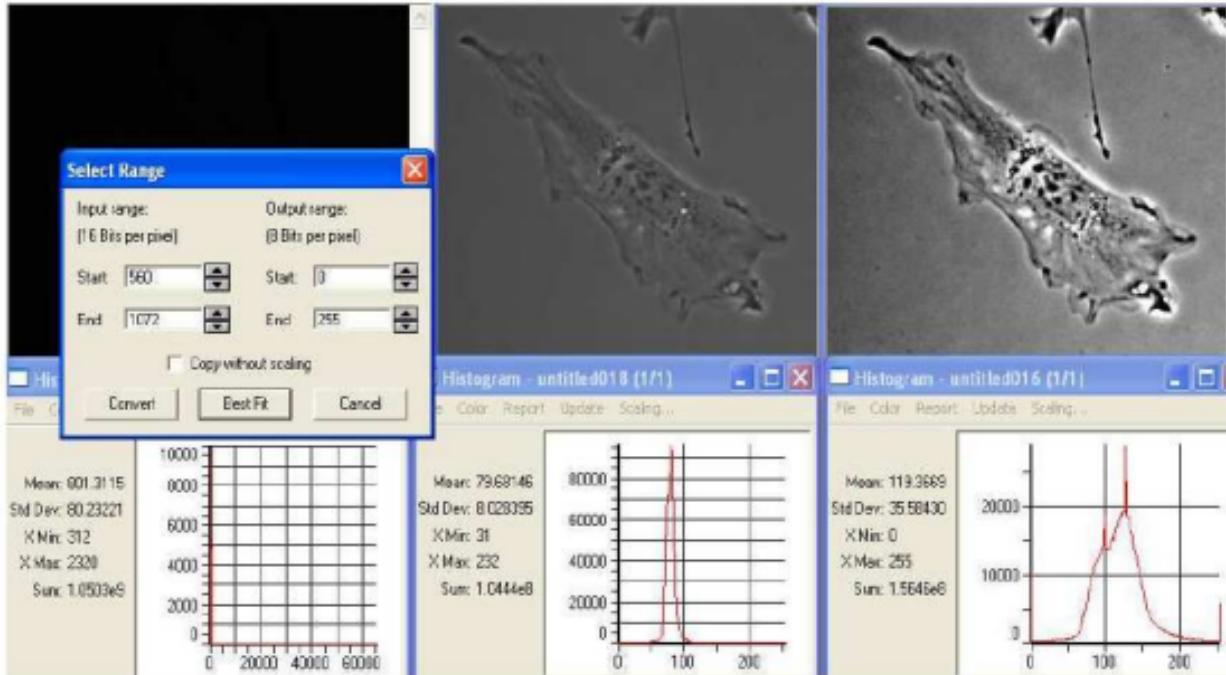
OK

Cancel

Help

Liniowe rozciąganie histogramu





Nieliniowe rozciąganie histogramów

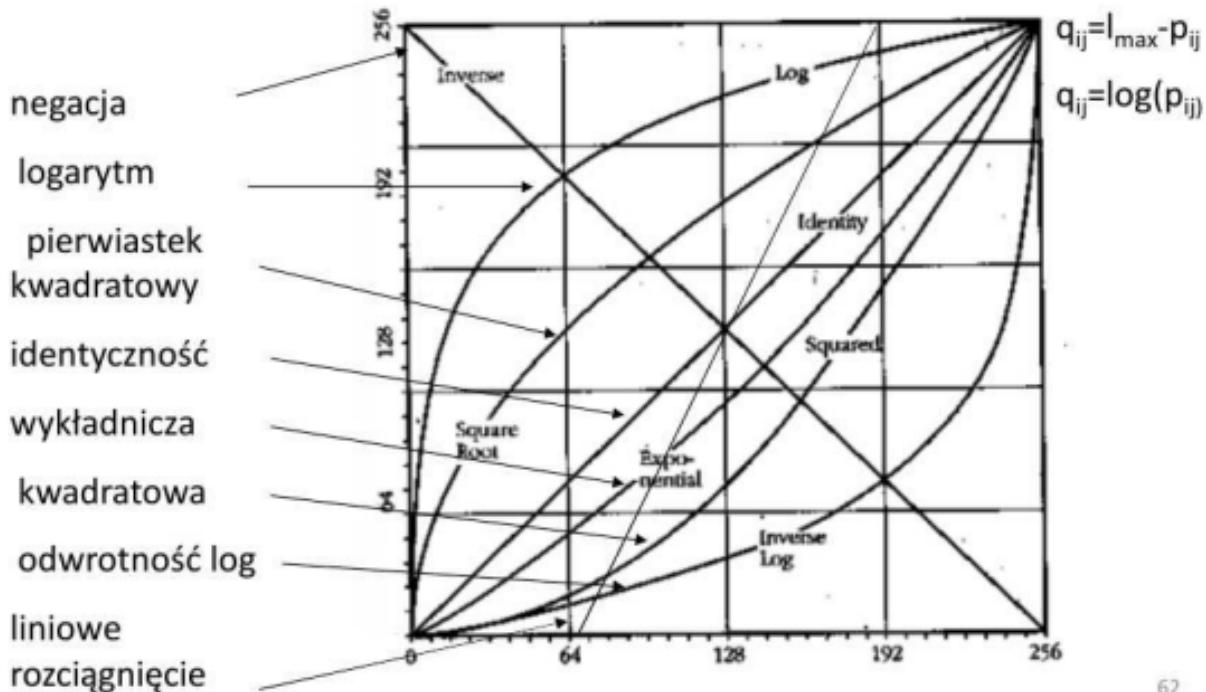
Służy do uwypuklenia pewnych zakresów poziomów szarości kosztem innych, według typowych funkcji matematycznych:

gamma, wykładniczej, pierwiastkowej,
kwadratowej, logarytmicznej czy odwrotności
funkcji logarytmicznej

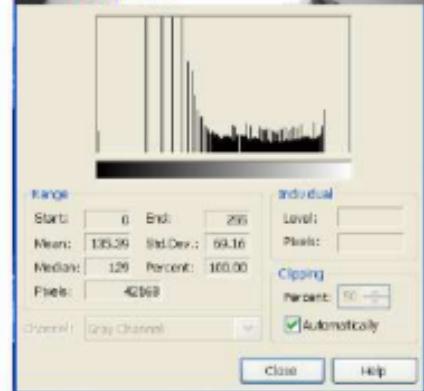
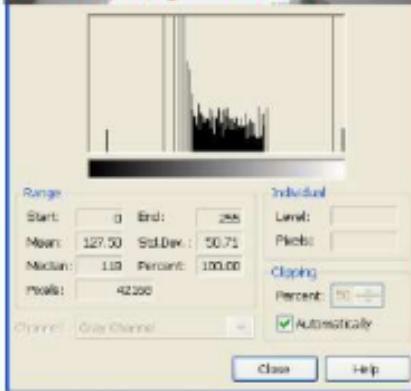
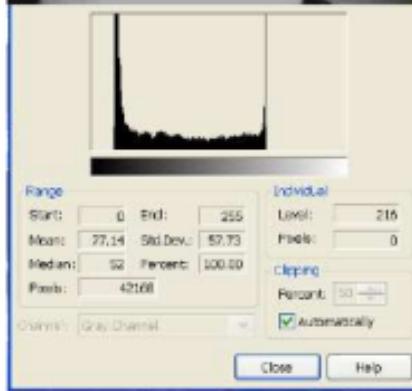
lub

według dowolnie zdefiniowanej funkcji UOP
(za pomocą: krzywej tonalnej lub tablicy LUT)

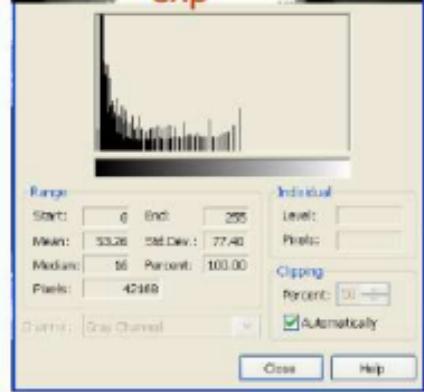
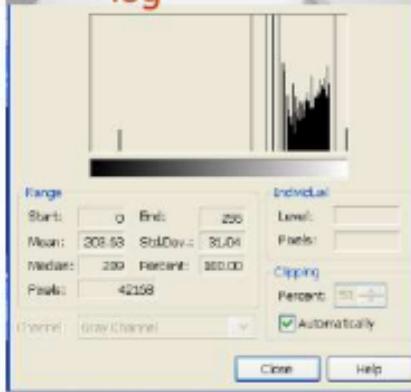
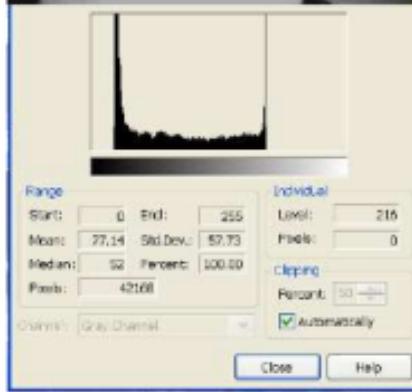
Transformacje jasności według znamy \acute{c} ch matematycznych funkcji



Nieliniowe rozciąganie histogramów

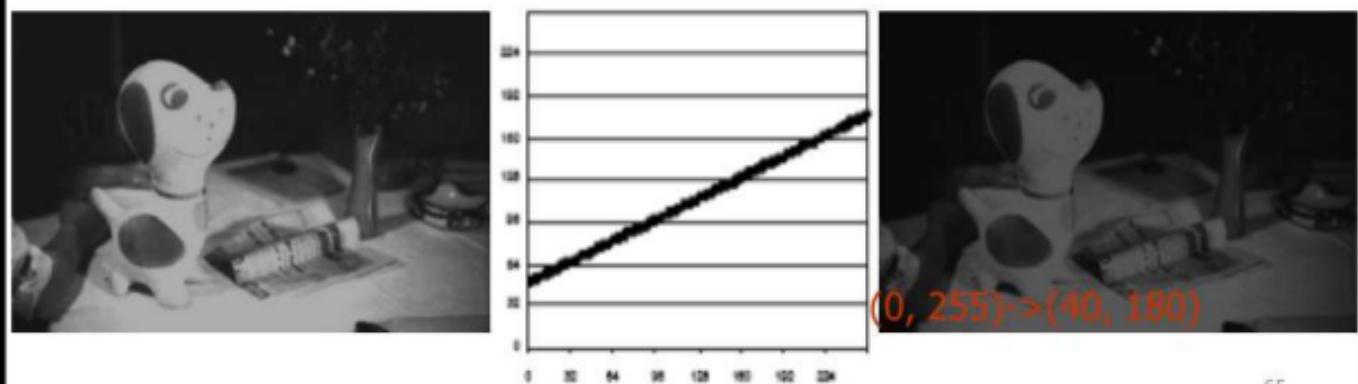


Nieliniowe rozciąganie histogramów

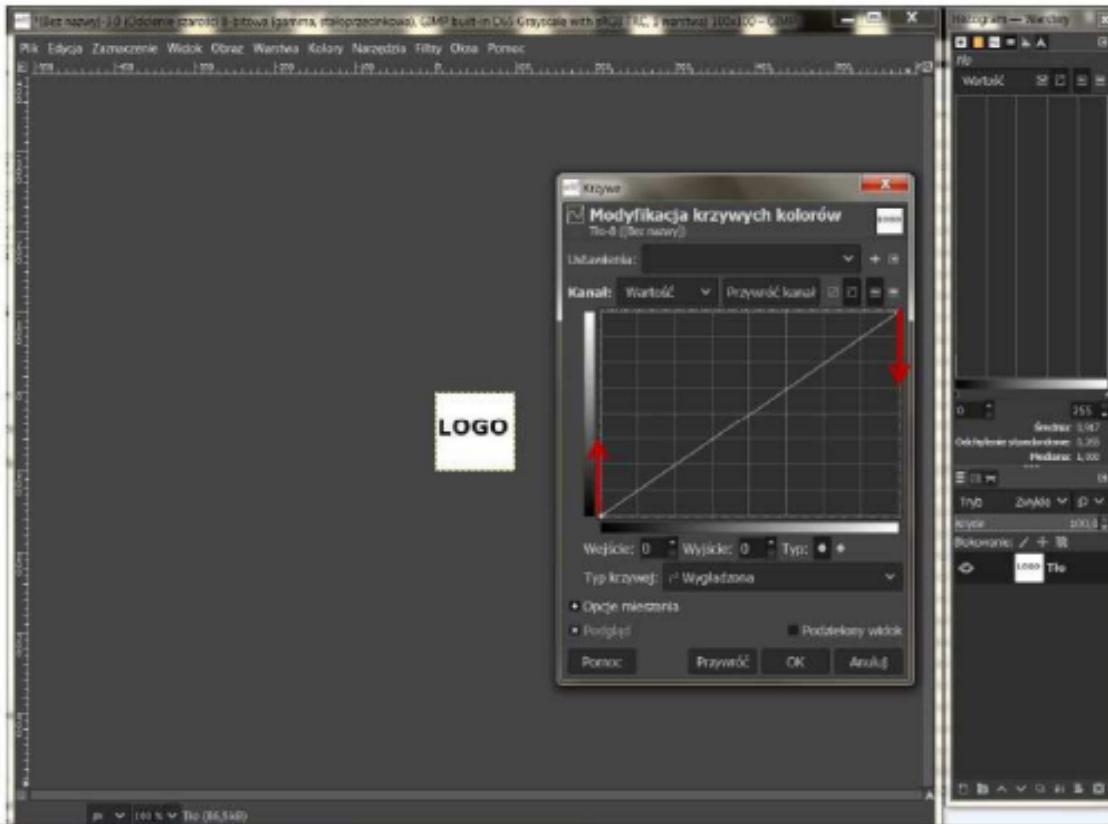


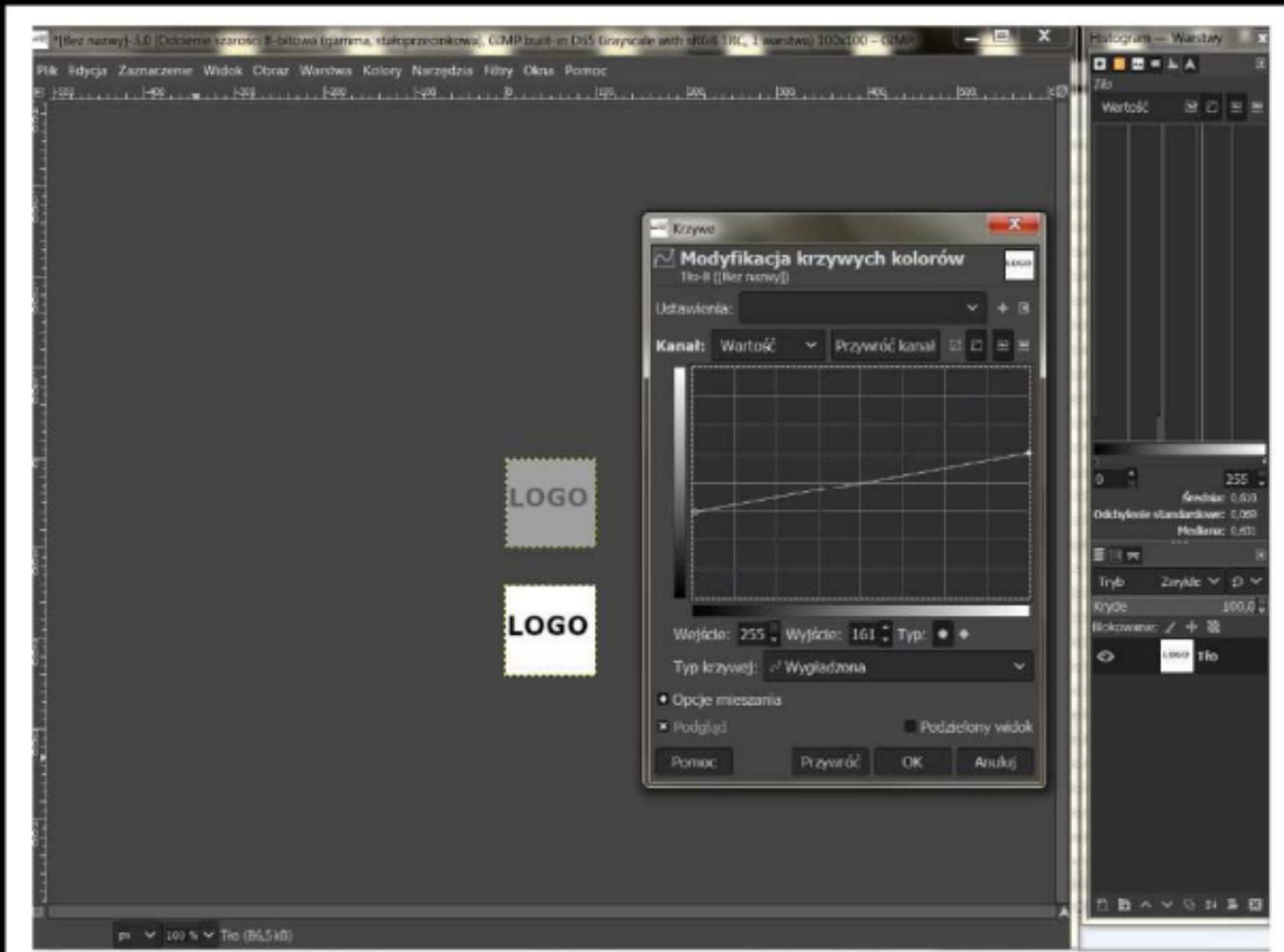
Zwężenie zakresu poziomów szarości

Przez zmniejszenie ilości poziomów faktycznie występujących w obrazie w stosunku do potencjalnie występujących

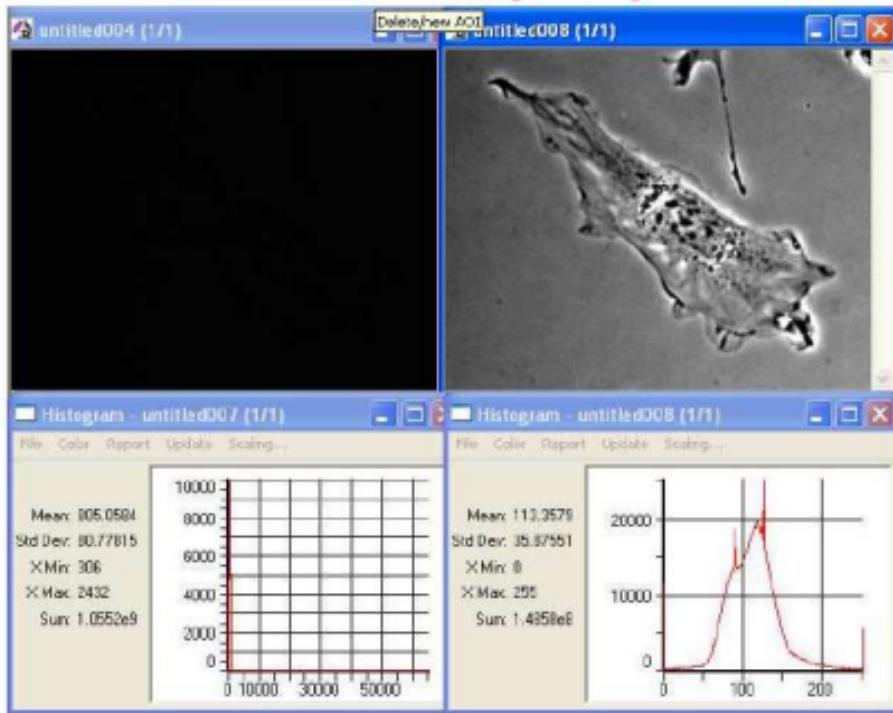


Manipulacja przez zwężanie zakresu histogramu





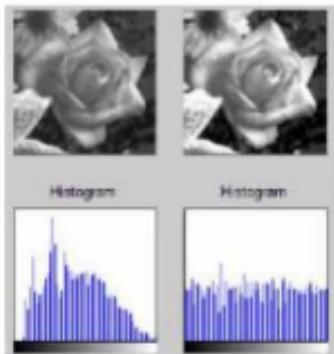
Zwężenie zakresu poziomów szarości - przykład



Wyrównywanie histogramu

Wyrównywanie histogramu:

takie przekształcenia jasności, aby wszystkie jasności była równomiernie reprezentowane w obrazie (narzędzie matematyczne to dystrybuanta czyli całka z histogramu)



Wyrównanie typu equalizacja (ang. Equalisation):

Rozszerzenia odległości między słupkami odpowiadającymi odcieniom szarości silniej reprezentowanymi (czyli o wysokich słupkach) a zawężenie pomiędzy słupkami o małej wysokości.

Redukcja liczby poziomów szarości- przykłady



Różne liczby przedziałów kwantowania



Redukcja liczby poziomów szarości - posteryzacja

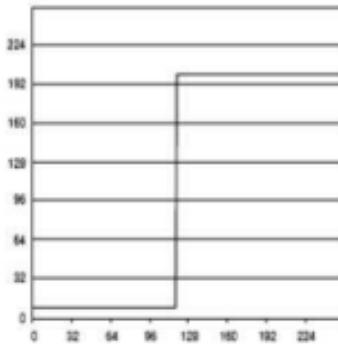
Powtórna kwantyzacja na mniejszą zadaną liczbę poziomów szarości (2, 3, 4, 5, ..., 255) – postryzacja (od ang. Posterize)

Cel:

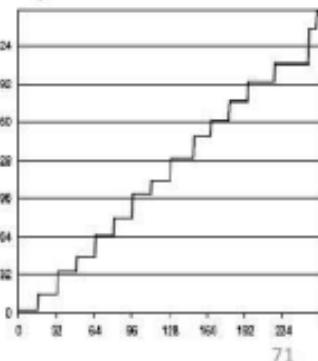
1. artystyczny – obraz posteryzowany o ostrych granicach i niewielkiej ilości tonów ,
2. prosta segmentacja,
3. kompresja

Inż. W. Romer „Izohelja”
w Kamera Polska

Dwa przedziały – binaryzacja

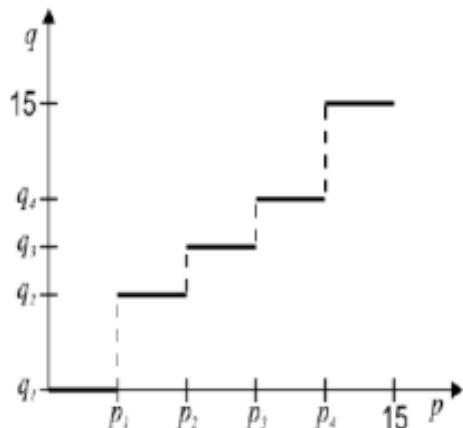


Szesnaście przedziałów



Operacja nierównomiernej redukcji poziomów szarości

Przykład dla $L_{min} = 0$, $L_{max} = 15$ (czyli $M=16$)



$$q = \begin{cases} 0 & \text{dla } p \leq p_1 \\ q_2 & \text{dla } p_1 < p \leq p_2 \\ q_3 & \text{dla } p_2 < p \leq p_3 \\ q_4 & \text{dla } p_3 < p \leq p_4 \\ 15 & \text{dla } p_4 < p \leq 15 \end{cases}$$

Wyznaczyć obraz $[q]$ i jego histogram dla:

$p1 = 2, p2 = 4, p3 = 6, p4 = 8; q2 = 3, q3 = 6, q4 = 9;$

Posteryzacja: $p2-p1=p3-p2=\dots; q2-q1=q3-q2=\dots;$

Progowanie

Jest to taka wersja operacji zmniejszenia ilości poziomów szarości do dwóch, dla której istnieje możliwość arbitralnego wyboru wartości progu (p_1) czyli szarości granicznej, od której przyporządkowujemy wyższy poziom szarości (najczęściej biel) i poniżej której przyporządkowujemy niższy próg szarości (najczęściej czerń).

$$q = \begin{cases} L_{\min} & \text{dla } p \leq p_1 \\ L_{\max} & \text{dla } p > p_1 \end{cases}$$

Różne typy progowania

- Progowanie z pojedynczym progiem segmentacji

$$q = \begin{cases} L_{\min} & \text{dla } p \leq p_1 \\ L_{\max} & \text{dla } p > p_1 \end{cases}$$

- Progowanie przedziałami

$$q = \begin{cases} L_{\max} & \text{dla } p_1 \leq p \leq p_2 \\ L_{\min} & \text{dla } p < p_1 \text{ lub } p > p_2 \end{cases}$$

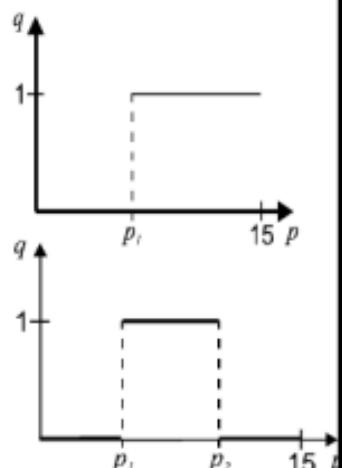
- Progowanie z warunkiem spójności

lub warunkiem nałożonym na wielkość obiektu

- Progowanie adaptacyjne

- Progowanie rekurencyjne

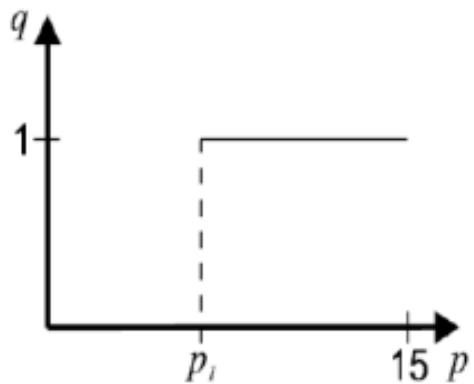
- Progowanie hierarchiczne (piramidowe, skalowalne)



Operacja progowania (binaryzacji)

$$q = \begin{cases} 0 & \text{dla } p \leq p_1 \\ 1 & \text{dla } p > p_1 \end{cases}$$

$$p_1 = 5$$



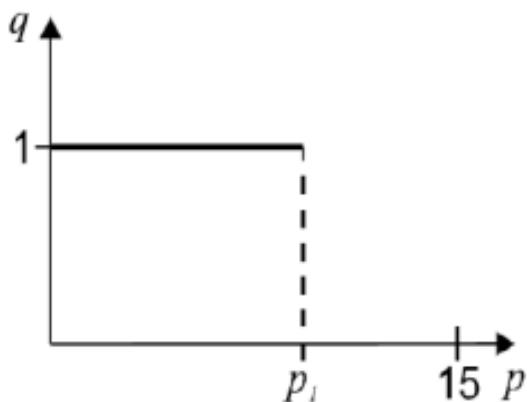
15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

1	1	0	0	0
1	1	1	0	0
0	0	1	1	1
0	0	0	0	0
1	1	1	1	1

Operacja odwrotnego progowania (binaryzacji)

$$q = \begin{cases} 1 & \text{dla } p \leq p_1 \\ 0 & \text{dla } p > p_1 \end{cases}$$

$$p_1 = 5$$

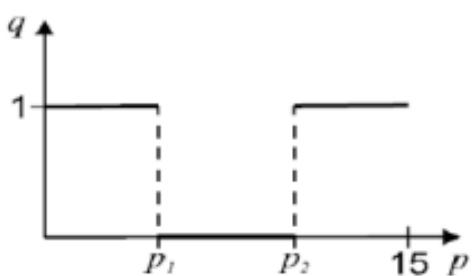


15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

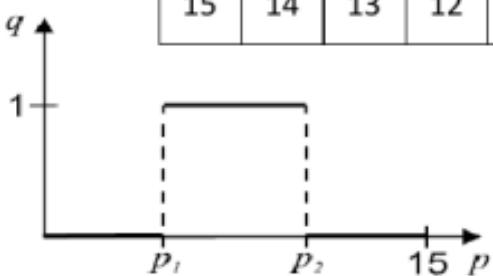
0	0	1	1	1
0	0	0	1	1
1	1	0	0	0
1	1	1	1	1
0	0	0	0	0

Operacje progowania przedziałami (binarne)

$$p_1 = 2, p_2 = 12$$



1	1	1	1	0
1	1	1	1	1
1	1	0	1	1
1	1	0	0	0
1	1	1	0	0



0	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	1	1	1
0	0	0	1	1

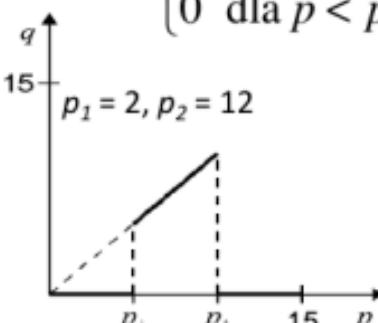
Operacje progowania z zachowaniem poziomów szarości

Dla $L_{min} = 0$, $L_{max} = 15$ (czyli $M=16$)

Z zachowaniem odwrotności
(negacji)

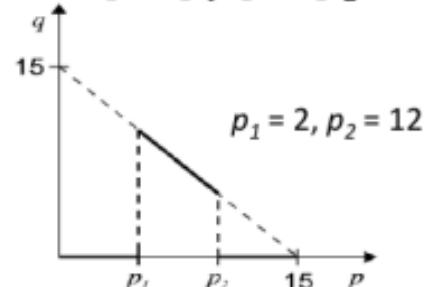
Z zachowaniem identyczności

$$q = \begin{cases} p & \text{dla } p_1 \leq p \leq p_2 \\ 0 & \text{dla } p < p_1, p > p_2 \end{cases}$$



0	0	0	0	2
0	0	0	0	0
0	0	7	0	0
0	0	2	3	4
0	0	0	12	11

$$q = \begin{cases} L_{max} - p & \text{dla } p_1 \leq p \leq p_2 \\ 0 & \text{dla } p < p_1, p > p_2 \end{cases}$$



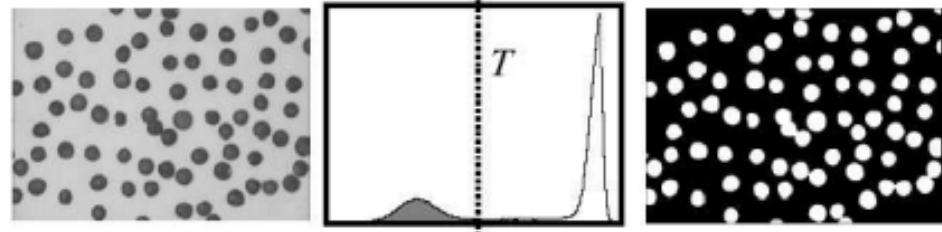
0	0	0	0	13
0	0	0	0	0
0	0	8	0	0
0	0	13	12	11
0	0	0	3	4

Przykład progowania

Progowanie dobrze segmentuje tylko wtedy, gdy:

- istnieje rozdzielność poziomów szarości lub kolorów obiektu i tła

- gdy „dolina” progu jest głęboka (najlepiej zerowa)



Operacje na obrazach (II)

WYKŁAD 3
Dla studiów niestacjonarnych
2020/2021

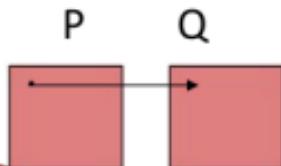
Dr hab. Anna Korzyńska, prof. IBIB PAN

Operacje na obrazach

- Operacje punktowe (jednopunktowe):

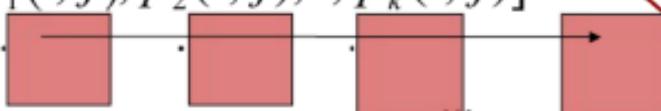
Jednoargumentowe

$$[q(i, j)] = f[p(i, j)]$$



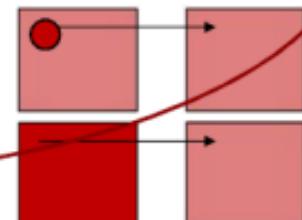
Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$



- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

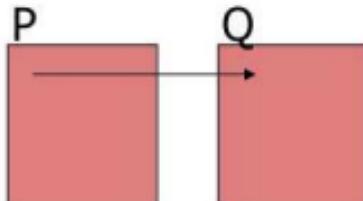


- Operacje globalne transformaty

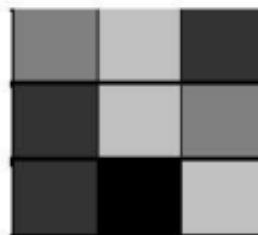
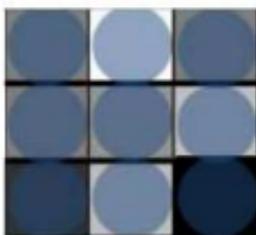
$$[q(i, j)] = f[P]$$

Operacje punktowe (lokalne, jednopunktowe) - powtórzenie

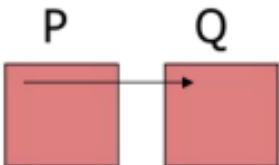
Operacje punktowe jednoargumentowe



Wartość piksla obrazu wyjściowego zależy od wartości piksla o takich samych współrzędnych na obrazie wejściowym (nie uwzględniamy sąsiedztwa)



Wykonywanie po wierszach, „piksel po pikselu”



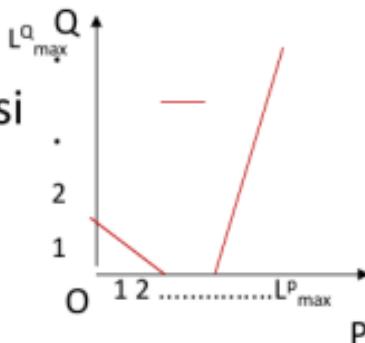
Operacje punktowe

Definiowane przez:

- Definicję funkcji; z jawnie postawionymi warunkami logicznymi np.: dla UOP (**Uniwersalnego Operatora Punktowego**);

- Wykres funkcji we współrzędnych OXY; na osi OX są **wszystkie potencjalne** wartości poziomów szarości obrazu pierwotnego P, a na OY obrazu po przekształceniu Q;

- Tablica przekodowań (LUT – *Look Up Table*).



0	1	2	L_P^{\max}	P
					Q

Operacje punktowe

Proste:

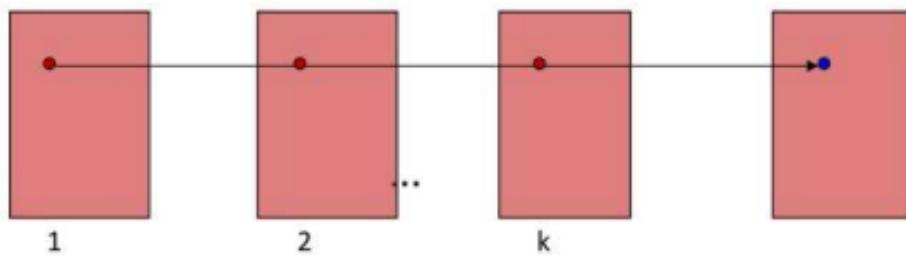
- operacje jednoargumentowe – transformacje jasności nielinowe - wykład 2 slajdy 50-67
- **operacje arytmetyczne: jedno-, dwu- i wieloargumentowe**
- manipulowanie histogramem (liniowe) - wykład 2 slajdy 50-67

Zaawansowane:

- metody progowania - wykład 2 slajdy 67-73
- wyrównywanie histogramu - wykład 2 slajdy 60 – dokładniej na APO
- klasyfikacja punktów obrazu – clustering, pseudokoloryzacja – tylko na projektach z APO

Operacje punktowe wieloargumentowe

$$[q(i,j)] = f[p_1(i,j), p_2(i,j), \dots, p_k(i,j)]$$



Realizacja:

1. Operacje (Calculation, Operation)
2. Na warstwach (add, subtract, difference,...)

Operacje jednopunktowe dwuargumentowe i wieloargumentowe

Są to operacje, w których na wartość zadanego piksela obrazu wynikowego o współrzędnych (i,j) mają wpływ tylko **wartości pikseli obrazów pierwotnych (argumentów) o współrzędnych (i,j) :**

$$c_{ij} = f_D(a_{ij}, b_{ij})$$

f_D - operacja **arytmetyczna** lub **logiczna** (dodawanie, odejmowanie, mnożenie, dzielenie, NOT, OR, AND, XOR...)

c_{ij} - wartość piksela obrazu wynikowego (w przypadku otrzymania ułamka następuje zaokrąglenie do najbliższej liczby całkowitej(lub obcięcie))

a_{ij}, b_{ij} - wartości pikseli obrazów pierwotnych

Operacje punktowe wieloargumentowe

Wykonywane na dwóch lub większej liczbie obrazów

- Dodawanie (uśrednianie)
- Odejmowanie (różnica i różnica bezwzględna)
- Mnożenie
- Dzielenie
- AND
- OR
- XOR
- Uśrednienie obrazów

Dodatkowo jednoargumentowe - pominięte

- NOT
- Pseudokolor

Dodawanie obrazów

Przekroczenie zakresów poziomów szarości regulujemy:

- Wagami (np. równymi - uśrednianie, nierównymi)
- Funkcją modulo,
- Skalowaniem wyniku
(min, max) → (0, L_{max})



+



=



Zastosowanie do: łączenia masek i efekty nałożenia obiektów i przenikania

Operacje jednopunktowe dwuargumentowe - arytmetyczne: uśrednianie (dodawanie ze skalowaniem)

$$c_{ij} = \frac{(a_{ij} + b_{ij})}{k}$$

k - liczba obrazów; przykład: *k* = 2,

$$\begin{bmatrix} 0 & 12 & 142 & 255 \\ 1 & 6 & 40 & 254 \\ 24 & 0 & 20 & 255 \\ 30 & 2 & 10 & 240 \end{bmatrix} \quad \begin{bmatrix} 14 & 11 & 9 & 253 \\ 3 & 5 & 39 & 254 \\ 11 & 1 & 19 & 255 \\ 18 & 2 & 11 & 256 \end{bmatrix} \quad \begin{bmatrix} 7 & 12 & 76 & 254 \\ 2 & 6 & 40 & 254 \\ 18 & 1 & 20 & 255 \\ 23 & 2 & 11 & 248 \end{bmatrix}$$

a_{ij} b_{ij} c_{ij}

Przykładowe zastosowanie: redukcja zakłóceń

Odejmowanie obrazów



- różnica



Przekroczenie zakresów poziomów szarości regulujemy:

- Funkcją modulo,
- Skalowaniem wyniku

- różnica bezwzględna



Obraz po liniowym rozciągnięciu histogramu do podwojenia zakresu

Zastosowanie do:

1. Pokazania różnicy między obrazami, zwłaszcza w przypadku, gdy porównywane obrazy są nieroóżnialne wzrokowo
2. Angiografii różnicowej

Najczęściej wykonywana operacja arytmetyczna na obrazach

Operacje jednopunktowe dwuargumentowe - arytmetyczne: Odejmowanie

$$\begin{bmatrix} 0 & 12 & 142 & 255 \\ 1 & 6 & 40 & 254 \\ 24 & 0 & 20 & 255 \\ 30 & 2 & 10 & 240 \end{bmatrix}_{a_{ij}} \quad \begin{bmatrix} 14 & 11 & 9 & 253 \\ 3 & 5 & 39 & 254 \\ 11 & 1 & 19 & 255 \\ 18 & 2 & 11 & 256 \end{bmatrix}_{b_{ij}}$$

1-szy etap: $a_{ij} - b_{ij}$

$$\begin{bmatrix} -14 & 1 & 133 & 2 \\ -2 & -19 & -60 & 254 \\ 15 & -1 & -60 & 254 \\ 0 & 0 & -100 & -15 \end{bmatrix}$$

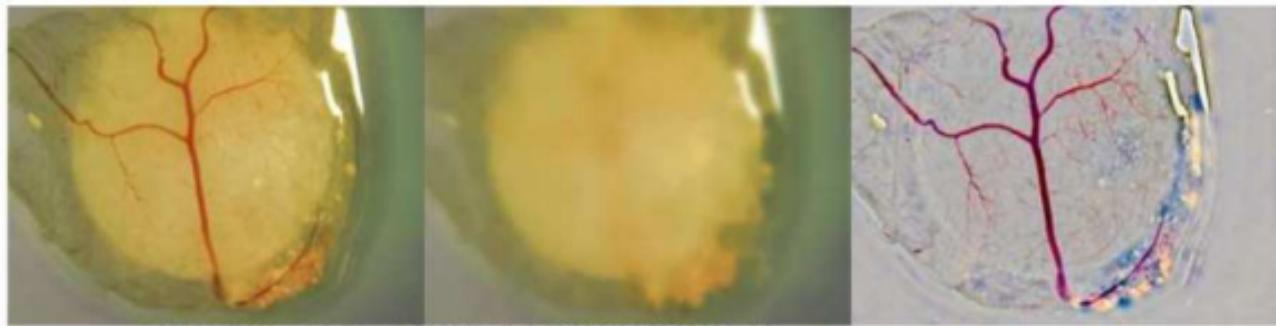
2-gi etap: $c_{ij} = |a_{ij} - b_{ij}|$

$$\begin{bmatrix} 14 & 1 & 133 & 2 \\ 2 & 19 & 60 & 254 \\ 15 & 1 & 60 & 254 \\ 0 & 0 & 100 & 15 \end{bmatrix}$$

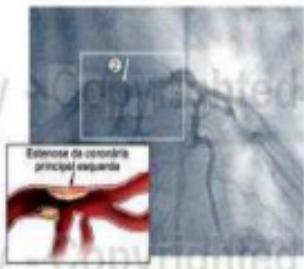
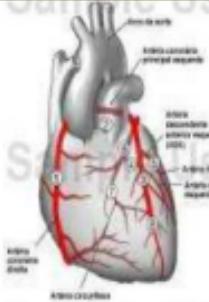
Przykładowe zastosowanie: porównywanie obrazów

Obrazy różnicowe

- Angiografia różnicowa gałki ocznej



Naczynia
serca



Naczynia
mózgu



Mnożenie obrazów

Zastosowanie:

- kompozycje artystyczne
- maskowanie szaroodcieniową lub binarną maską
- jako element operacji rekonstrukcji obrazów



*



=



*



=



Operacje jednopunktowe dwuargumentowe - arytmetyczne: Mnożenie

- tworzenie okna

$$c_{ij} = a_{ij} \times b_{ij} \quad b_{ij} = \begin{cases} 1 & \text{- wewnątrz okna} \\ 0 & \text{- poza oknem} \end{cases}$$

Przykład:

$$\begin{bmatrix} 0 & 12 & 142 & 255 \\ 1 & 6 & 40 & 254 \\ 24 & 0 & 20 & 255 \\ 30 & 2 & 10 & 240 \end{bmatrix}$$

obraz $[a_{ij}]$

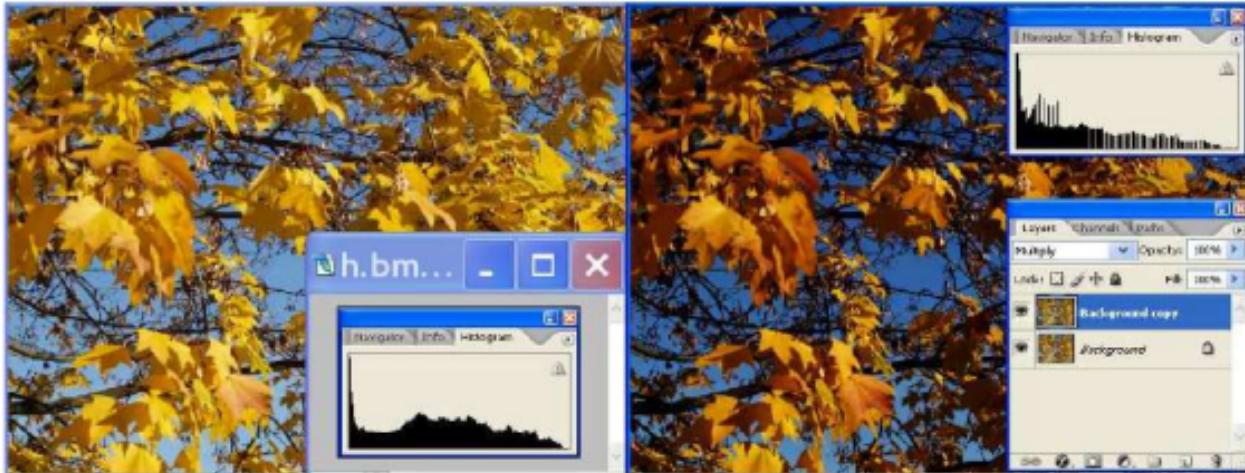
$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

współczynniki korekcji b_{ij}

$$\begin{bmatrix} 0 & 0 & 142 & 0 \\ 0 & 6 & 40 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

obraz $[c_{ij}]$

Przyciemnianie przez mnożenie



Jest to podnoszenia obrazu do kwadratu

Dzielenie obrazów

Przekroczenie zakresów poziomów szarości regulujemy:

- Wagami,
- Funkcją modulo,
- Skalowaniem wyniku



Dzielenie jest wykonywane na drodze mnożenia jednego obrazu przez odwrotność drugiego

Przez odwrotność, a nie negację!



Operacje logiczne

Poziom jasności n jest zapisany w kodzie dwójkowym jako kombinacja ośmiu 0 i 1:

Czerń 00000000

Biel 11111111

127 10000000

Operacje logiczne:

NOT NOT(1)=0; NOT(0)=1

AND 1 AND 1=1; 0 AND 0=0; 1 AND 0=0; 0 AND 1=0

OR 1 OR 1=1; 0 OR 0=0; 1 OR 0=1; 0 OR 1=1

XOR 1 XOR 1=0; 0 XOR 0=0; 1 XOR 0=1; 0 XOR 1=1

Jak wykonujemy operacje logiczne

W operacjach jednopunktowych dwuargumentowych logicznych na obrazach działania prowadzone są na odpowiednich pikselach obrazów stanowiących argumenty danej operacji.

W szczególności działania prowadzone są na bitach o tej samej wadze.

Przykład :

Operacja OR: $[p] \cup [q] = [r]$

Obraz $[p]$:

0	0	0	1	0
1	1	15	14	0
1	0	15	13	2
2	1	0	14	0
1	0	0	0	0

Obraz $[q]$:

14	14	14	14	14
14	14	14	14	14
14	14	14	14	14
14	14	14	14	14
14	14	14	14	14

0000	0000	0000	0001	0000
0001	0001	1111	1110	0000
0001	0000	1111	1101	0010
0010	0001	0000	1110	0000
0001	0000	0000	0000	0000

1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110

Obraz $[r]$:

1110	1110	1110	1110	1110
1111	1111	1111	1110	1110
1111	1110	1111	1111	1110
1110	1111	1110	110	1110
1110	1110	1110	1110	1110

14	14	14	14	14
15	15	15	14	14
15	14	14	15	14
14	15	14	14	14
14	14	14	14	14

Przykład :

Operacja XOR: $[p] > \ll [q] = [r]$

Obraz $[p]$:

0	0	0	1	0
1	1	15	14	0
1	0	15	13	2
2	1	0	14	0
1	0	0	0	0

Obraz $[q]$:

14	14	14	14	14
14	14	14	14	14
14	14	14	14	14
14	14	14	14	14
14	14	14	14	14

0000	0000	0000	0001	0000
0001	0001	1111	1110	0000
0001	0000	1111	1101	0010
0010	0001	0000	1110	0000
0001	0000	0000	0000	0000

1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110
1110	1110	1110	1110	1110

Obraz $[r]$:

1110	1110	1110	1111	1110
1111	1111	0001	0000	1110
1111	1110	0001	0011	0000
1100	1111	1110	0000	1110
1111	1110	1110	1110	1110

14	14	14	15	14
15	15	1	0	14
15	14	1	3	0
12	15	14	0	14
15	14	14	14	14

Maska



Operacja logiczna AND na obrazach

- Maskowanie, czyli selekcja fragmentów obrazów, zwanych ROI (ang. *region of interest*) na podstawie binarnej maski
- Zacieśnianie maski
- Przygotowanie do kodowania informacji (czyszczenie)



AND \cap



=



AND \cap



=



Operacja logiczna OR na obrazach

- Rozszerzanie maski
- Nakładanie informacji szyfrowanej



Operacja logiczna XOR na obrazach

XOR



OR





Inne operacje punktowe

- Operacja logiczna jednoargumentowa na obrazie NOT

$$q(i, j) = \text{NOT } p(i, j)$$

Równoważna negacji

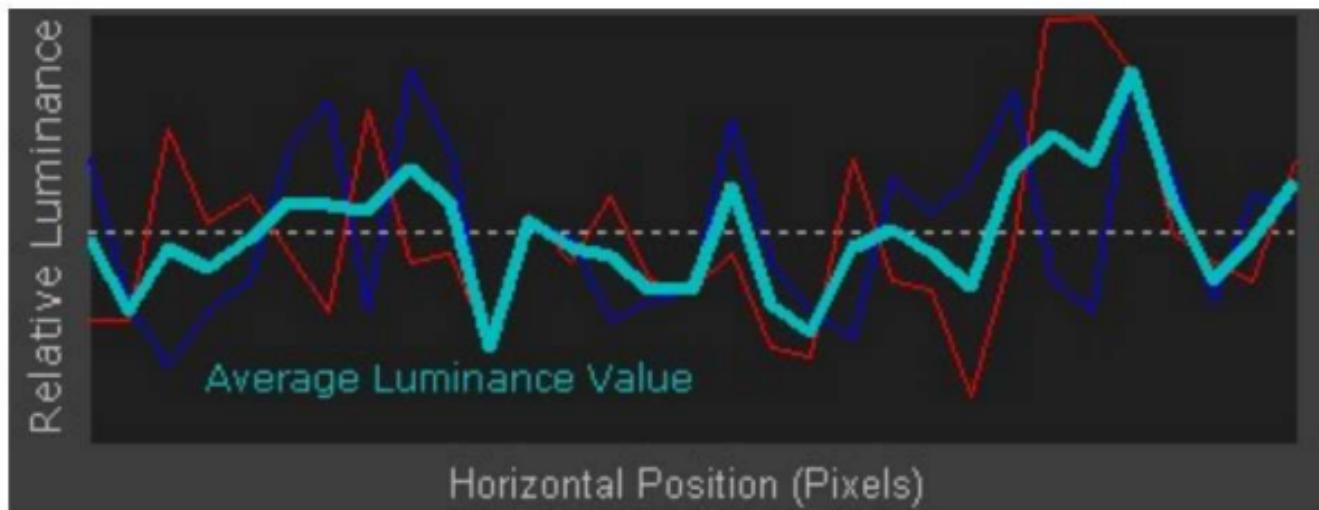
- Uśrednianie k obrazów

Stosowane do redukcji szumu (obrazy astronomiczne)



$$[\quad + \quad + \dots + \quad] / n = \quad$$

$$q(i, j) = \frac{1}{k} \sum_{l=1}^k p_l(i, j)$$



Inne operacje punktowe

- A. Konwersja koloru na obrazy szarooodcieniowe

np.: $g(\text{RGB}) = 60\% \text{ int } R + 30\% G + 10\% B$

- A. Pseudokolor – operacja podnoszenia widoczności niewielkich różnic w intensywności

2^{24}



A ↓ B
↑

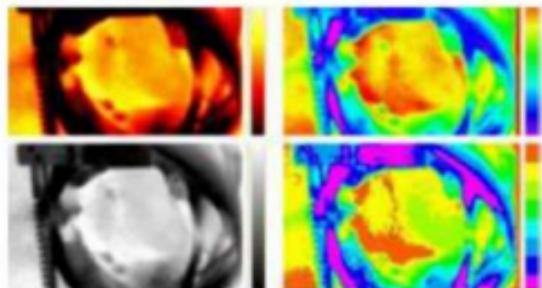
2^8



Zastępowanie kolorów (*Replace colour*)

Pseudokoloryzacja

Sztuczna wizualizacja
cechy pomiarowej innej
niż fale widzialne na
obrazie, np.: termografia

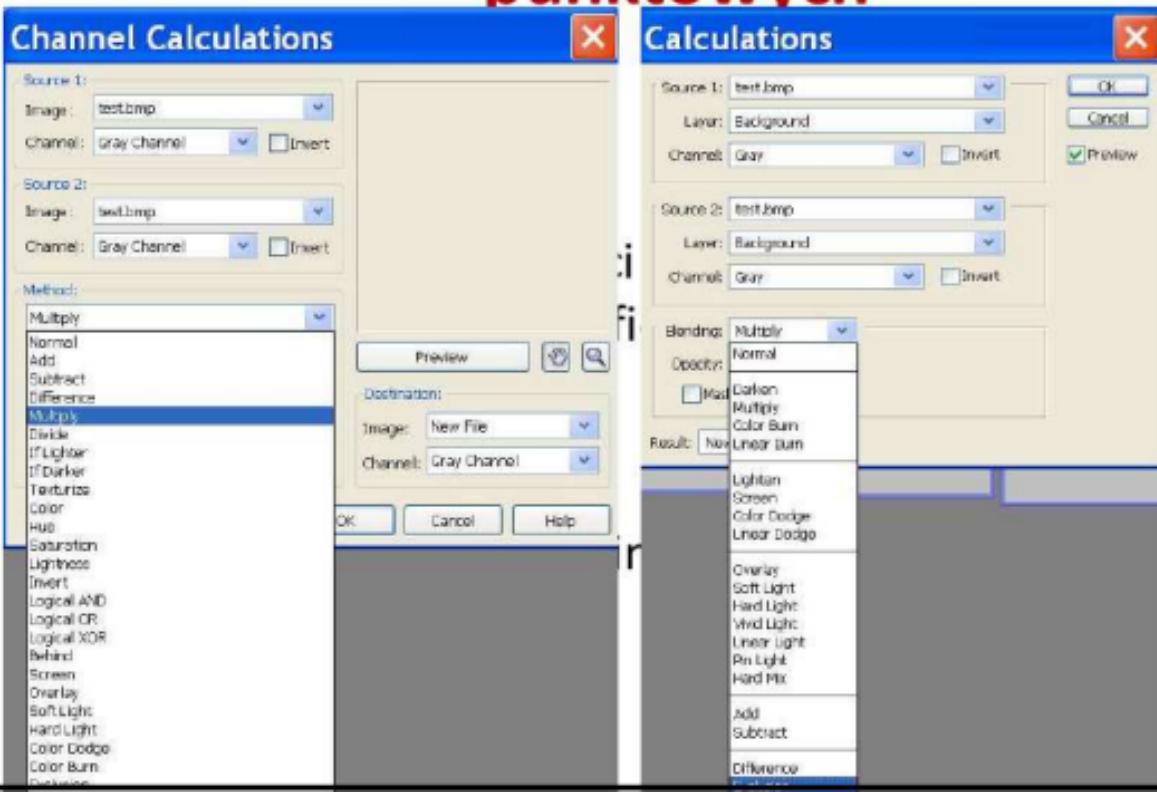


Typowe skale pseudokoloru:
Monochrome(1) - min od bieli
Monochrome(2) - min od czerni
Spectrum – widmo widzialne
Rainbow - tencza z inwersją zieleni
BlackBody – żółto-czerwona
Iron – tylko czerwona

Operacje punktowe - podsumowanie

- Zalety: łatwe do implementacji i szybkie
 - Implementowane przez LUT (manipulacje na palecie szarości lub barw) bo dziedzina funkcji jest skończona
 - Dla różnowartościowych funkcji (bezstratnych) istnieje przekształcenie odwrotne
 - Dla nieróżnowartościowych funkcji (stratnych) nie istnieje przekształcenie odwrotne – większość praktycznie stosowanych
- Wady: ignorują przestrzenną zależność wartości intensywności w obrazie tzn. lokalną charakterystykę przestrzenną obrazu

Praktyczne wykorzystanie operacji punktowych



Operacje sąsiedztwa (kontekstowe o małym otoczeniu)

Operacje na obrazach

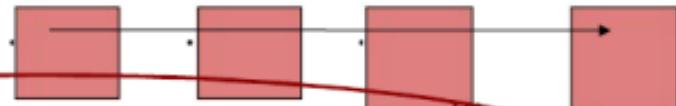
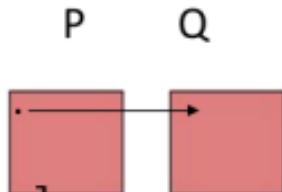
- Operacje punktowe (jednopunktowe):

Jednoargumentowe

$$[q(i, j)] = f[p(i, j)]$$

Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$

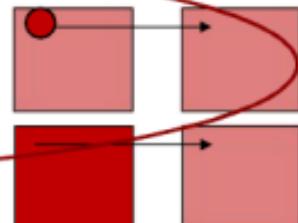


- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

- Operacje globalne transformaty

$$[q(i, j)] = f[P]$$

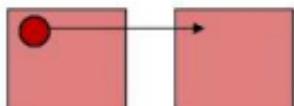


Operacje kontekstowe (sąsiedztwa)

Są to operacje, w których na wartość zadanego piksela obrazu wynikowego współrzędnych (i,j) mają wpływ wartości:

- pikseli o współrzędnych (i,j) obrazu pierwotnego p , oraz
- pikseli go otaczających, czyli tzw. **otoczenia**

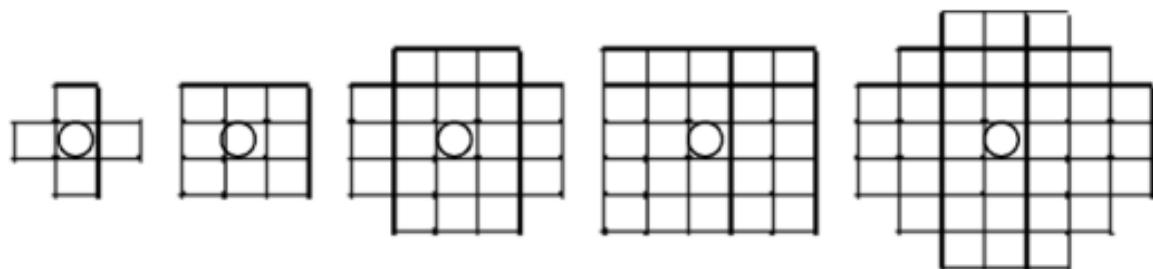
$$[q(i,j)] = f[p(i,j), p(i-1, j-1), p(i+1, j+1), ..]$$



Otoczenie

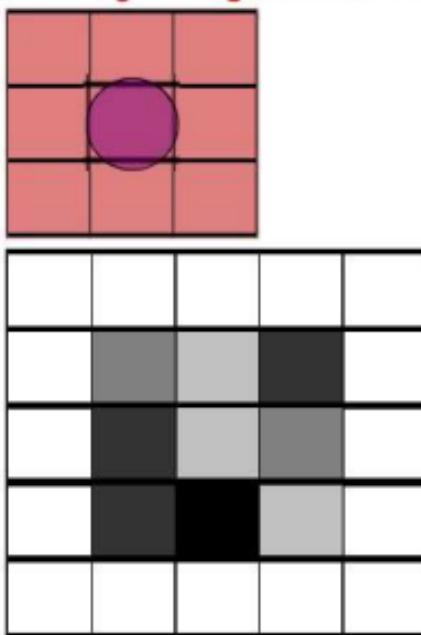
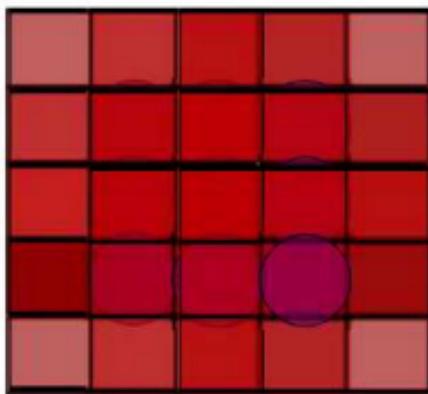
Otoczenie definiujemy określając jego:

- Kształt: kwadratowe, prostokątne, kołowe
- Wielkość: 3×3 , 5×5 , 7×7 , 9×9 , 11×11 , itd..



Symetria otoczenia

Proces liczenia operacji sąsiedztwa

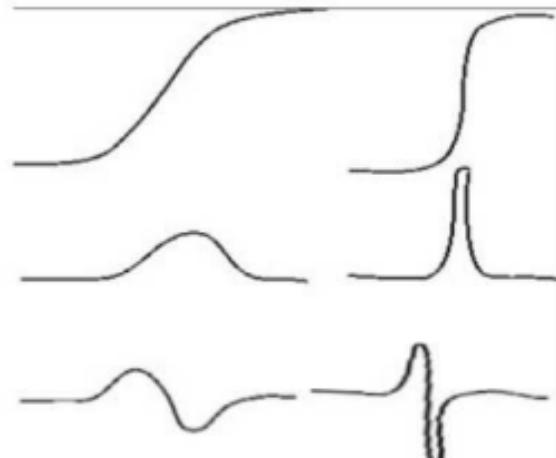
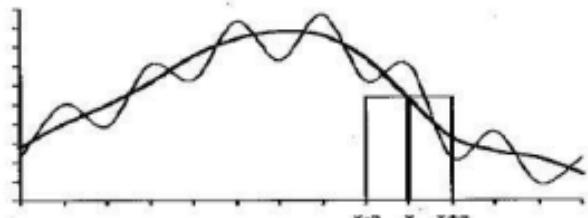


Problem
brzegów
marginów

Wynik operacji zależy od wielkości maski, ale głównie od funkcji zdefiniowanej w punkcie i jego otoczeniu.

Działania matematyczne uwzględniające sąsiedztwo

- Całkowanie
(sumowanie, uśrednianie)



- Różniczkowanie (pierwsza pochodna)
- Laplasjany (druga pochodna)

Funkcje operacji sąsiedztwa

- Wygładzanie (całkowanie przestrzenne)
wytlumianie szumu
- Uwypuklające krawędzie (różniczkowanie przestrzenne)
wydobywanie informacji przez zwiększenie kontrastu
- Konturowanie (operatory oparte na gradientach,
laplasjany, różniczki przestrzenne)
wydobywanie informacji przez wskazanie "zera"
- Analiza kształtu i położenia obiektów
(morphologia matematyczna)

Matematyczny podział operacji sąsiedztwa

- **Liniowe** (oparte na pewnej liniowej operacji, polegającej na wykonaniu liniowej kombinacji wartości wybranych pikseli obrazu wejściowego)
 - Są proste w implementacji, tak na poziomie oprogramowania, jak i procesorów sprzętowych.
 - Łączne i separowalne (rozdzielcze), przemienne.
 - Dają efekty odpowiadające manipulacji pewnymi zakresami częstotliwości.
- **Nieliniowe** (oparte na funkcjach nieliniowych, np. statystycznych, logicznych i morfologii matematycznej)
 - Są czasochłonne i często skomplikowane.
 - Mają bogatsze możliwości.
 - Nie można ich interpretować w kategoriach manipulowania tylko określonymi częstotliwościami

Operacje **wygładzania** stanowią praktyczną realizację *filtracji dolnoprzepustowej* (FD) i dzielą się na operacje filtracji *liniowej i nieliniowej*.

Operacje filtracji nieliniowej dzielą się na operacje filtracji *logicznej i medianowej*.

Operacje **wyostrzania** stanowią praktyczną realizację *filtracji górnoprzepustowej* (FG) i dzielą się na operacje filtracji *gradientowej i laplasjanowej*

Operacje wygładzania

Podstawowe zadanie wygładzania: usuwanie zakłóceń z obrazu

Filtracja liniowa (metody *konwolucyjne*, tzn. uwzględniające pewne otoczenie przetwarzanego piksela):

$$g(x, y) = \sum_{k=1}^n w_k f_k(x, y)$$

n - liczba punktów (pikseli) otoczenia wraz z pikselem przetwarzanym

$f(x, y)$ - wartość piksela o współrzędnych x, y obrazu pierwotnego

$g(x, y)$ - wartość piksela o współrzędnych x, y obrazu wynikowego

w_k - waga k -tego piksela otoczenia

Przykład

obraz $[f(x,y)]$

14	15	13	15
12	14	0	15
13	12	12	14
15	14	14	12

otoczenie 3×3

$[f(x,y)]$

15	13	15
14	0	15
12	12	14

$[g(x,y)]$

15	13	15
14	11	15
12	12	14

średnia ważona

$\rightarrow g(x, y)$

$$\begin{aligned}
 &= w_1 f(x-1, y-1) + w_2 f(x-1, y) \\
 &+ w_3 f(x-1, y+1) + w_4 f(x, y-1) \\
 &+ w_5 f(x, y) + w_6 f(x, y+1) \\
 &+ w_7 f(x+1, y-1) + w_8 f(x+1, y) + w_9 f(x \\
 &+ 1, y+1)
 \end{aligned}$$

w_1 $x-1, y-1$	w_2 $x-1, y$	w_3 $x-1, y+1$
w_4 $x, y-1$	w_5 x, y	w_6 $x, y+1$
w_7 $x+1, y-1$	w_8 $x+1, y$	w_9 $x+1, y+1$

y

Ruchoma średnia

Filtracja liniowa

X

Sposoby zapisu operacji filtracji liniowej (wygładzanie liniowe)

1. Za pomocą *macierzy wag*
2. Za pomocą *maski filtracji dolnoprzepustowej* (FD)

Macierz wag

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Maska filtracji
dolnoprzepustowej

1	1	1
1	1	1
1	1	1

K = 9

K - współczynnik maski

Przykłady macierzy wag i masek operacji filtracji liniowej

Macierz wag

1/10	1/10	1/10
1/10	2/10	1/10
1/10	1/10	1/10

Maska filtracji dolnoprzepustowej

1	1	1
1	2	1
1	1	1

K = 10

1/16	2/16	1/16
2/16	4/16	2/16
1/16	2/16	1/16

1	2	1
2	4	2
1	2	1

K = 16

Operacje wyostrzania

Metoda: konwolucja + maska *filtracji górnoprzepustowej(FG)*.

W wyostrzaniu stosuje się metody numeryczne aproksymujące pochodną.

Zadanie wyostrzania:

- podkreślenie na obrazie konturów obiektów
- podkreślenie na obrazie punktów informatycznych (np. wierzchołki dla wielokątów, zakończenia, skrzyżowania, rozgałęzienia linii dla rysunków technicznych, wykresów lub pisma).

Inne zadania wyostrzania: *wydobycie i uwypuklenie krawędzi obiektu.*

Detekcja (wykrywanie) krawędzi (edge detection)

Jest to technika segmentacji obrazu, polegająca na znajdowaniu pikseli krawędziowych przez sprawdzanie ich sąsiedztwa.

4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8

Krawędź

Zbiór pikseli na krzywej mający taką właściwość, że piksele w ich sąsiedztwie, lecz po przeciwnych stronach krzywej mają różne poziomy jasności.

Cel detekcji

znalezienie lokalnych nieciągłości w poziomach jasności obrazu oraz granic obiektów zawartych w obrazie.

Opis matematyczny operacji wyostrzania

Model krawędzi: linia prosta **separująca** dwa obszary o różnej **intensywności (jasności)** I_1 i I_2 .

Użycie funkcji $u(z)$ do matematycznego opisu krawędzi

$$u(z) = \begin{cases} 1 & \text{dla } z > 0 \\ \frac{1}{2} & \text{dla } z = 0 \\ 0 & \text{dla } z < 0 \end{cases}$$

Założenia:

Jeśli $\delta(t)$ - impuls Diraca, to:

$$u(z) = \int_{-\infty}^z \delta(t) dt$$

Krawędź leży wzdłuż linii prostej opisanej równaniem:

$$x \sin \varphi - y \cos \varphi + \rho = 0 \text{ (postać normalna prostej)}$$

Intensywność obrazu:

$$f(x, y) = I_1 + (I_2 - I_1) u(x \sin \varphi - y \cos \varphi + \rho)$$

Cyfrowa wersja gradientu

Pochodna pionowa G_x funkcji $f(x,y)$

$$G_x \stackrel{\text{def}}{=} [f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1)] - [f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)]$$

maska:

	$y-1$	y	$y+1$
$x-1$	-1	-2	-1
x	0	0	0
$x+1$	1	2	1

Pochodna pozioma G_y funkcji $f(x,y)$

$$G_y \stackrel{\text{def}}{=} [f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)] - [f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)]$$

maska:

	$y-1$	y	$y+1$
$x-1$	-1	0	1
x	-2	0	2
$x+1$	-1	0	1

$$G(x,y) = \sqrt{G_x^2 + G_y^2}$$

Cyfrowa wersja laplasjanu

$$L(x, y) = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)]$$

maska:

	$y-1$	y	$y+1$
$x-1$	0	1	0
x	1	-4	1
$x+1$	0	1	0

Własności:

Gradient: wrażliwy na intensywność zmiany; używany tylko do detekcji krawędzi;

Laplasjan: podaje dodatkową informację o położeniu piksela względem krawędzi (po jasnej czy po ciemnej stronie).

Uwaga: Dla operacji wyostrzania współczynnik maski **K=1**

Przykłady

$f(x,y)$

4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8
4	4	4	8	8	8	8



0	1	0
1	-4	1
0	1	0

$L(x,y)$

x	x	x	x	x	x	x
x	0	4	-4	0	0	x
x	0	4	-4	0	0	x
x	0	4	-4	0	0	x
x	x	x	x	x	x	x

jasna strona (-4)

$g(x,y)$

$f(x,y)$

4	4	8	8	8	8	8
4	4	4	8	8	8	8
4	4	4	4	8	8	8
4	4	4	4	4	8	8
4	4	4	4	4	4	8

$L(x,y)$

x	x	x	x	x	x	x
x	0	8	-4	0	0	x
x	0	0	8	-4	0	x
x	0	0	0	8	-4	x
x	x	x	x	x	x	x

jasna strona (-4)

ciemna strona (+8)

Laplasjany

Tekst

Dyskretna forma drugiej pochodnej

0	-1	0		-1	-1	-1		1	-2	1		-1	-1	-1		0	-1	0
-1	4	-1		-1	8	-1		-2	4	-2		-1	9	-1		-1	5	-1
0	-1	0		-1	-1	-1		1	-2	1		-1	-1	-1		0	-1	0

Tekst

Tekst

Tekst

Tekst

Tekst

Metody operacji na pikselach wchodzących w skład skrajnych kolumn i wierszy

1. Pozostawienie wartości pikseli bez zmian
2. Wartości pikseli są nieokreślone (xxxxxxxxxx)
3. Nadanie pikselom wartości arbitralnie zadanych przez użytkownika (np. same wartości „0”, „15”, „10” itd.)
4. Operacje z zastosowaniem kolumn i wierszy pomocniczych (zdublowanie (powielenie) skrajnych wierszy i kolumn)
5. Operacje z wykorzystaniem pikseli z **istniejącego sąsiedztwa**.
 - Lewa skrajna kolumna (oprócz pikseli górnego i dolnego rogu) – kierunki 0,1,2,6,7,
 - Lewa skrajna kolumna piksel w górnym rogu – kierunki 0, 6,7,
 - Lewa skrajna kolumna (piksel w dolnym rogu) – kierunki 0,1,2,
 - Prawa skrajna kolumna (oprócz pikseli górnego i dolnego rogu) – kierunki 2,3,4,5,6,
 - Prawa skrajna kolumna piksel w górnym rogu – kierunki 4,5,6,
 - Prawa skrajna kolumna (piksel w dolnym rogu) – kierunki 2,3,4,
 - Górnny skrajny wiersz (oprócz pikseli z lewego i prawego rogu) – kierunki 4,5,6,7,0
 - Dolny skrajny wiersz (oprócz pikseli z lewego i prawego rogu) – kierunki 0,1,2,3,4.

Zdublowanie/powielenie skrajnych wierszy i kolumn

x	x	x	x	x	x	x
x	0	4	-4	0	0	x
x	0	4	-4	0	0	x
x	0	4	-4	0	0	x
x	x	x	x	x	x	x

- Odbicie v2 (reflect_101) edcb|abcdef|edcb
- Odbicie v1 (reflect) edcba|abcdef|fedcb
- Powielenie (replicate) aaaa|abcde|eeee
- Stała n (constant) nnnn|abcdef|nnnn

Metody skalowania tablic obrazów wynikowych

Cel skalowania: sprowadzanie wartości pikseli do zakresu $[0, (M-1)]$

Metoda proporcjonalna

$$g'(x, y) = \frac{g(x, y) - g(x, y)_{\min}}{g(x, y)_{\max} - g(x, y)_{\min}} \cdot (M - 1)$$

Własność:

Równomierne przeskalowanie wszystkich pikseli obrazu.

Końcowy efekt: obraz z zakresu $[0, (M-1)]$

Metoda trójwartościowa

$$g'(x, y) = \begin{cases} 0 & \text{dla } g(x, y) < 0 \\ E[(M - 1) / 2] & \text{dla } g(x, y) = 0 \\ M - 1 & \text{dla } g(x, y) > 0 \end{cases}$$

Zastosowanie

obrazy o jednolitym tle i dobrze widocznych obiektach - np.
obrazy binarne. Efekt: czarno-biała krawędź na szarym tle.

Metoda obcinająca

$$g'(x, y) = \begin{cases} 0 & \text{dla } g(x, y) < 0 \\ g(x, y) & \text{dla } 0 \leq g(x, y) \leq M - 1 \\ M - 1 & \text{dla } g(x, y) > M - 1 \end{cases}$$

Filtracja statystyczna

Filtryle: **medianowy**, **maksymalny**, **minimalny**, oparty na **najbardziej prawdopodobnej wartości**

Filtryle działają na otoczeniu, ale ich wpływ na wartość w analizowanym punkcie po filtracji wyraża się warunkiem nie maską.

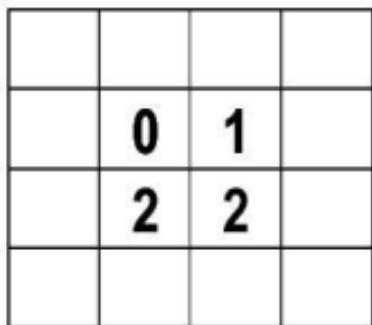
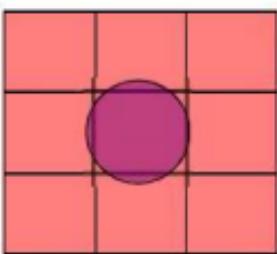
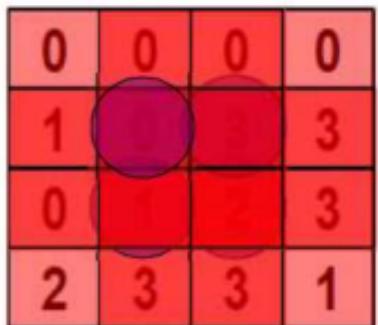
Medianą – wartość środkowa w uporządkowanym ciągu liczb

Maksymalna – największa w uporządkowanym ciągu liczb

Minimalna – najmniejsza w uporządkowanym ciągu liczb

Najbardziej prawdopodobna – najczęściej występująca w otoczeniu

Medianą i pozostałe filtry statystyczne



1, 2, 2, 4, 5, 6, 7, 8, 9

0, 0, 0, 0, 0, 1, 1, 2, 3 mediana=0, min=0, max=3, najbardziej prawdopodobna=0

0, 0, 0, 0, 1, 2, 3, 3, 3 mediana=1, min=0, max=3, najbardziej prawdopodobna=0

0, 0, 1, 1, 2, 2, 3, 3, 3 mediana=2, min=0, max=3, najbardziej prawdopodobna=3

0, 0, 1, 1, 2, 2, 3, 3, 3 mediana=2, min=0, max=3, najbardziej prawdopodobna=3

Metoda filtracji medianowej (wygładzanie medianowe)

[p(i,j)]				
15	15	14	13	14
14	14	13	12	12
1	1	1	11	12
0	0	1	1	10
0	0	0	0	1

p(i,j):

p(2,2): 1 1 12 13 14 14 14 15 15
p(2,3): 0 1 11 12 13 13 14 14 15
p(2,4): 0 11 12 12 12 13 13 14 14
p(3,2): 0 0 1 1 1
p(3,3): 0 0 1 1 1
p(3,4): 0 1 1 10 11
p(4,2): 0 0 0 0 0
p(4,3): 0 0 0 0 1
p(4,4): 0 0 1 1 1

q(i,j):

q(i,j):

15	15	14	13	14
14	14	13	12	12
1	1	1	11	12
0	0	1	1	10
0	0	0	0	1

Przykładowy obraz pierwotny

Obraz wynikowy

z pozostawieniem wartości pikseli marginesów -bez zmian

Filtracja medianowa (wygładzanie medianowe)

Usuwanie zakłóceń **bez rozmywania krawędzi**
(por. metodę filtracji liniowej)

Medianą - wartość **środkowa** (w sensie położenia w ciągu wartości uporządkowanych)

Filtracja dwuetapowa i łączenie filtrów

- 1szy etap – maska **f**,
- 2gi etap (filtracja obrazu otrzymanego po 1-szym etapie) – maska **g**

Filtracja jednoetapowa równoważna filtracji dwuetapowej:

-maska **m**

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	f1	f2	f3	0	0
0	0	f4	f5	f6	0	0
0	0	f7	f8	f9	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Maska **f** (rozmiar 3x3) ,
współczynnik maski **Kf**

g1	g2	g3
g4	g5	g6
g7	g8	g9

Maska **g** (rozmiar 3x3) ,
współczynnik maski: **Kg**

m1	M2	m3	m4	m5
m6	m7	m8	m9	m10
m11	m12	m13	m14	m15
m16	m17	m18	m19	m20
M21	m22	m23	m24	m25

Maska **m** (rozmiar 5x5) ,
współczynnik maski: **Km = Kf*Kg**

Przykładowe podstawienie:

$f_1=f_2=f_3=\dots=f_9=1$,
 $g_1=g_2=g_3\dots=g_9=1$

Maska g:

1	1	1
1	1	1
1	1	1

Maska wynikowa m:

1	2	3	2	1
2	4	6	4	2
3	6	9	6	3
2	4	6	4	2
1	2	3	2	1

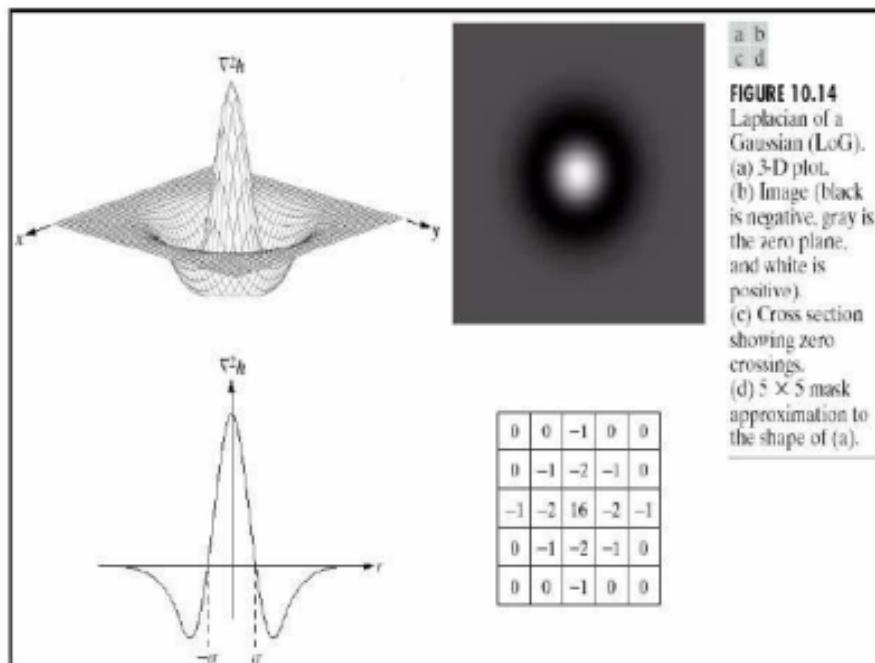
Maska f:

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	0	1	1	1	0	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

$$\begin{aligned}
 m1 &= 0*g1+0*g2+0*g3+0*g4+0*g5+ \\
 &\quad +0*g6+0*g7+0*g8+ f1*g9 \\
 m2 &= 0*g1+0*g2+0*g3+0*g4+0*g5+ \\
 &\quad +0*g6+0*g7+f1*g8+ f2*g9 \\
 m3 &= 0*g1+0*g2+0*g3+0*g4+0*g5+ \\
 &\quad +0*g6 +f1*g7+f2*g8+ f3*g9 \\
 &\vdots \\
 m25 &= f9*g1+0*g2+0*g3+0*g4+0*g5+ \\
 &\quad 0*g6+0*g7+0*g8+ 0*g9
 \end{aligned}$$

Współczynnik maski: $K_m = K_f * K_g = 1/9 * 1/9 = 1/81$

Laplasjan filtra gaussowskiego



Filtr jednocześnie wygładzający i wyostrzający

Filtracja oparta na operacjach logicznych

	a	
b	x	c
	d	

$$X' = \begin{cases} a & \text{if } a = d \\ \text{else } X \end{cases}$$

$$X' = \begin{cases} b & \text{if } b = c \\ \text{else } X \end{cases}$$

$$X' = \begin{cases} a & \text{if } a = b = c = d \\ \text{else } X \end{cases}$$

Czterospójne otoczenie punktu

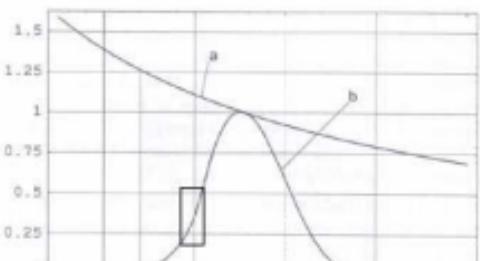
– eliminacja izolowanych punktów i poziomych linii o pojedynczej grubości

– eliminacja izolowanych punktów i pionowych linii o pojedynczej grubości

– eliminacja izolowanych punktów

1	1	1	0
1	1	1	1
1	0	1	1
0	0	0	0
1	1	0	1
1	1	1	1
1	1	0	1
1	1	0	1

Przykład działania filtracji logicznych



W celu zrealizowania zadania przedstawionego na rysunku skonstruujmy logikę wyjścia dla jednego z kanałów światła. W tym celu skonstruujmy funkcję logiczną, której wyjście będzie miało wartość 1, jeśli sygnał wejściowy z kanału światła jest większy od 0,55, a wyjście ma wartość 0, jeśli sygnał wejściowy z kanału światła jest mniejszy od 0,55. W celu zrealizowania zadania powinniśmy skorzystać z funkcji AND, funkcji OR oraz funkcji NOT. Oznaczmy sygnał wejściowy z kanału światła jako A . Wówczas sygnał wyjściowy powinien mieć postać: $Y = \overline{A} \cdot \text{funkcja AND} (A, 0,55) + A \cdot \text{funkcja OR} (A, 0,55)$.



MATERIAŁ DO WYKŁADU I ĆWICZEŃ LABORATORYJNYCH:

Wykład

Materiał podstawowy:

M. Doros, **Przetwarzanie obrazów**, Skrypt WSISiZ (Rozdz.1)

Materiał uzupełniający:

R.Tadeusiewicz, P.Korohoda: **Komputerowa analiza i przetwarzanie obrazów**, Kraków 1997. <http://winntbg.bg.agh.edu.pl/skrypty2/0098/>

Ćwiczenia Laboratoryjne

Materiał podstawowy:

M.Doros, A. Korzyńska, M.Przytulska, H.Goszczyńska: „**Przetwarzanie Obrazów, ćwiczenia laboratoryjne**”, Skrypt WSISiZ (Rozdz.1)

Operacje na obrazach (III)

WYKŁAD 4
Dla studiów niestacjonarnych
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

Operacje na obrazach

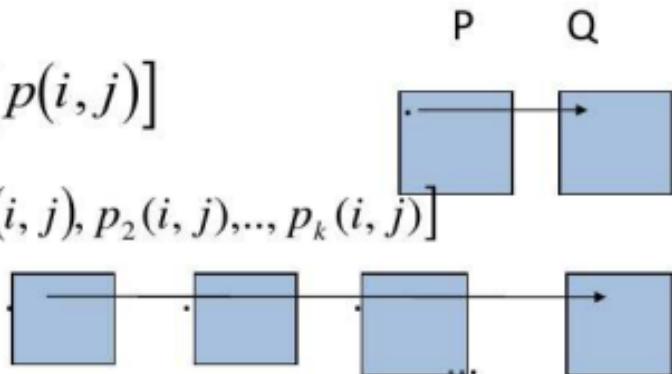
- Operacje punktowe (jednopunktowe):

Jednoargumentowe

$$[q(i, j)] = f[p(i, j)]$$

Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$

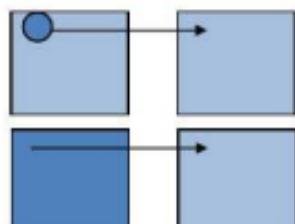


- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

- Operacje globalne transformaty

$$[q(i, j)] = f[P]$$



Z poprzedniego wykładu

Filtracja oparta na operacjach logicznych

	a	
b	x	c
	d	

$$X' = \begin{cases} a & \text{if } a = d \\ \text{else } X \end{cases}$$

$$X' = \begin{cases} b & \text{if } b = c \\ \text{else } X \end{cases}$$

$$X' = \begin{cases} a & \text{if } a = b = c = d \\ \text{else } X \end{cases}$$
 — eliminacja izolowanych punktów

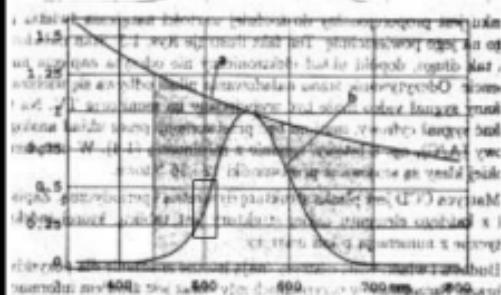
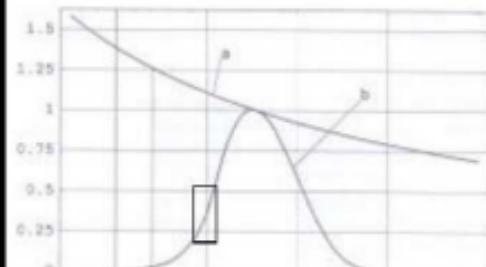
Czterospójne otoczenie punktu

— eliminacja izolowanych punktów i poziomych linii o pojedynczej grubości

— eliminacja izolowanych punktów i pionowych linii o pojedynczej grubości

1	1	1	0
1	1	1	1
1	0	0	1
0	0	0	0
1	1	1	1
1	0	0	0
1	1	0	1
1	1	0	1

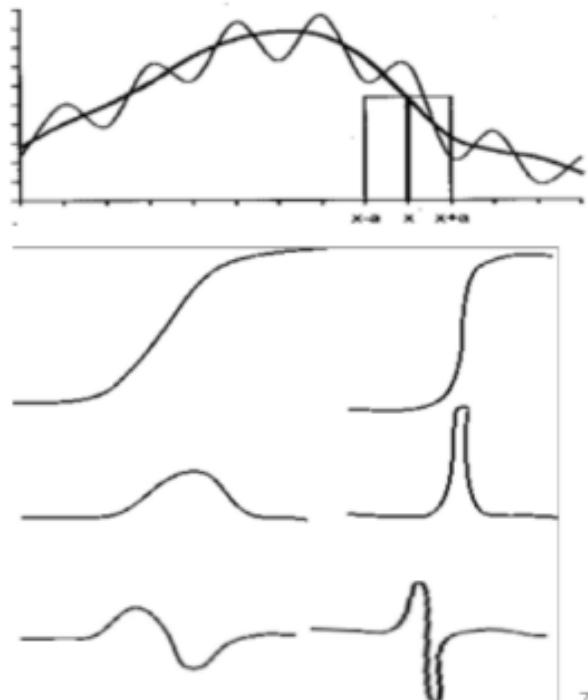
Przykład działania filtracji logicznych



Krawędziowanie i operatory krawędziujące

Działania matematyczne uwzględniające sąsiedztwo

- Całkowanie
(sumowanie, uśrednianie)
- Różniczkowanie (pierwsza pochodna)
- Laplasjany (druga pochodna)



Krawędź obrazu widoczna w przekroju (xz)

- obraz pierwotny
- po obróbce gradientowej
(1-sza pochodna)
- po obróbce laplasjanowej
(2-ga pochodna)

Krawędź



Pierwsza
pochodna

$$\left(\frac{\partial f}{\partial x} \right)$$



Druga
pochodna

$$\left(\frac{\partial^2 f}{\partial x^2} \right)$$



Operator
gradientowy

-1	0	+1
----	---	----



Operator typu
drugiej pochodnej

-1	0	+1
----	---	----



Rozróżnienie operacji związanych z krawędziami

- Krawędziowanie laplasjanem
- Uwypuklenia krawędzi (wystrzenia)
- Detekcja krawędzi w określonym kierunku
- Detekcja krawędzi w dowolnym kierunku



Detekcja (wykrywanie) krawędzi (ang. edge detection)

Jest to jedna z technik segmentacji obrazu, polegająca na znajdowaniu pikseli krawędziowych przez sprawdzanie ich sąsiedztwa.

Krawędź

Zbiór pikseli na krzywej mający taką właściwość, że piksele w ich sąsiedztwie, lecz po przeciwnych stronach krzywej mają różne poziomy jasności.

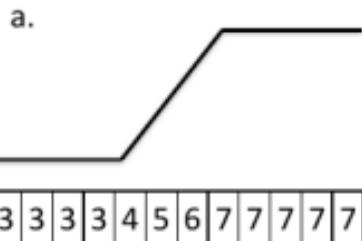
Cel detekcji

znalezienie lokalnych **nieciągłości** w poziomach jasności obrazu oraz granic obiektów zawartych w obrazie.

Wyostanie krawędzi obrazu widoczne na przekroju (xz)

- a. obraz pierwotny
- b. obraz wynikowy po obróbce i po dodaniu w sposób ważony jasności odpowiednich pikseli obrazu pierwotnego
- c. maska laplasjanu
- d. laplasjan

Rezultat **uwypuklenie** (wzmacnianie) krawędzi (edge enhancement)



c.

-1	2	-1
----	---	----

d.

0	0	-1	0	0	0	1	0	0	0
---	---	----	---	---	---	---	---	---	---

UWAGA!! Metoda ma podstawy w procesach fizjologicznych zachodzących w oku ludzkim zwanych hamowaniem obocznym, jaj efekty opisał w Ernst Mach w 1860 roku

Obraz o wzmacnionych konturach obiektów

Wyostrzenie przez **złożenie** obrazów:

- **wejściowego,**
- **po operacji zadanej laplasjanem lub filtrem rozmywającym/wygładzającym** następnie przeskalowanie stopni szarości

0	-1	0	-1	-1	-1	1	-2	1	-1	-1	-1	0	-1	0
-1	4	-1	-1	8	-1	-2	4	-2	-1	9	-1	-1	5	-1
0	-1	0	-1	-1	-1	1	-2	1	-1	-1	-1	0	-1	0

Wykonanie operacji konstrukcji tzn. **Unsharpen Mask**

Odjęcie obrazu wygładzonego (rozmytego np. filtrem gaussa) od obrazu oryginalnego daje obraz regionów wyostrzenia

Jak otrzymać podniesienie kontrastu na krawędziach z wykorzystaniem unsharp mask



Porównanie z obrazem o maksymalnie dobranym kontraście przez rozciągnięcie od bieli do czerni



Przykłady masek FG detekcji krawędzi

Tekst

maska 1

	$y-1$	y	$y+1$
$x-1$	1	-2	1
x	-2	5	-2
$x+1$	1	-2	1

maska 2

-1	-1	-1
-1	9	-1
-1	-1	-1

maska 3

0	-1	0
-1	5	-1
0	-1	0

ogólnie

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Obliczanie wartości piksela obrazu wynikowego

$$g(x, y) = w_1 f(x-1, y-1) + w_2 f(x-1, y) + w_3 f(x-1, y+1) + \\ + w_4 f(x, y-1) + w_5 f(x, y) + w_6 f(x, y+1) + \\ + w_7 f(x+1, y-1) + w_8 f(x+1, y) + w_9 f(x+1, y+1)$$

Uwaga

W przypadku operacji wyostrzania oraz detekcji krawędzi współczynnik maski K=1.

Filtry uwypuklające krawędzie (w założonym kierunkowe lub ogólnie)

- Zmniejszają udział lub wręcz usuwają informację o obszarach jednorodnych, uwypuklając informację o krawędziach
- Suma współczynników maski równa 0
- Konieczność skalowania

1	0	-1
1	0	-1
1	0	-1

Tekst

1	-1	-1
1	-2	-1
1	1	1

Tekst

Tekst

1	-2	1
-2	4	-2
1	-2	1

Tekst

Metoda specjalnego gradientu

Stosowana w przypadkach, gdy metody filtracji górnoprzepustowej (FG) powodują wzmacnienie zakłóceń w obszarach leżących **wewnątrz** konturu.

Zasada

Krawędź uznana jest za istniejącą, jeśli wartość gradientu intensywności w pewnych punktach przekracza ustalony próg.

Metody : **Roberts, Sobela, Prewitta, Kirsha, Canny** oraz **filtры резьбиace**

f_0	f_1	f_2
f_3	f_4	f_5
f_6	f_7	f_8

Metoda Robertsa

$$R(i,j) = \sqrt{(f_4 - f_8)^2 + (f_7 - f_5)^2}; \quad \alpha = -\frac{\pi}{4} + \operatorname{tg}^{-1}\left(\frac{f_7 - f_5}{f_4 - f_8}\right)$$

gdzie:

$R(i,j)$ - specjalny gradient w punkcie (i,j)

α - kierunek gradientu intensywności.

Metoda Sobela

dwie składowe gradientu:

$$S_x = (f_2 + 2f_5 + f_8) - (f_0 + 2f_3 + f_6)$$

$$S_y = (f_6 + 2f_7 + f_8) - (f_0 + 2f_1 + f_2)$$

$$S(x,y) = \sqrt{S_x^2 + S_y^2}$$

Maski konwolucyjne do konstrukcji operatorów krawędziowych

Roberts

1	0
0	-1

G_{225}

0	-1
1	0

G_{45}

Sobel:

-1	0	1
-2	0	2
-1	0	1

G_x

-1	-2	-1
0	0	0
1	2	1

G_y

Prewitt:

1	0	-1
1	0	-1
1	0	-1

G_x

1	1	1
0	0	0
-1	-1	-1

G_y

Kirsha:

3	3	-5
3	0	-5
3	3	-5

G_x

3	3	3
3	0	3
-5	-5	-5

G_y

Operatory krawędziowania

Konstruowane przez nieliniową kombinację dwóch prostopadłych kierunków gradientu (liniowych transformacji)

dokładnej

$$G = \sqrt{Gx^2 + Gy^2}$$

Robertса

Gx	1	0	lub	-1	0
	0	-1		0	1
	0	1	lub	0	-1
Gy	-1	0		1	0

przybliżonej

$$G = |Gx| + |Gy|$$

Sobela

	1	0	-1		-1	0	1
Gx	2	0	-2	lub	-2	0	2
	1	0	-1		-1	0	1
	1	2	1		-1	-2	-1
Gy	0	0	0	lub	0	0	0
	-1	-2	-1		1	2	1

Metoda Kirscha

Wartość piksela (i, j) jest zmieniana według wzoru:

$$g(i, j) = \max_{k=0}^7 \left\{ 1, \max \left[|5S_k - 3T_k| \right] \right\}$$

gdzie:

$$S_k = f_k + f_{k+1} + f_{k+2}$$

$$T_k = f_{k+3} + f_{k+4} + f_{k+5} + f_{k+6} + f_{k+7}$$

Numeracja pikseli:

f - obraz źródłowy,

g - obraz wynikowy

0	1	2
7	i,j	3
6	5	4

Indeksy punktów obrazu f - wartości modulo 8

Maski odpowiadające operatorowi Kirsch (maski Kirsch)

N

3	3	3
3	0	3
-5	-5	-5

NE

3	3	3
-5	0	3
-5	-5	3

E

-5	3	3
-5	0	3
-5	3	3

SE

-5	-5	3
-5	0	3
3	3	3

S

-5	-5	-5
3	0	3
3	3	3

SW

3	-5	-5
3	0	-5
3	3	3

W

3	3	-5
3	0	-5
3	3	-5

NW

3	3	3
3	0	-5
3	-5	-5

Operatory kierunkowe

$[p(i,j)]$					$j \rightarrow$
$i \downarrow$	15	15	14	13	14
	14	(14)	(13)	(12)	12
	4	(5)	(12)	(11)	12
	5	(3)	(4)	(5)	6
	1	0	3	2	4

SW

1	-1	-1
1	-2	-1
1	1	1

$$q_{22} = 15 + 14 + 4 + 15 + 12 - 28 - 15 - 14 - 13 = -10$$

$$q_{23} = 15 + 14 + 15 + 12 + 11 - 26 - 14 - 13 - 12 = 2$$

$$q_{32} = 14 + 4 + 5 + 3 + 4 - 30 - 14 - 13 - 12 = -39$$

$$q_{42} = 4 + 5 + 1 + 3 - 6 - 15 - 12 - 4 = -24$$

$$q_{33} = 14 + 15 + 3 + 4 + 5 - 24 - 13 - 12 - 11 = -19$$

$$q_{43} = 15 + 3 + 3 + 4 - 8 - 12 - 11 - 5 = -11$$

$$q_{24} = 14 + 13 + 12 + 11 + 12 - 24 - 13 - 14 - 12 = -11$$

$$q_{34} = 13 + 12 + 4 + 5 + 6 - 22 - 12 - 12 - 12 = -18$$

$$q_{44} = 12 + 4 + 3 + 4 + 2 - 10 - 11 - 12 - 6 = -14$$

$[q(i,j)]$					$j \rightarrow$
$i \downarrow$	x	x	x	x	x
	x	(-10)	(2)	(-1)	x
	x	(-39)	(-19)	(-18)	x
	x	(-24)	(-11)	(-14)	x
	x	x	x	x	x

Filtry kierunkowe rzeźbiące

Wschód			Południowy wschód			Południe			Południowy zachód					
-1	0	1		-1	-1	0		-1	-1	-1		0	-1	-1
-1	1	1		-1	1	1		0	1	0		1	1	-1
-1	0	1		0	1	1		1	1	1		1	1	0
E			SE			S			SW					

1	0	-1		1	1	0		1	1	1		0	1	1
1	1	-1		1	1	-1		0	1	0		-1	1	1
1	0	-1		0	-1	-1		-1	-1	-1		-1	-1	0
Zachód			Połnocny zachód			Połnoc (N)			Połnocny wschód					
23														

Obraz wykorzystujący efekty płaskorzeźby



NA RYBY!

Operator Canny-ego

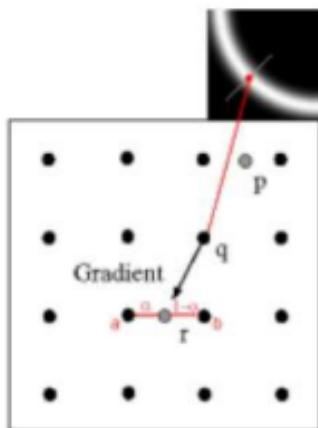
Wieloetapowy algorytm detekcji krawędzi zaproponowany przez twórcę teorii *Computational theory of edge detection* John F. Canny w 1986 r. Jego działanie oparte jest o znane detektory krawędzi (Roberts, Sobela i Prewitta) i **progowanie z histerezą**, które optimalizuje wynik wyrzucając krawędzie rozmyte, nachylone pod kątami niewiele odchylonymi od wcześniej wykrytych, zapobiega przerwaniu krawędzi w miejscach utraty kontrastu, itp..



1986. [*A computational approach to edge detection*](#). IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, 1986, pp. 679–698

Etapy działania operatora Cannyego

- Zamiana obrazu kolorowego na jego wersję monochromatczną w odcieniach szarości
- Rozmycie filtrem gaussowskim o zadanych parametrach: np.: sigma (δ)
- Obliczenie gradientu Sobole/Prewitta i kierunku gradientu według Robertsa
- Wytlumienie lokalnych nie maksymalnych gradientów w obrazie gradientu celem otrzymania cienkiej linii
- Dokonanie podwójnego progowania według zadanych parametrów T_{min} (i T_{max}):
 - $<T_{min}$ - brak krawędzi,
 - $>T_{max}$ – silne krawędzie
 - $>=T_{min}$ i $<=T_{max}$ – krawędzie słabe
- Wykonanie śledzenia krawędzi opartego na histerezie w celu ich uciąglenia: słabe krawędzie będące przedłużeniem silnych są dołączane a nie będące są oznaczane do czyszczenia
- Wyczyszczenie krawędzi



Operacje morfologii matematycznej na obrazach

Operacje morfologiczne

Są to operacje, w których modyfikacja danego piksla obrazu jest uzależniona od **kształtu elementu strukturalnego** (wzorca) użytego w danej operacji oraz od **spełnienia zadanego warunku logicznego**.

Metody te oparte o matematycznej teorii zbiorów i poszukują i uwypukla lub wyrzucają cech, które ujawniają się w obrazie/obiekcie w kontekście narzędzia badawczego, którym jest element strukturalny

Cel: przygotowanie poszczególnych elementów obrazu do etapu analizy obrazów

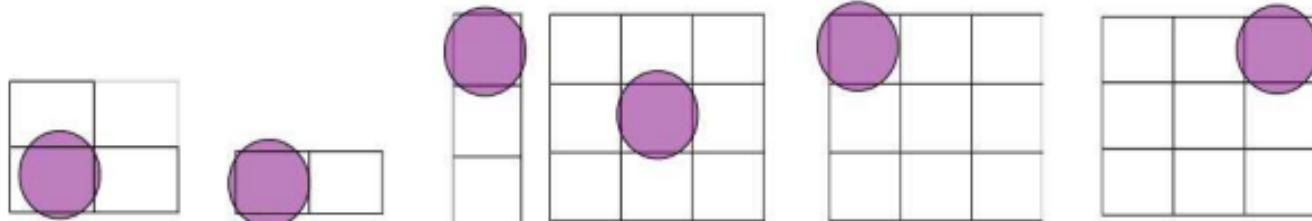
Główna cecha różniąca operacje morfologiczne od operacji punktowych i operacji sąsiedztwa:

W operacjach morfologicznych przekształcana jest tylko ta część pikseli obrazu, których otoczenie jest zgodne z elementem strukturalnym – koincydentne

Operacje morfologii matematycznej na obrazach

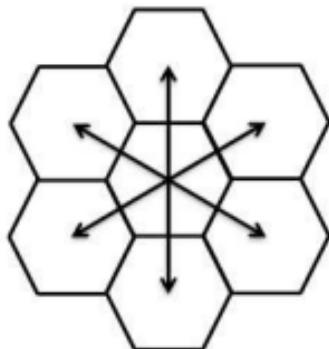
Pozwalają na budowanie złożonych operacji nieliniowych, do analizy kształtu i wzajemnego położenia obiektów.

Fundamentalne pojęcie: **element strukturalnym (strukturujący)** – podzbiór obrazu z wyróżnionym punktem, zwanym często punktem centralnym

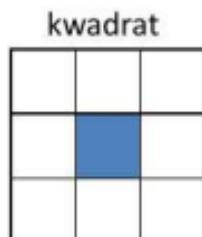


Podstawowy element strukturalny

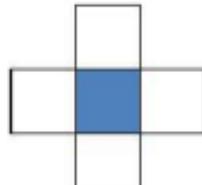
Siatka heksagonalna
(1 rodzaj)



Siatka prostokątna

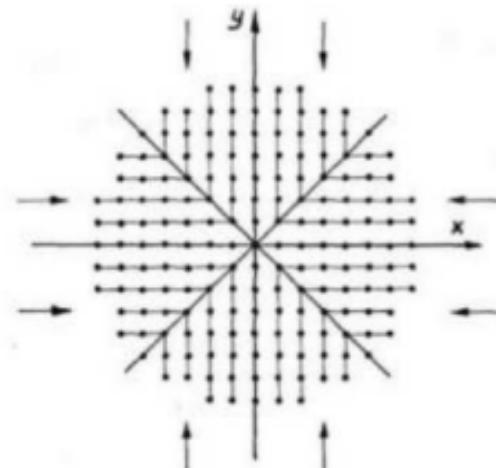


Romb/krzyż



koło

o promieniu
jednostkowym



koło o promieniu $r = 7$

Operacje morfologii matematycznej na obrazach

– w elemencie strukturalnym występują następujące symbole:

- **1** element wskazuje piksel zapalony tzn. wartość obiektu w masce binarnej
- **0** element wskazuje piksel wytłumiony tzn. wartość tła w masce binarnej
- **X** element wskazuje dowolną wartość tzn. wartość tła lub obiektu w masce binarnej

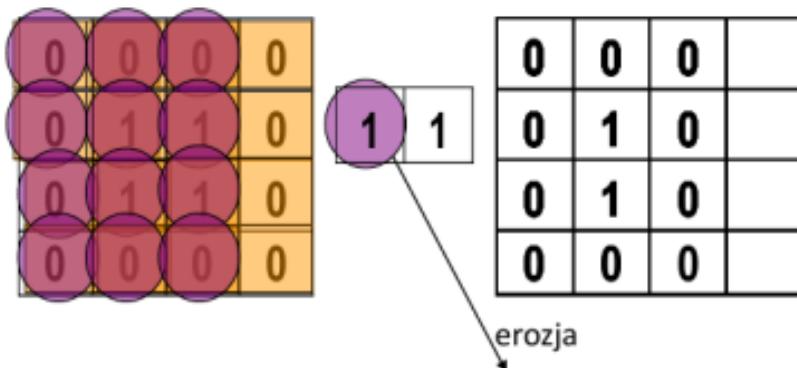
Przekształcenia polegają na pozostawieniu lub zmianie intensywności według pewnej funkcji skojarzonej z nazwą funkcji punktu przykrytego przez punkt centralny elementu strukturalnego jeśli jego otoczenie nie ma lub ma zgodność z założeniami zakodowanemu symbolami w elemencie strukturalnym.

Operacje morfologiczne przekształcają tylko część punktów obrazu

Operacje morfologiczne

1. Operacje morfologiczne binarne działają na tzn. maskach czyli **obrazach binarnych**, w którym obiekty oznaczone są „1” a tło „0” lub na **obrazach w odcieniach szarości**, ale korzystając z uogólnionej definicji
2. Element strukturalny jest przemieszczany po wszystkich punktach obrazu tak, że punkt centralny elementu strukturalnego jest nakładany na kolejne punkty w kolejnych wierszach,
3. W każdym położeniu elementu sprawdza się, czy rzeczywista konfiguracja punktów jest zgodna (koincydentna) ze wzorcem zawartym w elemencie strukturalnym zakodowanym symbolami 1, 0, X
4. W przypadku wykrycia zgodności jest wykonywana operacja związana z filtrem, a w przeciwnym przypadku wartość występująca w obrazie pierwotnym jest przepisywana.

Operacje morfologiczne

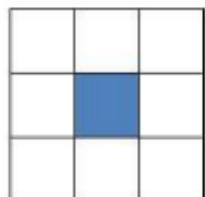


Jeśli punkt otoczenia jest wygaszony (równy wartości tła - 0) przy zapalonym (większym od tła - 1) elemencie centralnym, element pod elementem centralnym zostaje wygaszany – zrównany z tłem, a w przeciwnym wypadku zostawiamy jego poprzednią wartość

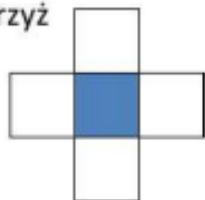
Przykładowe operacje morfologiczne

- **Erozja** – oparta o różnicę Minkowskiego;
Def.: jeżeli choć jeden piksel z sąsiedztwa określonego elementem strukturalnym ma wartość „0” to punkt centralny otrzymuje wartość „0”, w przeciwnym przypadku jego wartość nie ulega zmianie
- **dylatacja** – oparta o sumę Minkowskiego;
Def.: jeżeli choć jeden piksel z sąsiedztwa określonego elementem strukturalnym ma wartość „1” to punkt centralny otrzymuje wartość „1”, w przeciwnym przypadku przyjmuje wartość „0”
- **Otwarcie** - złożenie erozji i dylatacji,
- **Zamknięcie** – złożenie dylacji i erozji.

Kwadrat



Romb/krzyż



Wykorzystanie sąsiedztwa 8-spójnego i 4-spójnego

Podstawowe operacje morfologii matematycznej

0-zgaszony; 1-zapalony; X-o dowolnej wartości.

- Erozja

$$q(i, j) = \min_{i_n, j_m \in B(i, j)} (p(i_n, j_m))$$

- Dylatacja (dylacja) negatyw erozji

$$q(i, j) = \max_{i_n, j_m \in B(i, j)} (p(i_n, j_m))$$

$B(i,j)$ element strukturalny z punktem centralnym o współrzędnych (i,j)

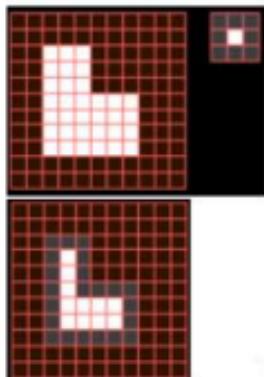
Dylatacja jest operacją **dualną** do erozji i na odwrót

1	1	1
1	1	1
1	1	1

X	X	X
X	0	X
X	X	X

Przykłady operacji erozji

Algorytm erozji opiera się na wybraniu piksela o wartości najmniejszej / minimalnej i wstawieniu go w miejsce elementu centralnego.



O małym (2x2)/dużym(6x6) kwadratowym elemencie strukturalnym

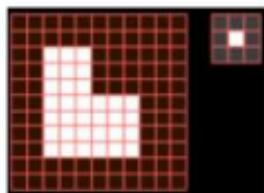


O wertykalnym/horyzontalnym elemencie strukturalnym

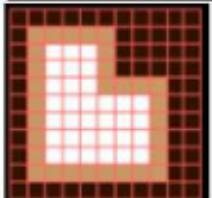
Minimum / Ściemnienie

Przykłady operacji dylacji

Algorytm dylatacji opiera się na wybraniu piksela o wartości największej/ maksymalnej i wstawieniu go w miejsce elementu centralnego.



O małym (2x2)/dużym(6x6) kwadratowym elemencie strukturalnym



O wertykalnym/horyzontalnym elemencie strukturalnym

Otwarcie

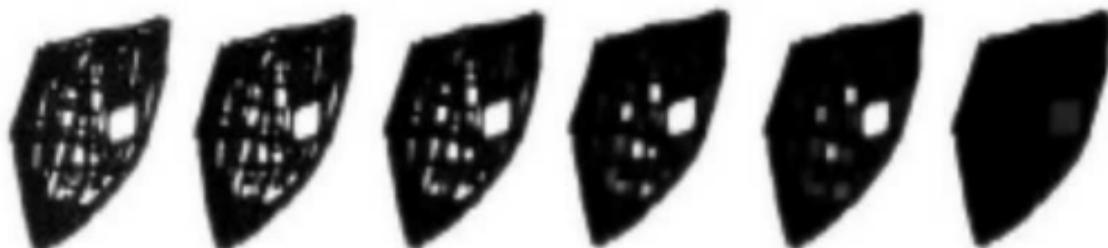
Operacja morfologiczna, która opiera się na dwóch innych operacjach: erozji i dylatacji. Otwarcie polega na wykonaniu na obrazie najpierw erozji (minimum), a następnie na tak przetworzonym obrazie należy zastosować dylatację (maksimum) na tym samym elemencie strukturyzującym.

Zamknięcie

Operacja morfologiczna, która opiera się na dwóch innych operacjach: dylatacji i erozji. Zamknięcie polega na wykonaniu na obrazie najpierw dylatacji (maksimum), a następnie na tak przetworzonym obrazie należy zastosować erozję (minimum) na tym samym elemencie strukturyzującym.

Operacje otwarcia i zamknięcia na obrazach w odcieniach szarości

Otwarcie



Zamknięcie



Element strukturalny: 3x3

5x5

9x9

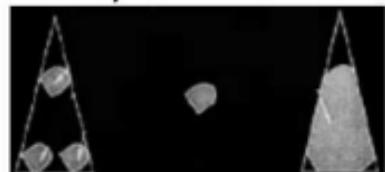
13x13

21x21

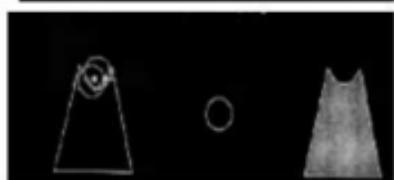
Komplementarność operacji

Otwarcie i zamknięcie – interpretacja geometryczna

Obrazy binarne



Otwarcie: toczenie elementu strukturalnego/strukturyzującego od wewnętrz



Zamknięcie: toczenie elementu strukturalnego/strukturyzującego od zewnątrz

Obrazy monochromatyczne
w odcieniach szarości



Otwarcie/Zamknięcie: toczenie kulki/piłki czyli elementu strukturalnego/strukturyzującego od dołu/góry linii profilu dla obrazów w skali szarości

Pocienianie

Zmniejszenie obiektu o piksele będące jego krawędzią.

Pogrubianie

Zwiększenie obiektu o dodatkowe piksele tła stykające się z krawędzią obiektu.

Ekstrakcja konturu

Kolejność działań: 1) operacja erozji obrazu, 2) odjęcie wyniku erozji od obrazu pierwotnego. W wyniku otrzymujemy kontur obiektu.

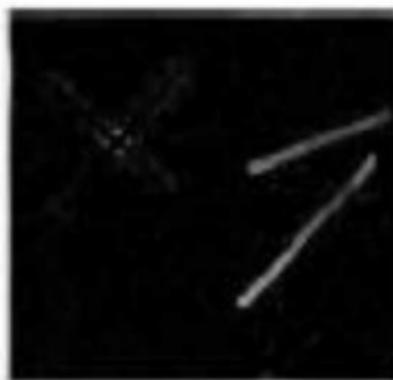
Szkieletyzacja

Operacja, która wykrywa szkielet obiektu. Przykładowy algorytm szkieletyzacji: 1) obliczyć, ile erozji można wykonać, aby obraz nie został sprowadzony do tła, 2) wykonać obliczoną ilość razy erozję i otwarcie. Wyniki kolejnych kroków erozji i otwarcia należy od siebie odjąć. Wyniki odejmowania z kolejnych kroków należy wstawić w obraz wynikowy.



obraz wynikowy

Top Hat



White Top Hat



Black Top Hat

Rekonstrukcja



Otwarcie

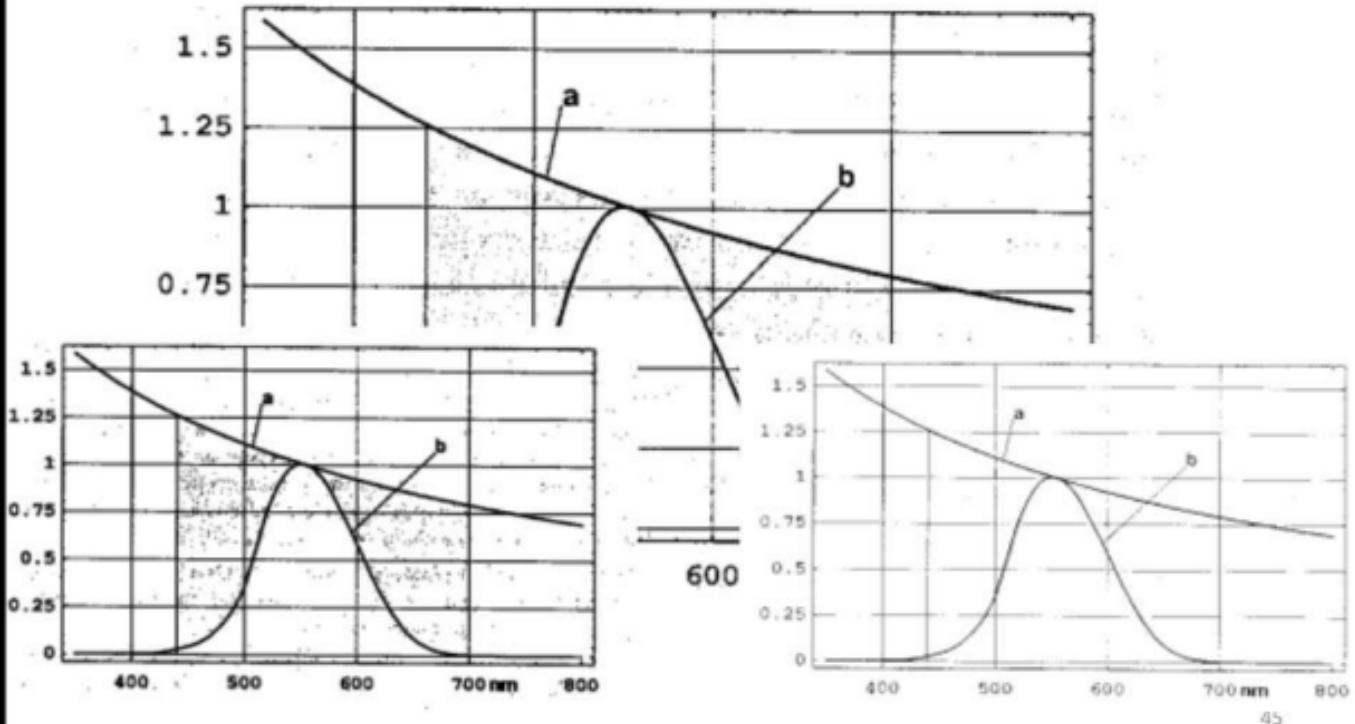


Rekonstrukcja

Pozostałe operacje morfologii matematycznej

- Gradient morfologiczny (= Otwarcie+Zamknięcie)
- Wygładzanie morfologiczne (=Dylacja–Erozja)
- Odcinanie gałęzi (artefaktów z nieregularności obiektów szkieletyzowanych)
- Detekcja centroidów (punktów centralnych obiektu)
- Dylatacja bez styków (SKIZ ang. Skeleton by influence zone)
- Erozja warunkowa
- Automediana

Przykład działania operacji morfologicznych



Transformaty Fouriera i Hough'a, kompresja obrazów

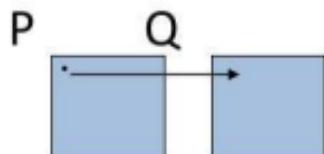
WYKŁAD 5
Dla studiów niestacjonarnych
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

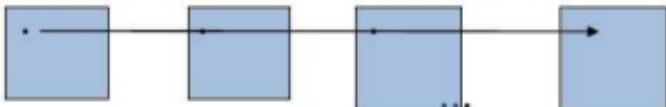
Operacje na obrazach

- Operacje punktowe (jednopunktowe):

Jednoargumentowe
 $[q(i, j)] = f[p(i, j)]$



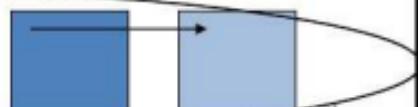
Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$


- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

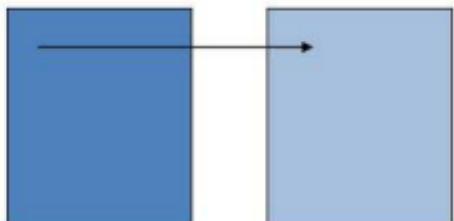

- Operacje globalne

$$[q(i, j)] = f[P]$$


Operacje globalne

Są to operacje, w których na wartość zadanego piksela obrazu wynikowego q o współrzędnych (i, j) mają wpływ wartości wszystkich pikseli obrazu pierwotnego p , czyli są to operacje kontekstowe dla których otoczeniem piksela jest cały obraz p .

$$[q(i, j)] = f[p]$$



Operacje globalne na obrazach

- Transformacje
 - Fouriera (matematyczny pryzmat)
 - kosinusowa
 - Falkowa (matematyczny mikroskop)
 - Hough'a
 - Odległościowa (morfologia matematyczna)
- Inne globalne
 - Skalowanie rozmiarów
 - Rejestracja multymodalna i multisesyjna

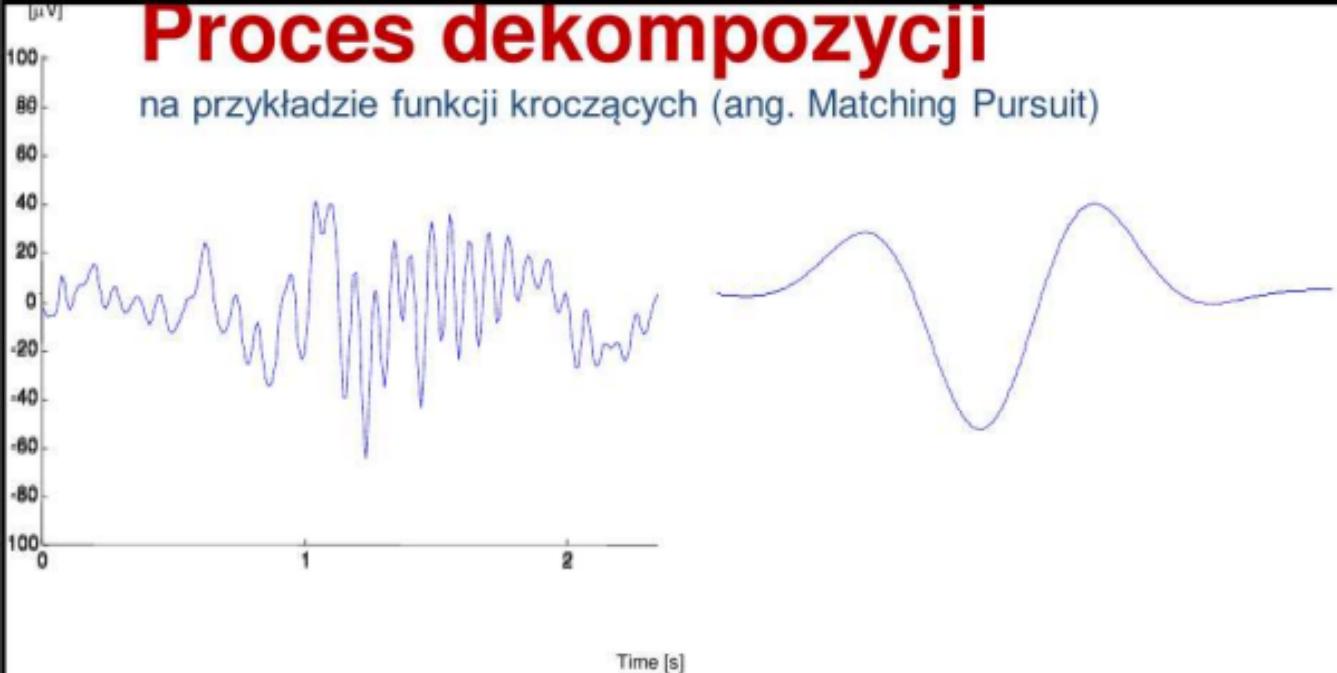
Transformata

Przekształcenie matematyczne **odwracalne i bezstratne**, przenoszące sygnał z jednej przestrzeni w inną, w której wygodniej dokonywać pewnych analiz lub procesów (np.: kompresji informacji), ponieważ w przestrzeni docelowej uwypuklane są cechy sygnału istotne z punktu widzenia celu analizy lub przetwarzania .

W przestrzeni docelowej sygnał jest **dekomponowany ze względu na zbiór funkcji bazowych tej przestrzeni**.

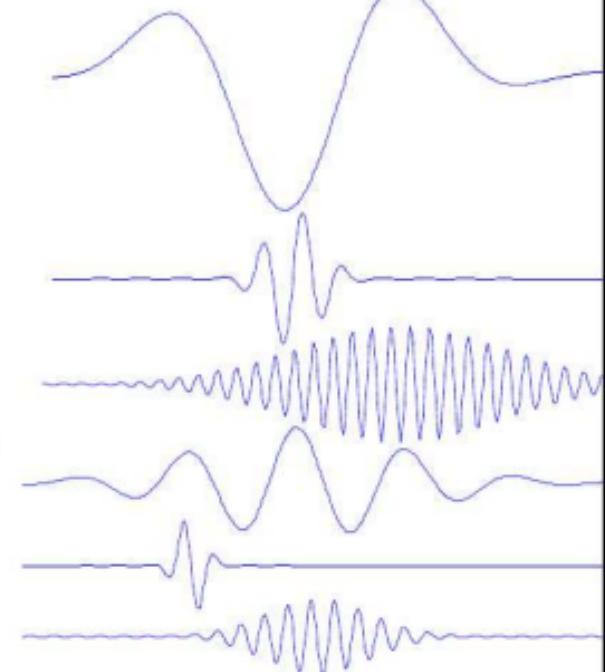
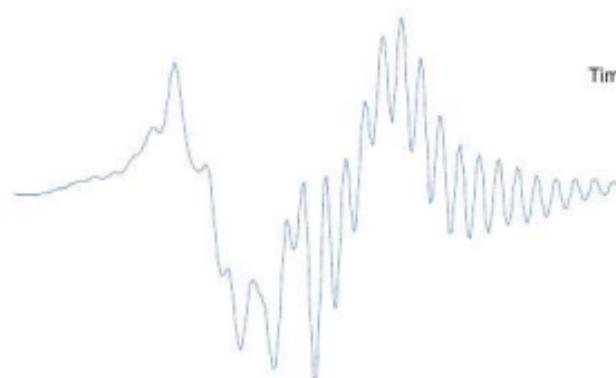
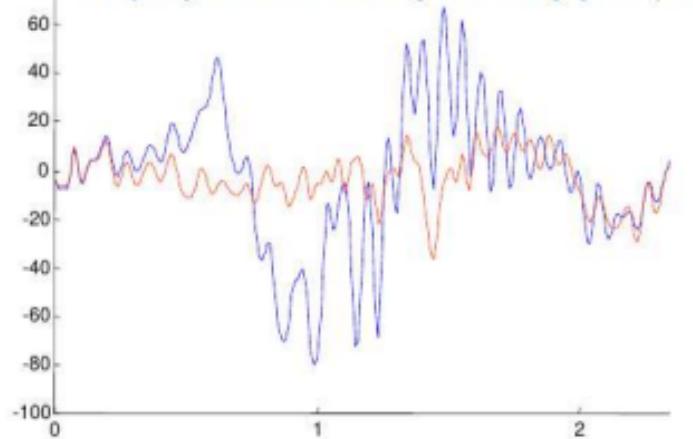
Proces dekompozycji

na przykładzie funkcji kroczących (ang. Matching Pursuit)



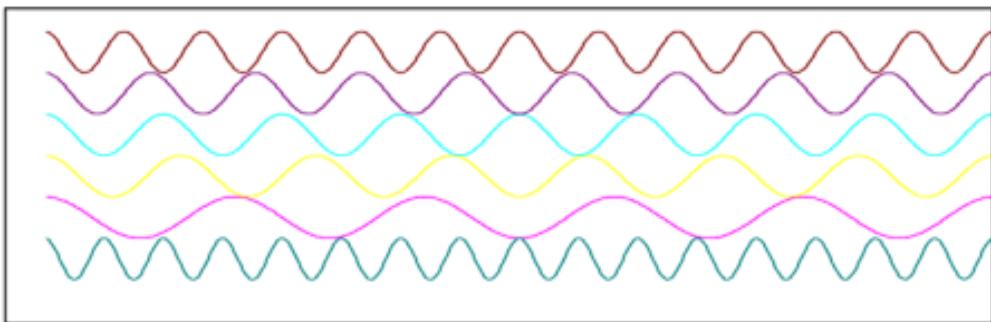
Proces dekompozycji

na przykładzie funkcji kroczących (ang. Matching Pursuit)



Transformacja Fouriera

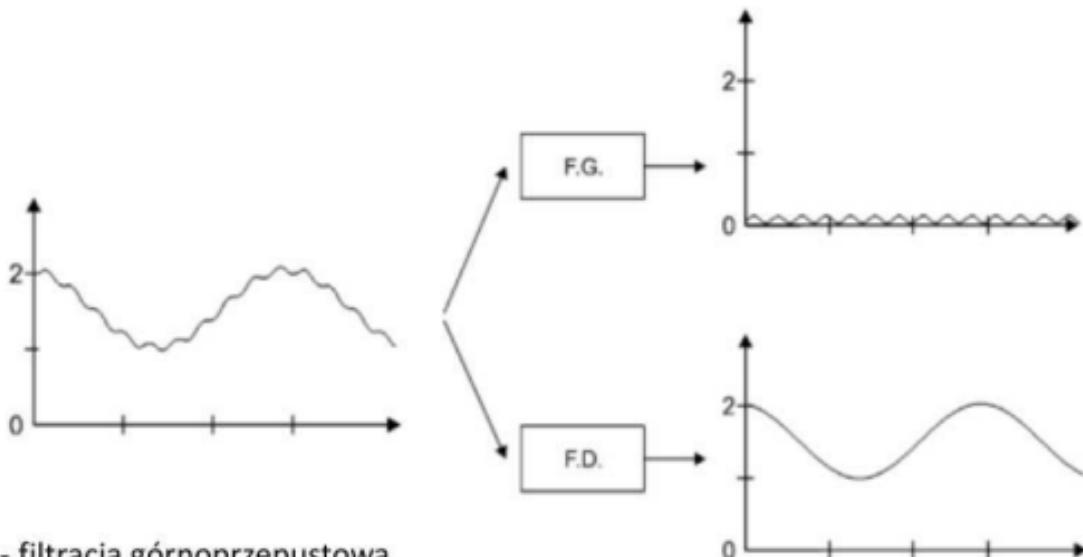
Funkcje bazowe dla przestrzeni jednowymiarowej 1D



Transformata Fouriera

Główne zastosowanie: – poprawa jakości sygnału według
przyjętego kryterium

Interpretacja częstotliwościowa filtracji



F.G. - filtracja górnoprzepustowa

F.D. - filtracja dolnoprzepustowa

Transformata Fouriera (opis matematyczny)

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp\left(-j \frac{2\pi u x}{N}\right)$$

gdzie: $u = 0, 1, \dots, N-1$, $j = \sqrt{-1}$ u - argument widma

Funkcje bazowe
- składniki sumy
w wyrażeniu
opisującym
transformatę

Złożoność obliczeniowa:

$O(N^2)$ - tzn. N^2 dodawań i mnożeń

$O(N * \log_2 N)$ - szybka transformata Fouriera
(Fast Fourier Transform (FFT))

Transformata odwrotna:

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) \exp\left(j \frac{2\pi u x}{N}\right)$$

gdzie: $x = 0, 1, \dots, N-1$; x - argument funkcji

Dwuwymiarowa transformata Fouriera

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \exp\left(-j2\pi(ux + vy)\right)$$

gdzie: $u,v = 0, 1, \dots, N-1$; u,v - argumenty widma F

Dwuwymiarowa odwrotna transformata Fouriera

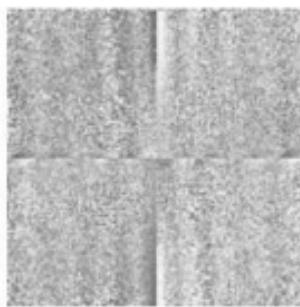
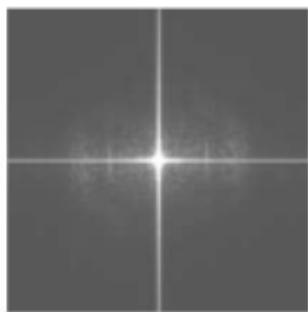
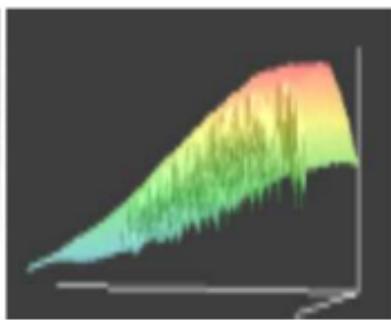
$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) \exp\left(\frac{j2\pi(ux + vy)}{N}\right)$$

gdzie: $x,y = 0, 1, \dots, N-1$ x,y - argumenty funkcji f

Duża złożoność obliczeniowa.

Transformata Fouriera

To transformacja która przenosi obraz (dyskretny sygnał dwuwymiarowy) z dziedziny oryginalnej do dziedziny częstotliwości przestrzennych.



Dziedzina przestrzenna jest taką dziedziną obrazu, w której odległości między obiektemi na obrazie odpowiadają odległościom między odpowiadającymi im obiektemi w przestrzeni (z uwzględnieniem skrótów perspektywicznych, - ang. spatial or pixels domain, którą przedstawia obraz).

W dziedzinie częstotliwości przestrzennych wartość w określonej pozycji przedstawia, jaki jest udział zmian poziomów szarości zgodnie z funkcją sinusoidalną, o kierunku i częstotliwości odpowiadającej pozycji tego punktu w układzie współrzędnych widma amplitudowego.

Wyliczanie transformaty Fouriera dla obrazów

Dzięki separowalności:

- Krok 1 po wierszach

$$P(k, m) = \frac{1}{N} \sum_{i=0}^N f(i, m) \exp\left[-j \frac{2\pi}{N}(ki)\right]$$

- Krok 2 po kolumnach

$$F(k, l) = \frac{1}{M} \sum_{m=0}^M P(k, m) \exp\left[-j \frac{2\pi}{M}(lm)\right]$$

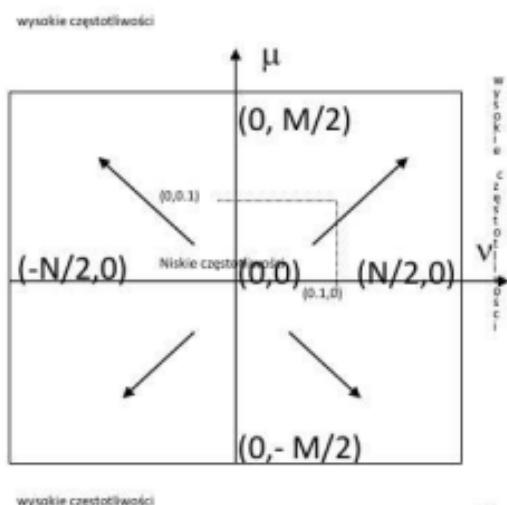
Gdzie:

$$j = \sqrt{-1} \quad \exp\left[-j \frac{2\pi}{N}(ki)\right] = \cos\left(\frac{2\pi}{N}(ki)\right) - j \sin\left(\frac{2\pi}{N}(ki)\right)$$

Algorytmy liczenia szybkiej transformaty Fouriera FFT

Wartość transformaty Fouriera jest liczbą zespoloną.
Informacja o obrazie jest zawarta w dwóch tablicach lub odpowiadających im obrazach:

- albo części rzeczywistej i części urojonej liczb zespolonych
- albo amplitudy i fazy liczb zespolonych.



Interpretacja w odniesieniu do operacji na obrazach

u, v - interpretuje się jako częstotliwości przestrzenne obrazu $f(x,y)$ odpowiednio wzdłuż osi x, y . Oznacza to, że **drobne przedmioty, detale, szum i gwałtowne** przejścia od czerni do bieli widoczne na obrazie są odpowiedzialne głównie za fragmenty funkcji $F(u,v)$ występujące dla **dużych** wartości jej argumentów. Przeciwnie, składowe $F(u,v)$ odpowiadające **małym** wartościom u,v kodują **duże obiekty lub obszary tła** obrazu.

Schemat filtracji:

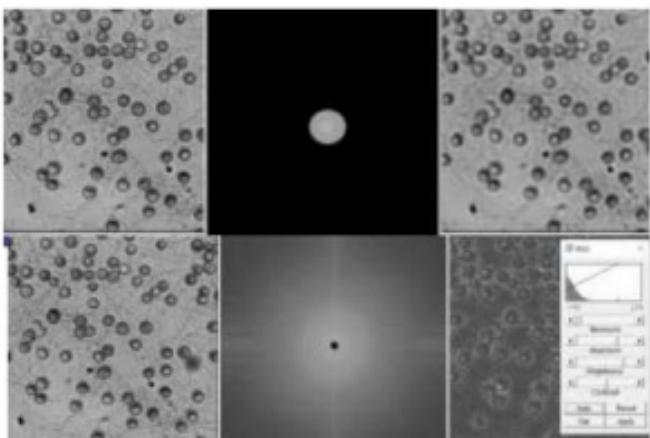
- $f(x, y) \xrightarrow{TF} F(u, v)$ funkcja filtru (usuwanie zakłóceń; podnoszenie kontrastu).
- $F' (u, v) = F(u, v) \cdot H(u, v)$ zmodyfikowana funkcja widma
- $F' (u, v) \xrightarrow{OTF} f' (x, y)$

$H(u, v)$ realizuje filtrację:

dolnoprzepustową (FD)

lub

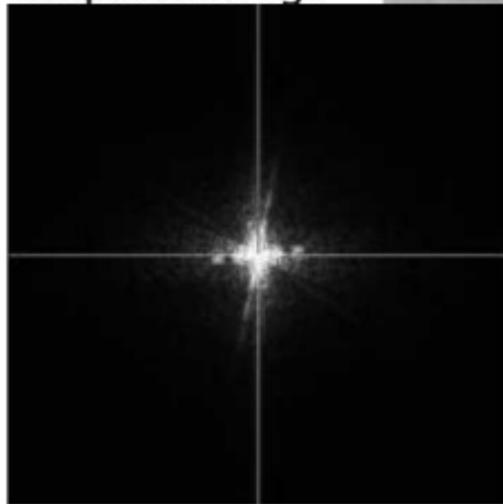
górnoprzepustową (FG).



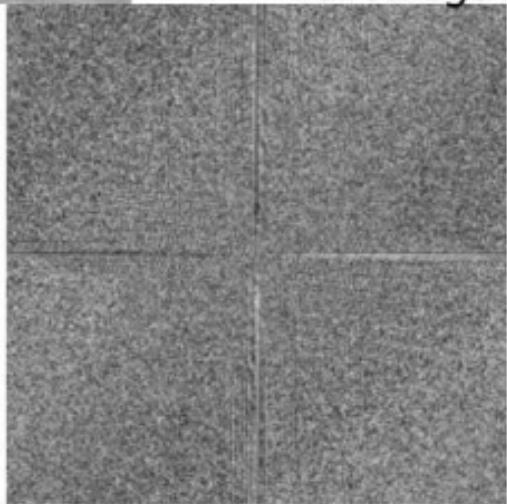
Przykład

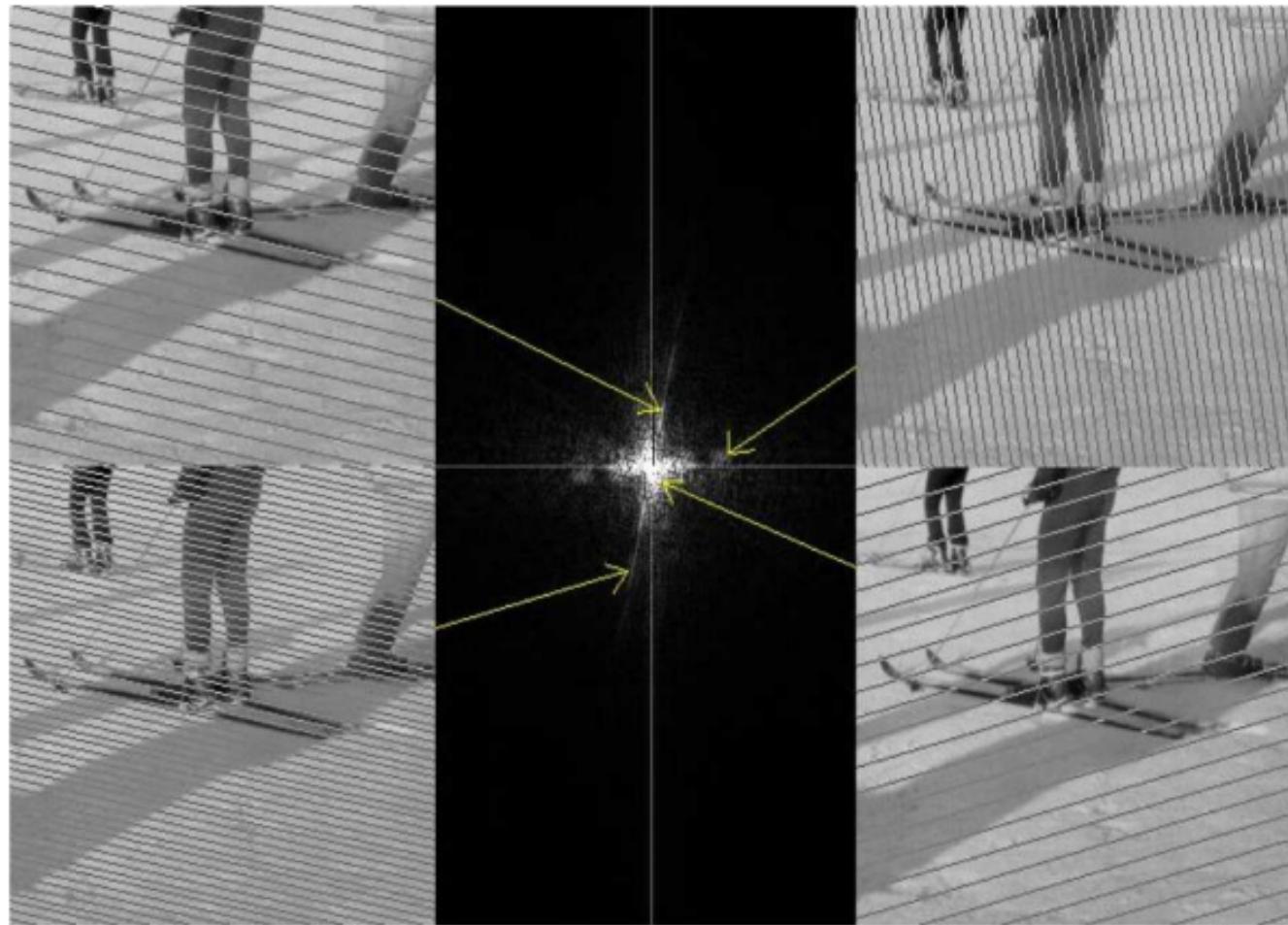


Obraz widma
amplitudowego



Obraz widma
fazowego



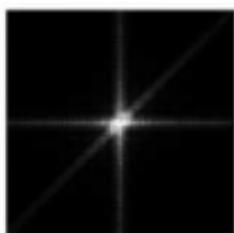
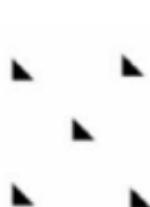


Do czego wykorzystujemy FFT w przetwarzaniu obrazów

- Do obserwacji **periodyczności** w obrazie
- Do wyznaczania **kierunku struktur** w obrazie
- Do **wytłumiania lub wzmacniania** pewnych kategorii informacji
- Do **wygładzania i wyostrzania** obrazu

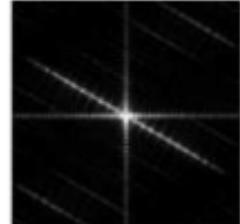
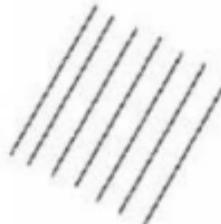
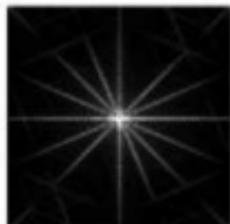
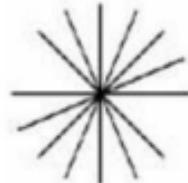
Przykład 1

Dwa zestawy periodycznie ułożonych trójkątów oraz ich widma amplitudowe



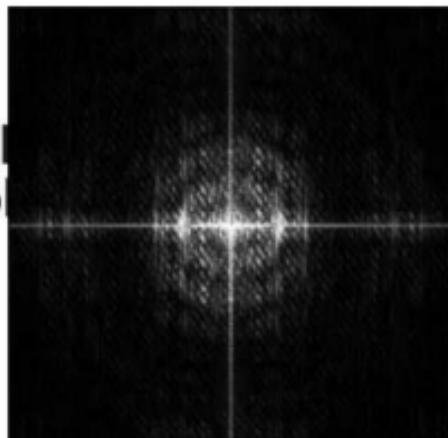
Przykład 2

Dwa zestawy linii o grubości 2 piksele i długości 150 pikseli każda oraz ich widma amplitudowe. Obraz z liniami równoległyymi ma widmo z jednym dominującym kierunkiem, prostopadłym do kierunku linii na obrazie pierwotnym.



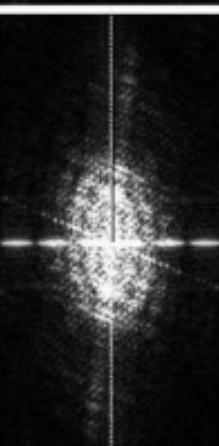
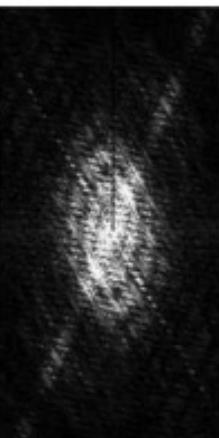
To jest tekst do
jest tekst do pro
tekst do prob.

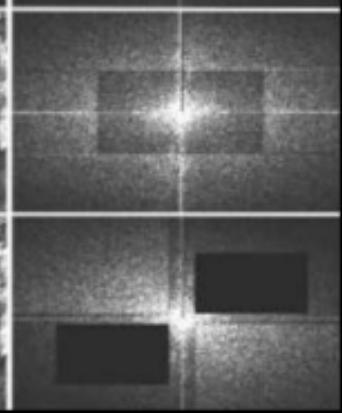
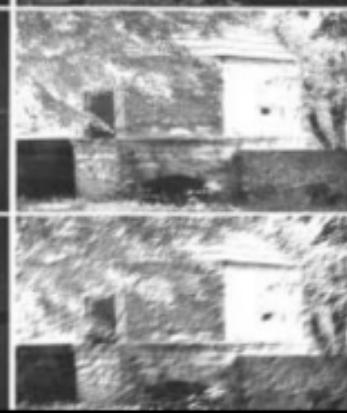
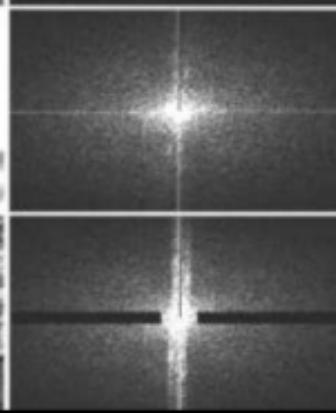
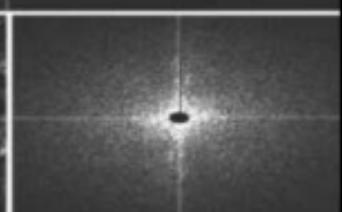
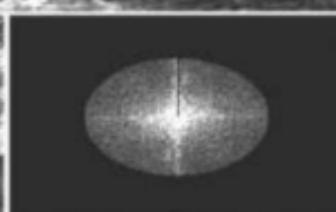
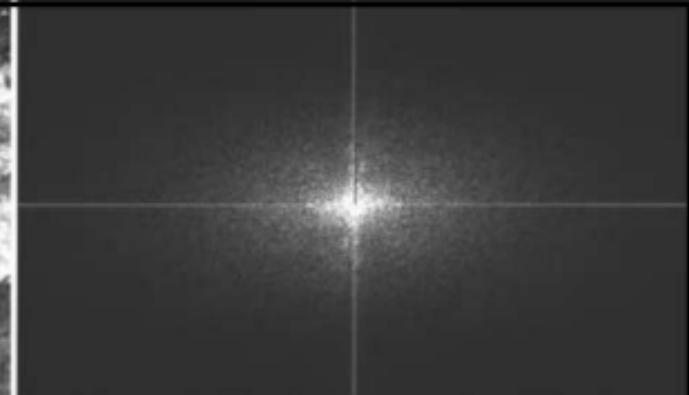
To jest tekst do pr
est tekst do prob.
t do prob.



ekst do pr
ekst do prob

est tekst do
tekst do prob
do prob



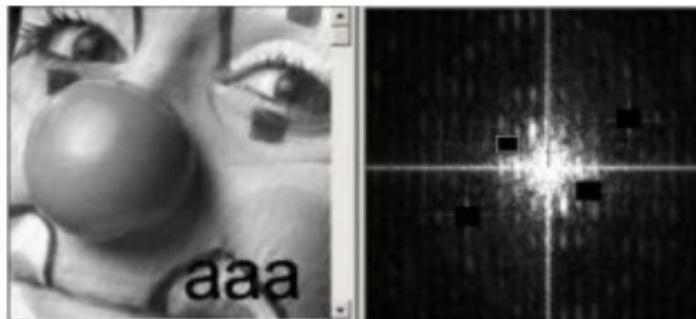


Przykład 4

Usunięcie zakłóceń w postaci regularnie powtarzających się skośnych prążków poprzez wytłumienie odpowiednich obszarów widma



Obraz i jego widmo amplitudowe po usunięciu zakłóceń



W przestrzeni docelowej sygnał jest dekomponowany ze względu na zbiór funkcji bazowych tej przestrzeni.

Do czego jeszcze może przydać się analiza częstotliwości w obrazach?

- Kompresji obrazów
- Rozszyfrowania tajemnicy uśmiechu Mony Lizy

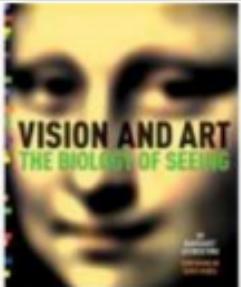
Inne cele niż te typowe dla przetwarzania obrazów

Tajemnica obrazu Mony Lizy



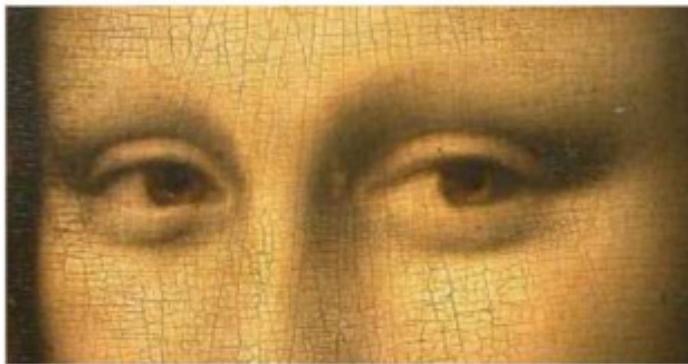
Notatki na marginesie książki malarza, zrobione w październiku 1503 roku, mówią, że to Liza del Giocondo żona florenckiego kupca Francesco del Giocondo



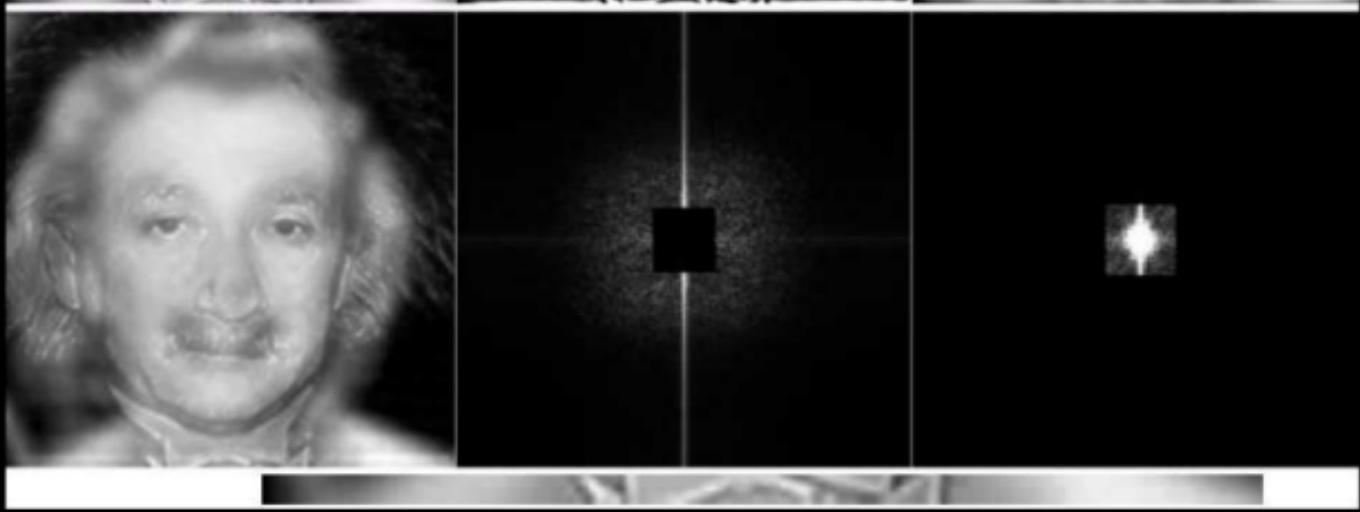


Dr. Margaret Livingstone nurobiolog wyjaśniła fenomen uśmiechu Mony Lisy tym,
jak działa ludzki wzrok.

Francuz, Pascal Cotte za pomocą urządzenia o rozdzielczości **240 mln pikseli/cal** zbadał różne warstwy obrazu w świetle białym, ultrafioletowym i podczerwonym.



Stwierdził, że kiedyś na obrazie były namalowane brwi i rzęsy kobiety, że jej dekolt ozdobiony był białą koronką, oraz że, miała **bardziej wyrazisty uśmiech**. Detale te zniknęły prawdopodobnie na skutek starzenia się obrazu.



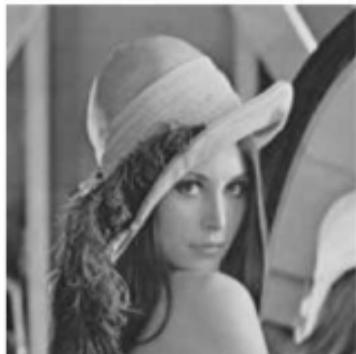
Transformata Hougha

Detekcja krawędzi z wykorzystaniem Transformaty Hougha

Detekcja linii daje rezultaty w postaci
fragmentów prostej lub krzywej, grupy
pikeseli (piksele) poszukiwanej krawędzi.

Kolejny etap:

aproksymacja linii (prostych i krzywych) na



Transformata Hough'a

Metoda detekcji krzywych analitycznych
(jak w zapisie wektorowym obrazu) nie pikseli (!))
przez transformatę wynikającą z **dualności** pomiędzy
punktami na krzywej a **parametrami** tej krzywej.
Krzywa analityczna o postaci $f(x, a) = 0$
gdzie x – punkt obrazu, a - wektor parametrów.

Zaleta:

działa dobrze nawet wówczas, gdy ciągłość krawędzi
nie jest zachowana i oprócz punktów krawędzi
występują dodatkowe elementy (najczęściej z powodu
szumów).

Założenia:

Piksele o niezerowej wartości są elementami krawędzi (obraz po krawędziowaniu).
Jeśli piksel (x, y) leży na prostej \rightarrow znaleźć zbiór wartości (ρ, ϕ) w przestrzeni parametrów tej prostej.

(x, y) - stałe, (ρ, ϕ) - zmienne \rightarrow równanie normalne prostej przedstawia relację pomiędzy krzywą w przestrzeni parametrów, a punktem w obrazie.

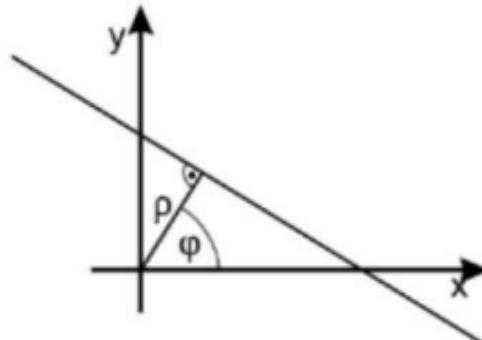
Punkt (x, y) leży na prostej \rightarrow krzywa w przestrzeni parametrów to sinusoida

TH - transformacja pomiędzy punktami obrazu a przestrzenią parametrów poszukiwanej krzywej.

Równanie *normalne* prostej:

$$x \cdot \cos\varphi + y \cdot \sin\varphi = \rho$$

Przestrzeń obrazu



Transformacja pomiędzy punktami obrazu, a przestrzenią parametrów poszukiwanej prostej

Właściwości transformacji punkt

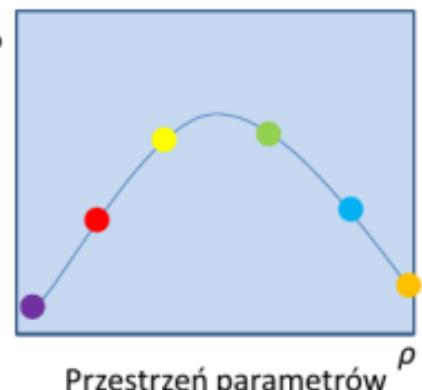
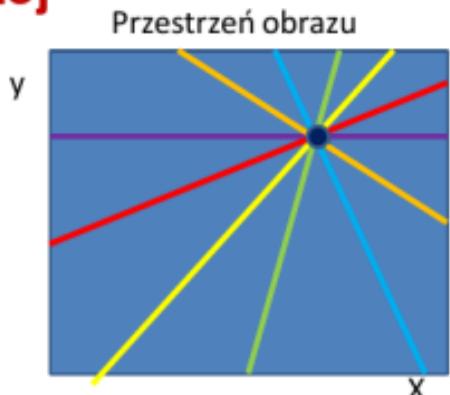
- sinusoida (przy detekcji prostej):

- punkt **obrazu** koresponduje z sinusoidą w przestrzeni parametrów,

- punkt **w przestrzeni parametrów** koresponduje z linią prostą w obrazie,

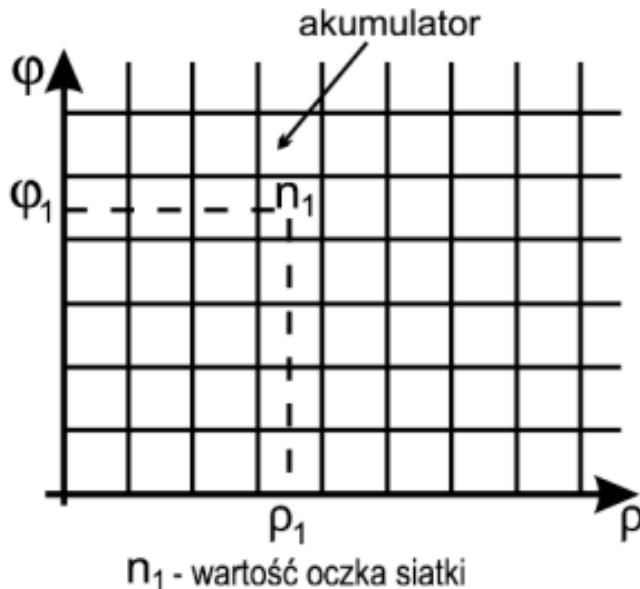
- punkty leżące na tej samej prostej w obrazie korespondują z sinusoidami przechodzącymi przez wspólny punkt w przestrzeni parametrów,

- punkty leżące na tej samej sinusoidzie w przestrzeni parametrów korespondują z liniami prostymi przechodzącymi przez ten sam punkt na obrazie.



Redukcja złożoności obliczeniowej \Rightarrow **dyskretyzacja przestrzeni parametrów ρ, ϕ** , utworzenie regularnej siatki ich wartości

Zdyskretyzowana przestrzeń Hougha dwuwymiarowa tablica akumulatorów:



1 krok:

Dla każdego punktu (x_1, y_1) obrazu oznaczonego jako należący do prostej korespondująca krzywa (sinusoida dana równaniem

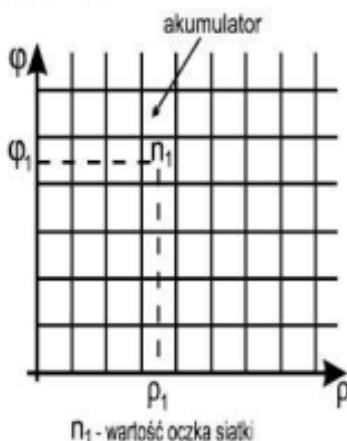
$$x_1 \cdot \cos \phi + y_1 \cdot \sin \phi = \rho$$

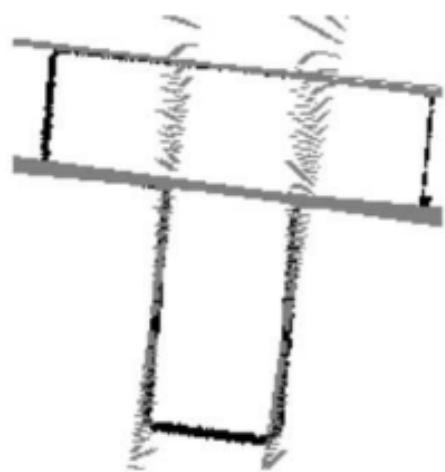
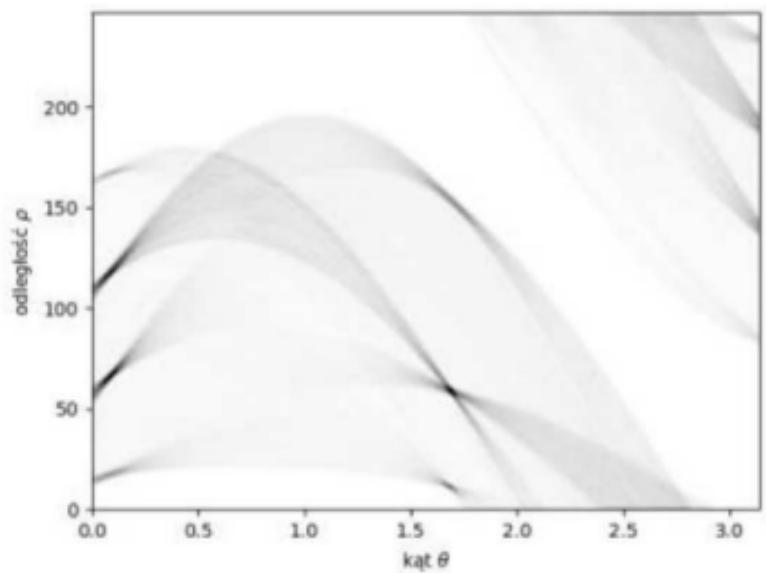
jest wprowadzana do tablicy poprzez powiększenie o 1 wartości oczek siatki w punkcie (ϕ, ρ) .

Dzięki temu każde oczko (ϕ, ρ) tablicy przechowuje liczbę równą liczbie sinusoid przechodzących przez to oczko.

2 krok:

Przeszukiwanie tablicy w celu znalezienia oczek o największych wartościach (miejsca przecięć największej liczby krzywych). Jeśli któryś element tablicy akumulatorów ma wartość k , oznacza to, że dokładnie k punktów obrazu leży wzdłuż linii, której parametrami są ϕ i ρ (pomijając błąd kwantyzacji).





Dla okręgu:

$$(x - a)^2 + (y - b)^2 = c^2$$

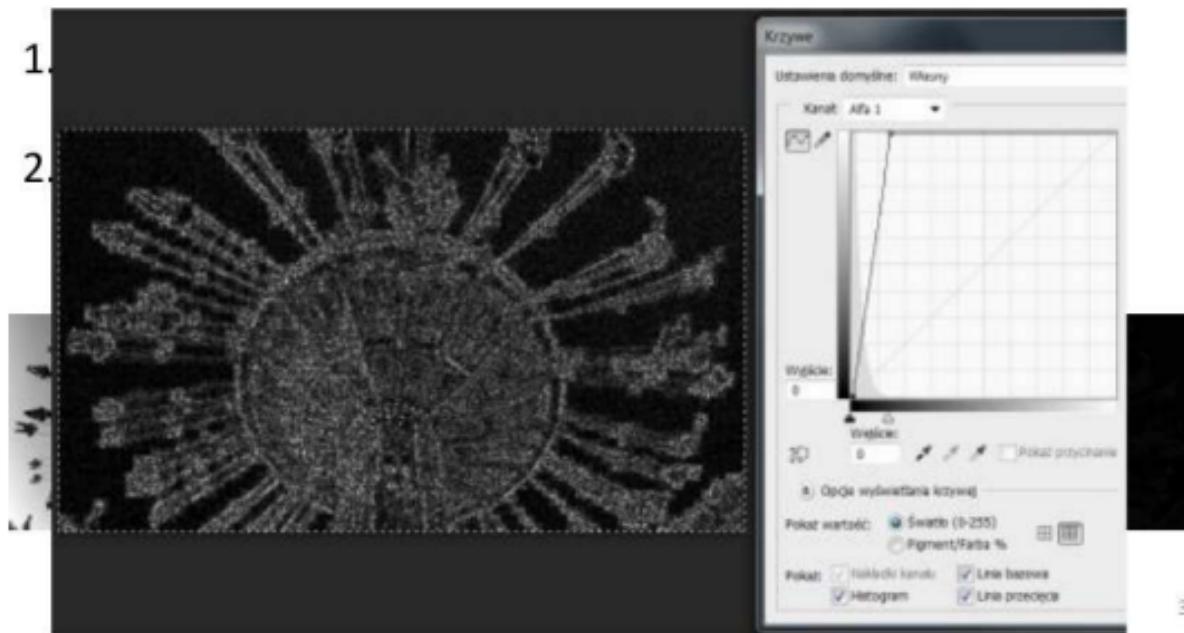
⇒ trójwymiarowa tablica akumulatorów.

współrzędne oczek: (a,b,c)

Kompresji obrazów

Jak ocenić podobieństwo obrazów?

Liczenie różnic pomiędzy obrazami – 2 sposoby



Obraz jako wektor (punkt w przestrzeni n-wymiarowej)

$$\underline{d} = [d_1, \dots, d_n]^T \quad \underline{d} \in D$$

n – liczba pikseli obrazu , np. N^2

D – przestrzeń obrazów

Wartości odpowiednich składowych wektora zależą od sposobu przeglądu obrazu.

Przykładowe sposoby przeglądu obrazu:

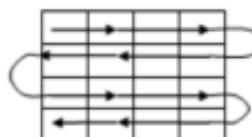
- "linia po linii" - 1-szy sposób

Bez przeplotu

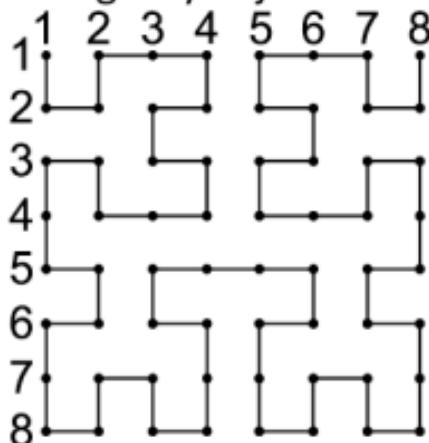


- "linia po linii - 2-gi sposób

Z przeplotem



- według krzywej Hilberta



Krzywa Hilberta rzędu 3. Obraz o rozmiarach 8x8 ($N=8$)

Metryka odwzorowanie $\rho : X \times X \rightarrow R_*$ spełniające dla wszystkich wektorów $\underline{x}^\mu \in X$ (μ założenia) (warunki):

$$\rho(\underline{x}^\mu, \underline{x}^\nu) = 0 \Leftrightarrow \underline{x}^\mu \equiv \underline{x}^\nu \quad - \text{tożsamość}$$

$$\rho(\underline{x}^\mu, \underline{x}^\nu) = \rho(\underline{x}^\nu, \underline{x}^\mu) \quad - \text{symetria},$$

$$\rho(\underline{x}^\mu, \underline{x}^\nu) \leq \rho(\underline{x}^\mu, \underline{x}^\eta) + \rho(\underline{x}^\eta, \underline{x}^\nu) \quad - \text{warunek trójkąta.}$$

Gdzie:

X - przestrzeń wektorów

R_* zbiór liczb nieujemnych

Zastosowanie praktyczne

obliczanie **różnic** pomiędzy poszczególnymi obrazami

Metryka Euklidesowa

$$\rho_1(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{\nu=1}^n (x_\nu^\mu - x_\nu^\eta)^2}$$

Wady i zalety metryki Euklidesowej

- odpowiada obiegowej definicji odległości
- ignorowanie składowych o bardzo małych wymiarach, długie czasy obliczeń (pierwiastkowanie, podnoszenie do kwadratu)

Uogólniona metryka Euklidesowa

$$\rho_2(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{\nu=1}^n [\lambda_\nu (x_\nu^\mu - x_\nu^\eta)]^2}$$

λ_ν - mnożniki normalizujące

Metryka uliczna (Manhattan, city block distance):

$$\rho_3(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n |x_v^\mu - x_v^\eta|$$

Uogólniona metryka uliczna

$$\rho_4(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n \lambda_v |x_v^\mu - x_v^\eta|$$

Metryka Czebyszewa (maksymalna)

$$\rho_5(\underline{x}^\mu, \underline{x}^\eta) = \max_{1 \leq v \leq n} |x_v^\mu - x_v^\eta|$$

Oznaczenia $\underline{x}^\mu = [x_1^\mu, \dots, x_v^\mu, \dots, x_n^\mu]^T$ $\underline{x}^\eta = [x_1^\eta, \dots, x_v^\eta, \dots, x_n^\eta]^T$

Obliczyć różnicę (stosując metrykę euklidesową) pomiędzy dwoma obrazami trzypikslowymi $\underline{d}^1 = [d_1^1, d_2^1, d_3^1]^T = [2, 3, 5]$, $\underline{d}^2 = [d_1^2, d_2^2, d_3^2]^T = [1, 4, 5]$

$$\rho(\underline{d}^1, \underline{d}^2) = \sqrt{\sum_{i=1}^3 (d_i^1 - d_i^2)^2} = \sqrt{(2-1)^2 + (3-4)^2 + (5-5)^2} = \sqrt{2}$$

KOMPRESJA OBRAZÓW

Cele kompresji

- archiwizacja,
- przesyłanie.

Stopień kompresji obrazu

$$SK = \frac{KP}{KW}$$

SK - stopień kompresji obrazu.

KP [abajt] - obszar pamięci zajmowany przez **kod pierwotny** obrazu.

Kod pierwotny - reprezentacja rastrowa lub wektorowa.

KW [abajt] - obszar pamięci zajmowany przez **kod wynikowy** obrazu.

Rodzaje kompresji:

- kompresja bezstratna (*lossless compression*)
- kompresja stratna (*lossy compression*)

Rodzaje kompresji (kodowania)

- **Bezstratna** (ang. lossless coding) odwracalna (ang. reversible) redukcja redundancji statystycznej (ang. statistical redundancy) w czasie i przestrzeni
- **Stratna** (ang. lossy coding) nieodwracalna (ang. irreversible) redukcja redundancji subiektywnej (ang. subjective redundancy), dotycząca nieistotnej informacji (ang. irrelevancy), z lub bez uwzględnienia charakterystyki słuchu i wzroku człowieka
- „Prawie” bezstratna lub **percepcyjnie bezstratna**, tzn. stratna, ale poziom zmian względem oryginału nie jest odczuwany przez człowieka

Kompresja bezstratna i stratna

lossless compression

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) = 0$$

lossy compression

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) \neq 0$$

\underline{x}^{μ} - wektor reprezentujący obraz pierwotny

\underline{x}^{ν} - wektor reprezentujący obraz odtworzony (zrekonstruowany)

ρ - wartość metryki

Wybrane metody kompresji bezstratnej obrazów rastrowych:

- Kompresja obrazów z obszarami o jednolitej jasności
- Kompresja obrazów z obszarami o niejednolitej jasności

Wybrane metody kompresji stratnej obrazów rastrowych

- *kodowanie różnic,*
- *kodowanie blokowe*

Kompresja stratna - uzyskiwany jest większy stopień kompresji (SK) niż przy stosowaniu kompresji bezstratnej.

Kompresja obrazów z obszarami o jednolitej jasności

Kodowanie ciągów identycznych symboli (RLE)

ciąg identycznych symboli - para zawiera 1 symbol i liczbę jego powtórzeń

Przykład:

Kodowanie dziesiętne:

7,7,8,9,10,10,10,10,9,9,9,8,7,7,7

7(2), 8(1), 9(1), 10(4), 9(3), 8(1), 7(3).

Kodowanie binarne: 111, 111

Długość kodu zależy:

- **od rozmiaru obrazu ($N_1 \times N_2$),**
- **od liczby poziomów jasności obrazu (M)**

Uwaga: dla obrazów o dużych obszarach o jednolitej jasności przegląd według krzywej Hilberta daje z reguły dłuższe ciągi identycznych symboli niż w przypadku przeglądu "linia po linii"

Praktyczna realizacja

metody kodowania ciągów identycznych symboli:

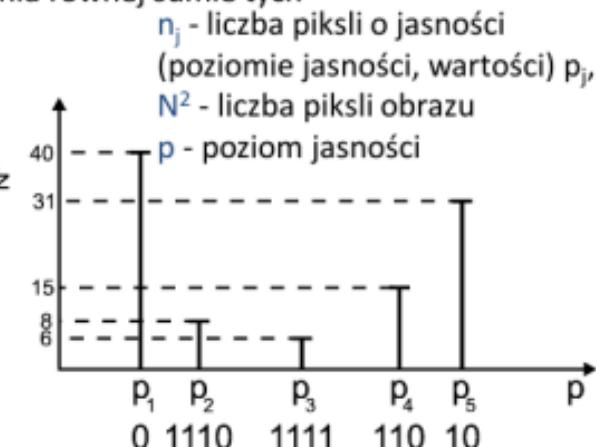
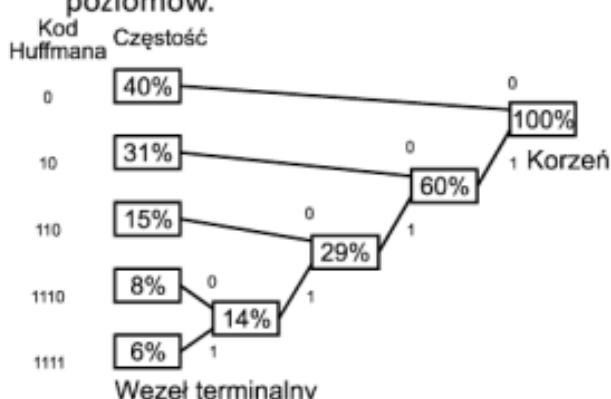
- formaty obrazowe: *.PCX, *.PIC

Kompresja obrazów z obszarami o niejednolitej jasności

Kod Huffmana - kod o zmiennej długości słowa.

Sposób postępowania:

- Przypisanie każdemu poziomowi jasności częstości występowania pikseli o tym poziomie jasności (na podstawie histogramu).
- Wyszukanie 2 poziomów o najmniejszej częstości występowania i połączenie w jeden o częstości występowania równej sumie tych poziomów.



Schemat łączenia - drzewo Huffmana.

Praktyczna realizacja kodu Huffmana - format obrazowy *.TGA

Metody słownikowe

Kodowanie *ciągów symboli* (pikseli) za pomocą odwołań do *słownika* zawierającego takie ciągi.

Stopień kompresji (SK) rośnie, gdy **długość ciągu pikseli** możliwych do zastąpienia *indeksem* do słownika rośnie.

Podział metod słownikowych:

- *statyczne* (słownik nie zmienia się w trakcie kodowania),
- *adaptacyjne* (słownik zmienia się w trakcie kodowania)

Kodowanie różnic

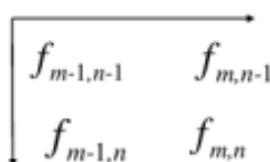
$$\epsilon_{m,n} = f_{m,n} - f_{m,n-1}$$

gdzie:

$f_{m,n}$ - poziom jasności piksela o współrzędnych m,n ,

$f_{m,n-1}$ - poziom jasności piksela (**kolejnego**) o współrzędnych $m,n-1$.

Położenie pikseli



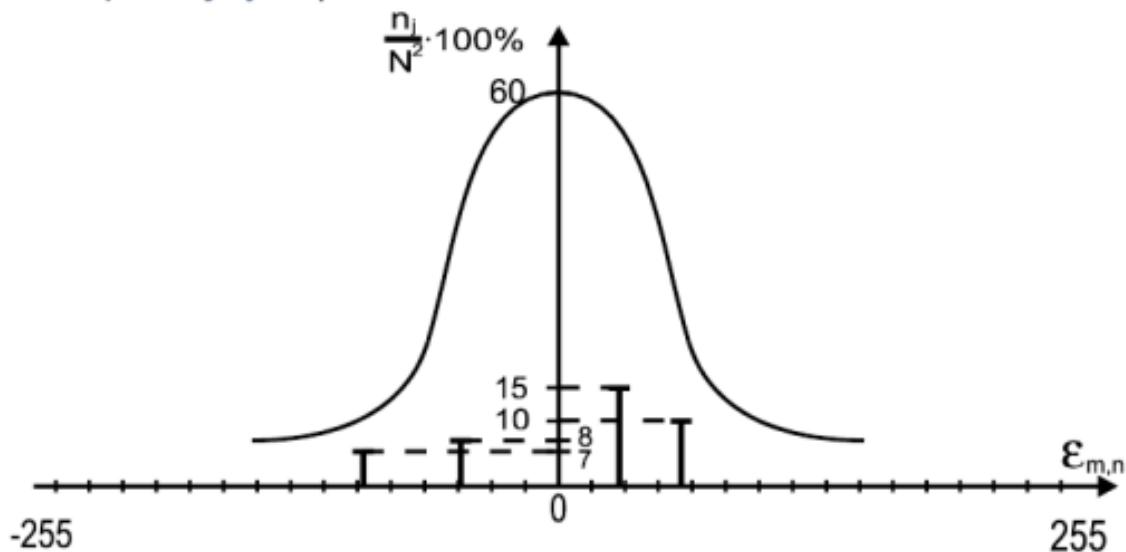
Przykład:

Obraz pierwotny : $L = 255$ ($M = 256$)

- kodowanie wartości pikseli: $I = 0,1,2,3,\dots,255$; zajętość pamięci: 8 bitów/piksel,
- kodowanie różnic pomiędzy wartościami kolejnych pikseli:

$$\epsilon_{m,n} = -255, -254, \dots, 0, \dots, 254, 255 \quad \text{zajętość pamięci: 9 bitów / piksel.}$$

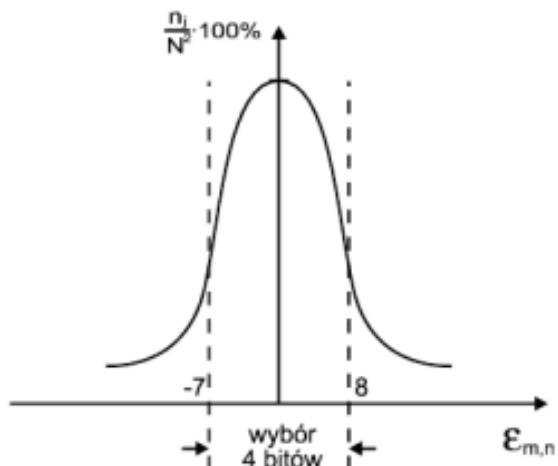
Histogram różnic $\varepsilon_{m,n}$ pomiędzy wartościami (poziomami jasności) kolejnych pikseli



Histogram różnic ma charakter *krzywej Gaussa* z maksimum w okolicy zera.

Kompresja wynika z zakodowanie najczęściej występujących różnic

Histogram różnic z zaznaczonym obszarem najczęściej występujących różnic



Kryteria wyboru obszaru:

- wymagania na wielkość stopnia kompresji (SK),
- wymagania na dokładność rekonstrukcji (określoną wielkością „ p ”)

$$KP = 9, KW = 4$$

$$SK = KP/KW = 9/4 = 2,25$$

Obraz zakodowany $\epsilon_{m,n} = -7, -6, \dots, 7, 8$ - 16 symboli zamiast 512

Niepożądany efekt: zależnie od rodzaju obrazu - większe lub mniejsze rozmywanie (*blurring*) ostrych krawędzi.

Kodowanie blokowe

Podział obrazu na jednakowe bloki, najczęściej 4×4 piksele. Obliczenie dla każdego bloku średniej arytmetycznej jasności. Podział pikseli na dwie grupy:

- a) o jasności większej lub równej jasności średniej,
- b) mniejszej niż jasność średnia.

Obliczenie nowej jasności średniej dla każdej z grup (**wartość góra** dla (a) i **wartość dolna** dla (b)). Przypisanie wszystkim pikselom danej grupy obliczonej jasności średniej (górną lub dolną), stąd blok zostaje zakodowany jako *mapa bitowa* określająca podział na grupy, plus dwie wartości jasności.

The diagram illustrates the process of block-based coding. On the left, a 4x4 grid of pixel values is shown. An arrow points from this grid to a 4x4 binary matrix on the right, which represents the bit map for the upper group. The pixel values are: Row 1: 13, 11, 10, 13; Row 2: 8, 9, 12, 15; Row 3: 6, 8, 11, 14; Row 4: 4, 7, 9, 12. The binary matrix has columns separated by vertical lines and rows by horizontal lines. The first column contains 1, 0, 0, 0. The second column contains 1, 0, 0, 0. The third column contains 0, 1, 1, 0. The fourth column contains 1, 1, 1, 1.

13	11	10	13
8	9	12	15
6	8	11	14
4	7	9	12

1	1	0	1
0	0	1	1
0	0	1	1
0	0	0	1

Średnia=10,125

Wartość góra= 12,625≈13
Wartość dolna = 7,625≈ 8

Kompresje obrazów oparte na transformatach

- Wybieram takie transformaty, które dają w dziedzinie transformacji takie upakowanie informacji, aby większość współczynników była zerowa (bliska零)
- Po kwantyzacji niezerowych współczynników dają minimalne zniekształcenia obrazu
- Prowadzą do dekorelacji sygnału
- Są ortogonalne, bo to zapewnia pełną transmisję informacji
- Są separowalne, bo to zapewnia szybki algorytm dekompozycji

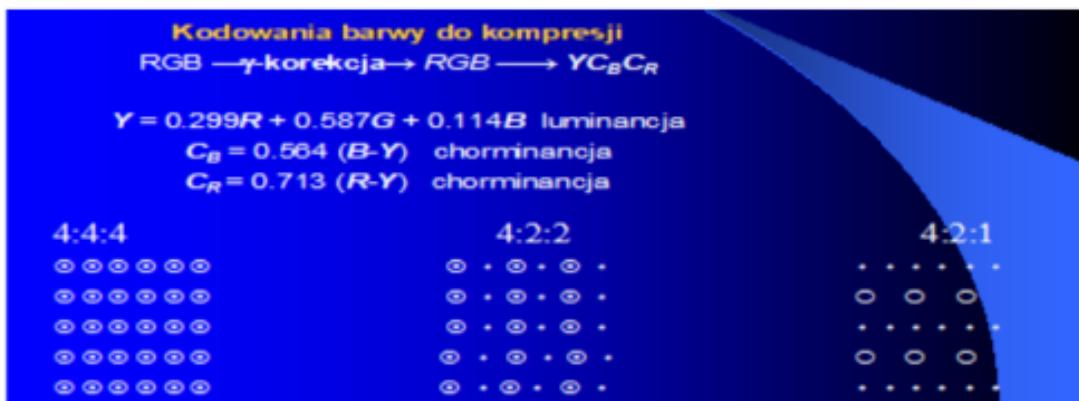
Kompresja stratna oparta na transformacjach

- Transformacja kosinusowa w kompresji JPEG
- Transformacja falkowa w kompresji JPEG 2000

Kolejność działań (i):

Konwersja obrazu kanałów (RGB) na jasność (luminancję) i 2 kanały barwy (chrominancje).

Odrzucenie części pikseli kanałów barwy, podział kanałów na bloki 8×8 pikseli.



Kolejność działań (ii)

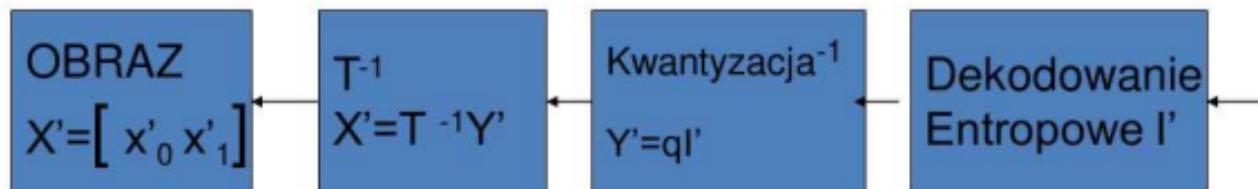
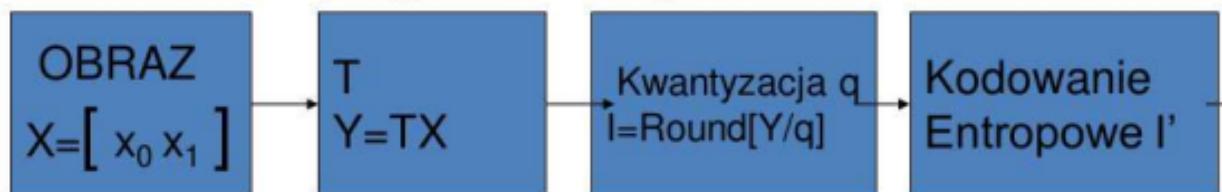
Na blokach wykonywana jest dyskretna transformata kosinusowa (DCT).

Zastąpienie średnich wartości bloków przez różnice wobec wartości poprzedniej (DPCM). Poprawia to w pewnym stopniu współczynnik kompresji.

Kwantyzacja, czyli zastąpienie danych zmiennoprzecinkowych przez liczby całkowite.

Schemat kompresji opartej na transformacjach

Kodowanie (ang. encoder)



Dekodowanie (ang. decoder)

Kompresja JPEG

- Odpowiada kompresji stratnej
- Oparty na dyskretnej transformacji kosinusowej (DCT)

$$\exp\left[-j \frac{2\pi}{N} (ki)\right] = \cos\left(\frac{2\pi}{N} (ki)\right) - j \sin\left(\frac{2\pi}{N} (ki)\right)$$

$$F(u, v) = \frac{4c(u, v)}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)\pi u}{2N} \cos \frac{(2y+1)\pi v}{2N}$$

$$c(u, v) = \begin{cases} \frac{1}{2} & u = 0 \text{ lub } v = 0 \\ 1 & \text{pozostało} \end{cases}$$

- Wraz ze wzrostem stopnia kompresji (SK) coraz bardziej widoczny jest podział na bloki w obrazie zrekonstruowanym

Przykład



Obraz	Jakość	Rozmiar (abajty)	Współczynnik kompresji	Komentarz
	Pełna jakość (Q = 100)	83,261	2,6:1	Niezauważalnie małe artefakty



	Niska jakość (Q = 10)	4,787	46:1	Duże straty rozciągłości tonalnej; artefakty na krawędziach obiektów
--	-----------------------	-------	------	--

Innowacyjność algorytmu JPEG: możliwość kontroli stopnia kompresji w jej trakcie

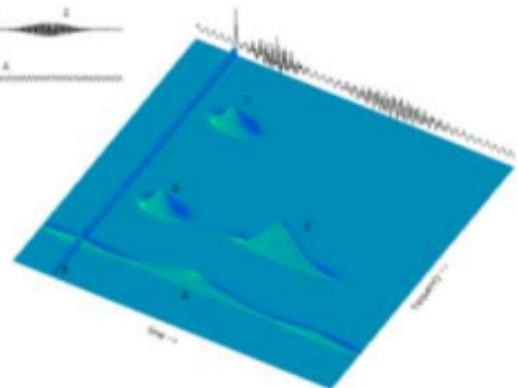
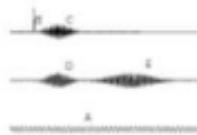
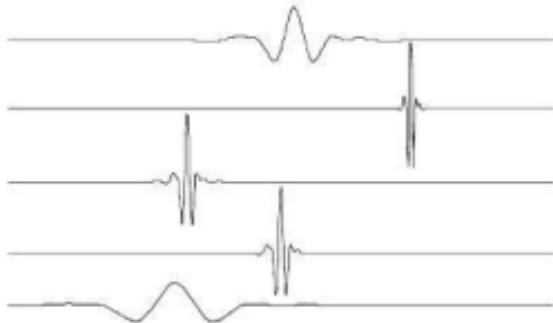
Kompresja JPEG200

- Koduje różne typy obrazów i filmów: binarne, szare, kolorowe, wielospektralne o różnych właściwościach, transmisja w czasie rzeczywistym, archiwizacja biblioteki obrazów.
- Umożliwia progresywne dekodowanie: pozwala widzowi zobaczyć wersję danego obrazu o niższej jakości, podczas gdy cały plik nie został w pełni odebrany lub pobrany. Jakość obrazu stopniowo się poprawia, gdy więcej bitów danych jest pobieranych ze źródła.
- Zapewnia zachowanie przejrzystości obrazów.
- Daje niezwykle wysoki poziom skalowalności.
- Obrazy mogą być kodowane w dowolnej jakości od bardzo skompresowanych do kompresji bezstratnej.
- Firma udostępnia oprogramowanie open-source tzn. OpenJPEG kod napisany w C do wykorzystania przez programistów.
- Stosuje transformację falkową

Co to są falki?

Komplet funkcji umożliwiających hierarchiczną wielorodzicielną reprezentację aproksymowanej funkcji, składający się z :

- Funkcji skalującej $\Phi(x)$
- Rodziny falek generowanych na podstawie falki podstawowej



Falki i transformacja falkowa w przetwarzaniu obrazów

- Kompresja obrazów i sekwencji wideo
- Odszumianie obrazów
- Analiza tekstury
- Segmentacja
- Rozpoznawanie np.: ręcznego pisma, biologicznego wieku dzieci, charakterystycznych cech biologicznego sygnału (EEG)
- Rekonstrukcja obrazów biomedycznych

Metody oceny jakości kodowania

- Subiektywne – notowanie wrażeń widzów lub słuchaczy
- Obiektywne – porównanie ilościowe, na podstawie matematycznie wyznaczonych parametrów, opisujących różnicę między sygnałem oryginalnym i po dekompresji
- Metody oparte na modelach percepcji informacji przez człowieka:
 - perceptualny model, zapewniający wierność percepcji słuchowej
 - modelowanie układu percepcji wzrokowej człowieka
(HVS ang. Human Visual System)

Subiektywna ocena jakości kodowania

- MOS (ang. Mean Opinion Score) – uśredniona ocena słuchaczy lub widzów, nie będących ekspertami

Norma dla sygnałów telewizyjnych ITU-R, zalecenie BT.500:

15 obserwatorów ogląda nie dłużej niż 30 min szereg 10-15 sekundowych prezentacji obrazów statycznych i sekwencji wizyjnych, będących międzynarodowymi standardami.

Zastosowane wymuszenia: pojedyncze (bez obrazu odniesienia) lub podwójne (z obrazem odniesienia), oddzielane szarym obrazem.

Bezpośrednio po prezentacji oceniają w ciągłej lub skokowej skali pięciostopniowej dwie cechy: jakość obrazu i poziom zakłóceń.

Inne protokoły miar: absolutnych i porównawczych (szeregujących)

Wady: zależność oceny od kontekstu i kolejności

Zalety: lepsze niż obiektywne

Obiektywna ocena jakości kodowania

- Średnia różnica
- Maksymalna różnica
- Błąd średniokwadratowy (ang. Mean Square Error)
- Szczytowy błąd średniokwadratowy (ang. Peak Mean Square Error)
- Znormalizowany błąd średniokwadratowy (ang. Normalized Mean Square Error)
- Laplasjanowy błąd średniokwadratowy (ang. Laplacian Mean Square Error)
- Jakość korelacji (ang. Correlation Quality)
- Znormalizowana korelacja skośna (ang. Normalized Cross-Correlation)
- Zawartość strukturalna (ang. Structural Content)
- Wierność obrazu (ang. Image Fidelity)
- Ważone odległości obrazów według różnych norm, np. normy Minkowskiego
- Inne, oparte na graficznych miarach jakości obrazów, liczone np. na podstawie histogramu lub wykresu Hosaka, obrazującego degeneracje jakości

Metody kompresji obrazów

Modelowanie koloru, segmentacja obrazu

WYKŁAD 6
Dla studiów niestacjonarnych
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

KOMPRESJA OBRAZÓW

Cele kompresji

- archiwizacja,
- przesyłanie.

Stopień kompresji obrazu

$$SK = \frac{KP}{KW}$$

SK - stopień kompresji obrazu.

KP [abajt] - obszar pamięci zajmowany przez **kod pierwotny** obrazu.

Kod pierwotny - reprezentacja rastrowa lub wektorowa.

KW [abajt] - obszar pamięci zajmowany przez **kod wynikowy** obrazu.

Rodzaje kompresji:

- kompresja bezstratna (*lossless compression*)
- kompresja stratna (*lossy compression*)

Rodzaje kompresji (kodowania)

- **Bezstratna** (ang. lossless coding) odwracalna (ang. reversible) redukcja redundancji statystycznej (ang. statistical redundancy) w czasie i przestrzeni
- **Stratna** (ang. lossy coding) nieodwracalna (ang. irreversible) redukcja redundancji subiektywnej (ang. subjective redundancy), dotycząca nieistotnej informacji (ang. irrelevancy), z lub bez uwzględnienia charakterystyki słuchu i wzroku człowieka
- „Prawie” bezstratna lub **percepcyjnie bezstratna**, tzn. stratna, ale poziom zmian względem oryginału nie jest odczuwany przez człowieka

Kompresja bezstratna i stratna

lossless compression

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) = 0$$

lossy compression

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) \neq 0$$

\underline{x}^{μ} - wektor reprezentujący obraz pierwotny

\underline{x}^{ν} - wektor reprezentujący obraz odtworzony (zrekonstruowany)

ρ - wartość metryki

Wybrane metody kompresji bezstratnej obrazów rastrowych:

- Kompresja obrazów z obszarami o jednolitej jasności
- Kompresja obrazów z obszarami o niejednolitej jasności

Wybrane metody kompresji stratnej obrazów rastrowych

- *kodowanie różnic,*
- *kodowanie blokowe*

Kompresja stratna - uzyskiwany jest większy stopień kompresji (SK) niż przy stosowaniu kompresji bezstratnej.

Kompresja obrazów z obszarami o jednolitej jasności

Kodowanie ciągów identycznych symboli (RLE)

ciąg identycznych symboli - para zawiera 1 symbol i liczbę jego powtórzeń

Przykład:

Kodowanie dziesiętne:

7,7,8,9,10,10,10,10,9,9,9,8,7,7,7

7(2), 8(1), 9(1), 10(4), 9(3), 8(1), 7(3).

Kodowanie binarne: 111, 111

Długość kodu zależy:

- **od rozmiaru obrazu ($N_1 \times N_2$),**
- **od liczby poziomów jasności obrazu (M)**

Uwaga: dla obrazów o dużych obszarach o jednolitej jasności przegląd według krzywej Hilberta daje z reguły dłuższe ciągi identycznych symboli niż w przypadku przeglądu "linia po linii"

Praktyczna realizacja

metody kodowania ciągów identycznych symboli:

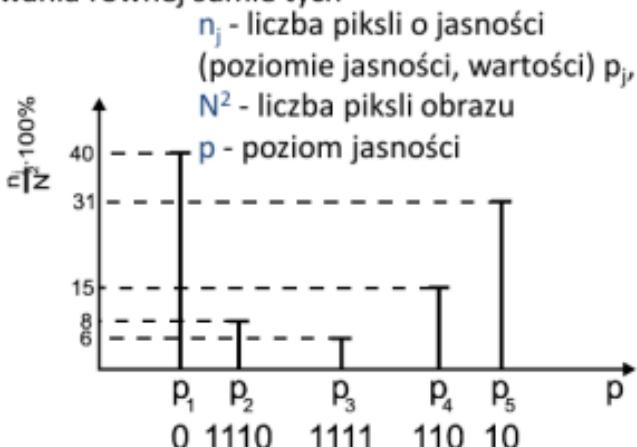
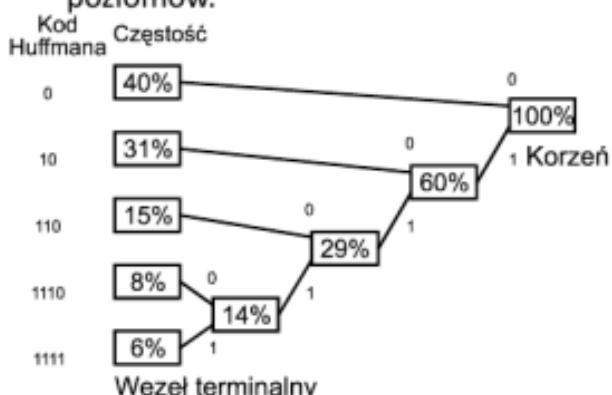
- formaty obrazowe: *.PCX, *.PIC

Kompresja obrazów z obszarami o niejednolitej jasności

Kod Huffmana - kod o zmiennej długości słowa.

Sposób postępowania:

- Przypisanie każdemu poziomowi jasności częstości występowania pikseli o tym poziomie jasności (utworzenie histogramu).
- Wyszukanie 2 poziomów o najmniejszej częstości występowania i połączenie w jeden o częstości występowania równej sumie tych poziomów.



Schemat łączenia - drzewo Huffmana.

Praktyczna realizacja kodu Huffmana - format obrazowy *.TGA

Metody słownikowe

Kodowanie *ciągów symboli* (pikseli) za pomocą odwołań do *słownika* zawierającego takie ciągi.

Stopień kompresji (SK) rośnie, gdy **długość ciągu pikseli** możliwych do zastąpienia *indeksem* do słownika rośnie.

Podział metod słownikowych:

- *statyczne* (słownik nie zmienia się w trakcie kodowania),
- *adaptacyjne* (słownik zmienia się w trakcie kodowania)

Kodowanie różnic

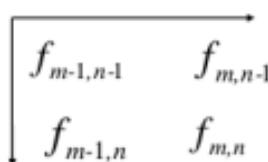
$$\epsilon_{m,n} = f_{m,n} - f_{m,n-1}$$

gdzie:

$f_{m,n}$ - poziom jasności piksela o współrzędnych m,n ,

$f_{m,n-1}$ - poziom jasności piksela (**kolejnego**) o współrzędnych $m,n-1$.

Położenie pikseli



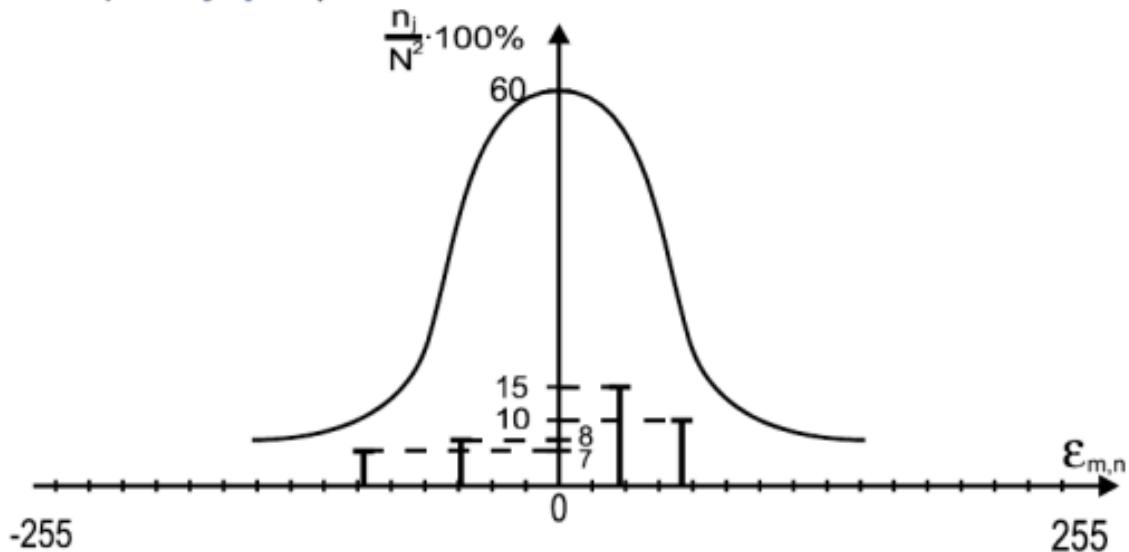
Przykład:

Obraz pierwotny : $L = 255$ ($M = 256$)

- kodowanie wartości pikseli: $I = 0,1,2,3,\dots,255$; zajętość pamięci: 8 bitów/piksel,
- kodowanie różnic pomiędzy wartościami kolejnych pikseli:

$\epsilon_{m,n} = -255, -254, \dots, 0, \dots, 254, 255$ zajętość pamięci: 9 bitów /piksel.

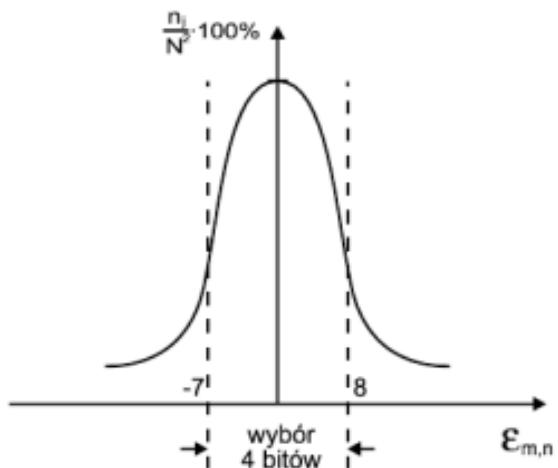
Histogram różnic $\varepsilon_{m,n}$ pomiędzy wartościami (poziomami jasności) kolejnych pikseli



Histogram różnic ma charakter *krzywej Gaussa* z maksimum w okolicy zera.

Kompresja wynika z zakodowanie najczęściej występujących różnic

Histogram różnic z zaznaczonym obszarem najczęściej występujących różnic



Kryteria wyboru obszaru:

- wymagania na wielkość stopnia kompresji (SK),
- wymagania na dokładność rekonstrukcji (określoną wielkością „ p ”)

$$KP = 9, KW = 4$$

$$SK = KP/KW = 9/4 = 2,25$$

Obraz zakodowany $\epsilon_{m,n} = -7, -6, \dots, 7, 8$ - 16 symboli zamiast 512

Niepożądany efekt: zależnie od rodzaju obrazu - większe lub mniejsze rozmywanie (*blurring*) ostrych krawędzi.

Kodowanie blokowe

Podział obrazu na jednakowe bloki, najczęściej 4×4 piksele. Obliczenie dla każdego bloku średniej arytmetycznej jasności. Podział pikseli na dwie grupy:

- a) o jasności większej lub równej jasności średniej,
- b) mniejszej niż jasność średnia.

Obliczenie nowej jasności średniej dla każdej z grup (**wartość góra** dla (a) i **wartość dolna** dla (b)). Przypisanie wszystkim pikselom danej grupy obliczonej jasności średniej (górnjej lub dolnej), stąd blok zostaje zakodowany jako *mapa bitowa* określająca podział na grupy, plus dwie wartości jasności.

The diagram illustrates the process of block-based coding. On the left, a 4x4 grid of pixel values is shown:

13	11	10	13
8	9	12	15
6	8	11	14
4	7	9	12

An arrow points from this grid to the right, where a 4x4 binary matrix is displayed:

1	1	0	1
0	0	1	1
0	0	1	1
0	0	0	1

Średnia=10,125

Wartość góra= 12,625≈13
Wartość dolna = 7,625≈ 8

Kompresje obrazów oparte na transformatach

- Wybieram takie transformaty, które dają w dziedzinie transformacji takie upakowanie informacji, aby większość współczynników była zerowa (bliska零)
- Po kwantyzacji niezerowych współczynników dają minimalne zniekształcenia obrazu
- Prowadzą do dekorelacji sygnału
- Są ortogonalne, bo to zapewnia pełną transmisję informacji
- Są separowalne, bo to zapewnia szybki algorytm dekompozycji

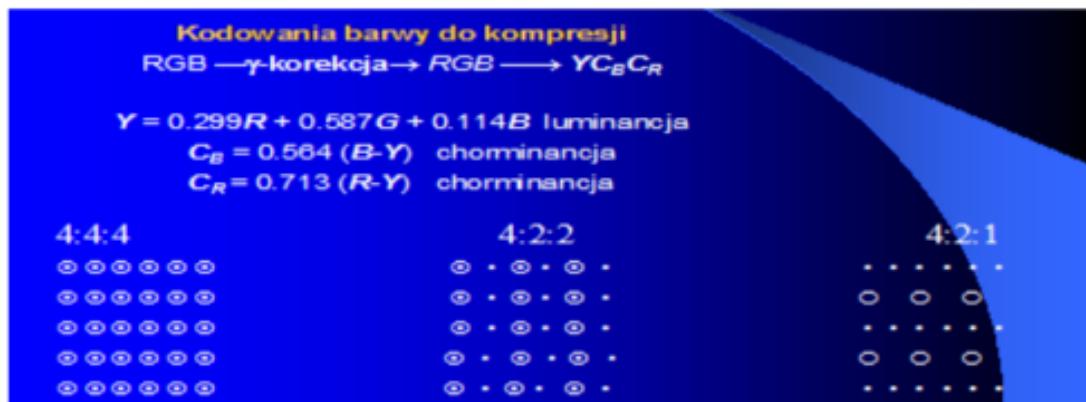
Kompresja stratna oparta na transformacjach

- Transformacja kosinusowa w kompresji JPEG
- Transformacja falkowa w kompresji JPEG 2000

Kolejność działań (i):

Konwersja obrazu kanałów (RGB) na jasność (luminancję) i 2 kanały barwy (chrominancje).

Odrzucenie części pikseli kanałów barwy, podział kanałów na bloki 8×8 pikseli.



Kolejność działań (ii)

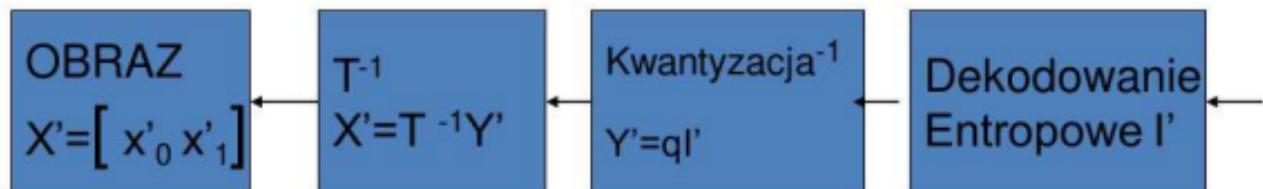
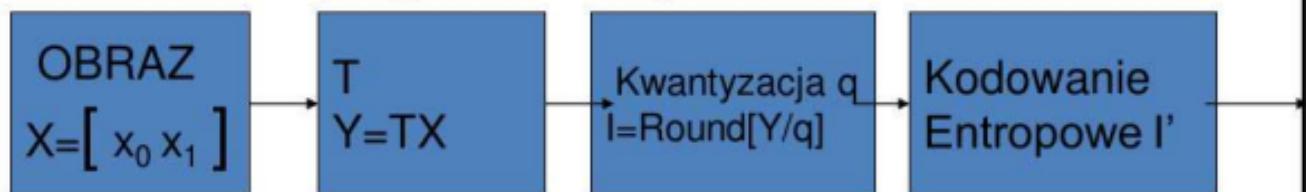
Na blokach wykonywana jest dyskretna transformata kosinusowa (DCT).

Zastąpienie średnich wartości bloków przez różnicę wobec wartości poprzedniej (DPCM). Poprawia to w pewnym stopniu współczynnik kompresji.

Kwantyzacja, czyli zastąpienie danych zmiennoprzecinkowych przez liczby całkowite.

Schemat kompresji opartej na transformacjach

Kodowanie (ang. encoder)



Dekodowanie (ang. decoder)

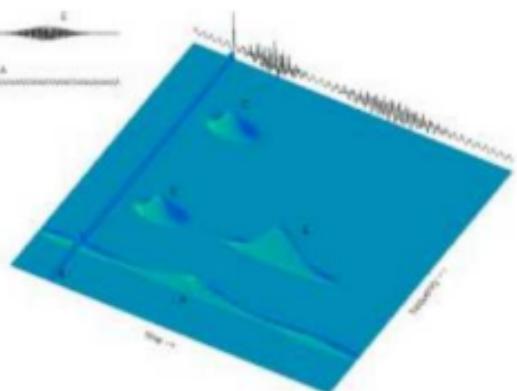
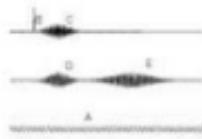
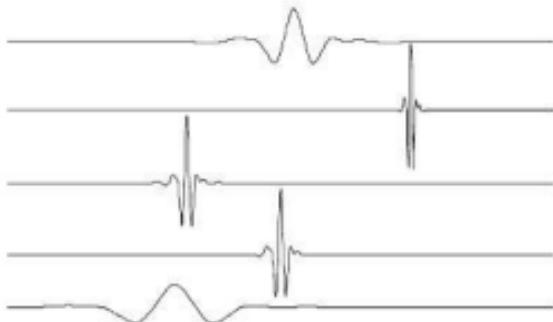
Kompresja JPEG200

- Koduje różne typy obrazów i filmów: binarne, szare, kolorowe, wielospektralne o różnych właściwościach, transmisja w czasie rzeczywistym, archiwizacja biblioteki obrazów.
- Umożliwia progresywne dekodowanie: pozwala widzowi zobaczyć wersję danego obrazu o niższej jakości, podczas gdy cały plik nie został w pełni odebrany lub pobrany. Jakość obrazu stopniowo się poprawia, gdy więcej bitów danych jest pobieranych ze źródła.
- Zapewnia zachowanie przejrzystości obrazów.
- Daje niezwykle wysoki poziom skalowalności.
- Obrazy mogą być kodowane w dowolnej jakości od bardzo skompresowanych do kompresji bezstratnej.
- Firma udostępnia oprogramowanie open-source tzn. OpenJPEG kod napisany w C do wykorzystania przez programistów.
- Stosuje transformację falkową

Co to są falki?

Komplet funkcji umożliwiających hierarchiczną wielorodzicielną reprezentację aproksymowanej funkcji, składający się z :

- Funkcji skalującej $\Phi(x)$
- Rodziny falek generowanych na podstawie falki podstawowej



Falki i transformacja falkowa w przetwarzaniu obrazów

- Kompresja obrazów i sekwencji wideo
- Odszumianie obrazów
- Analiza tekstury
- Segmentacja
- Rozpoznawanie np.: ręcznego pisma, biologicznego wieku dzieci, charakterystycznych cech biologicznego sygnału (EEG)
- Rekonstrukcja obrazów biomedycznych

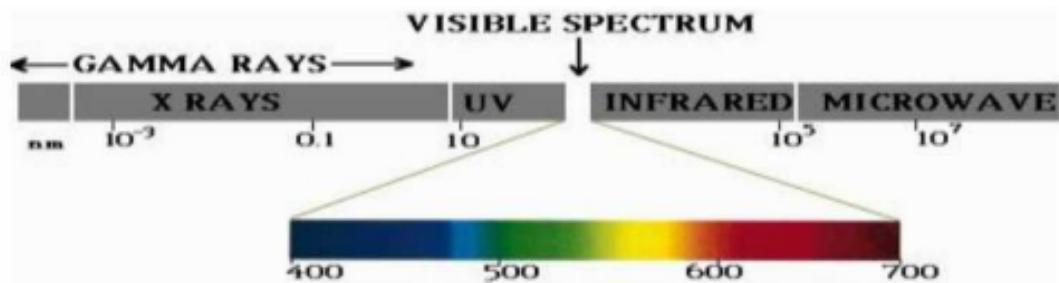
Metody oceny jakości kodowania

- Subiektywne – notowanie wrażeń widzów lub słuchaczy
- Obiektywne – porównanie ilościowe, na podstawie matematycznie wyznaczonych parametrów, opisujących różnicę między sygnałem oryginalnym i po dekompresji
- Metody oparte na modelach percepcji informacji przez człowieka:
 - perceptualny model, zapewniający wierność percepcji słuchowej
 - modelowanie układu percepcji wzrokowej człowieka
(HVS ang. Human Visual System)

Barwa

Barwa to zarówno kombinacja fizycznych (spektralnych) właściwości światła jak i sposób jego interpretacji przez ludzkie oko i mózg.

Różne długości fali elektromagnetycznej z zakresu widma widzialnego (380 - 760 nanometrów) są odbierane jako barwy, a ich mieszanina jako światło białe



Uwaga! Jednakowe bodźce barwne wywołują jednakowe wrażenia, ale takie samo wrażenie mogą wywołać bodźce różniące się fizycznym (spektralnym) składem promieniowania elektromagnetycznego (metameryzm)

Barwa

Fizycznie: *długość fali*

Subiektywnie: to co odróżnia zieleń od błękitu a jest wspólnie dla różnych *odcieni* czerwonego.

Jasność

stopień podobieństwa do barwy białej (dla *odcieni* jasnych) lub czarnej (dla *odcieni* ciemnych).

Nasycenie

czystość barwy np. stopień zbliżenia do **barw zasadniczych** występujących w widmie słonecznym: **czerwona, zielona, niebieska, żółta (RGBY)**.

Rozróżnialność barw

(uwzględniając jasności) **ok. 400000 kolorów (przy porównywaniu)**.

Z pamięci: **kilkadziesiąt barw**. Subiektywna ocena barw obrazu na monitorze przy różnym oświetleniu pomieszczenia.

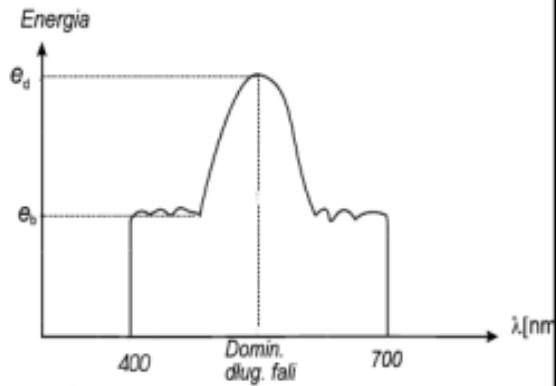
Cechy bodźca świetlnego wywołującego wrażenie barwy

Achromatyczne (widzenie **skotopowe**)

- Ilość (skuteczność pobudzenia) luminancja/jasność/jaskrawość czyli stopień podobieństwa do białej barwy, dającej maksymalne pobudzenie lub czarnej, dającej pobudzenie minimalne

Chromatyczne (widzenie **fotopowe**)

- Jasność (intensywność pobudzenia R+G+B):
- Barwa/odcień/wałor/Kolor/ (fizycznie: długość dominującej fali)
- Nasycenie/rozbłyszczenie (czystość barwy czyli podobieństwo do barw widma słonecznego, bladość)



Widzenie pośrednie **mezopowe**; przy słabym świetle

Standard barw podstawowych

Standard CIE 1931r. (Międzynarodowa Komisja Oświetleniowa).

Standardowe barwy podstawowe nie odpowiadają żadnej rzeczywistej barwie, za to dowolną widzialną barwę daje się wyrazić jako ich średnią ważoną.

A,B,C - ilości poszczególnych barw podstawowych CIE dających w sumie pewną barwę.

Wielkości

$$a = \frac{A}{A+B+C}; \quad b = \frac{B}{A+B+C}; \quad c = \frac{C}{A+B+C}$$

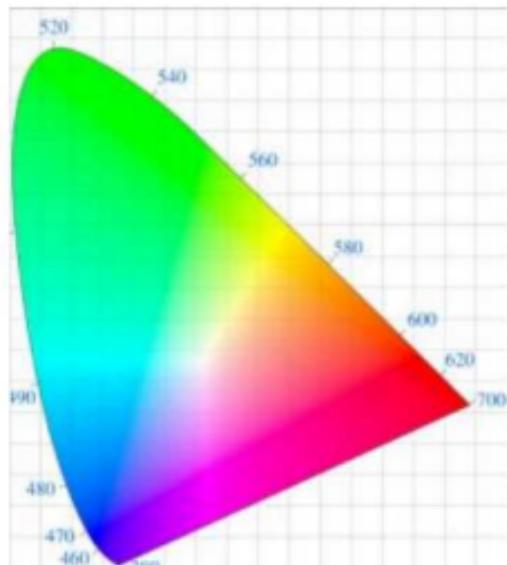
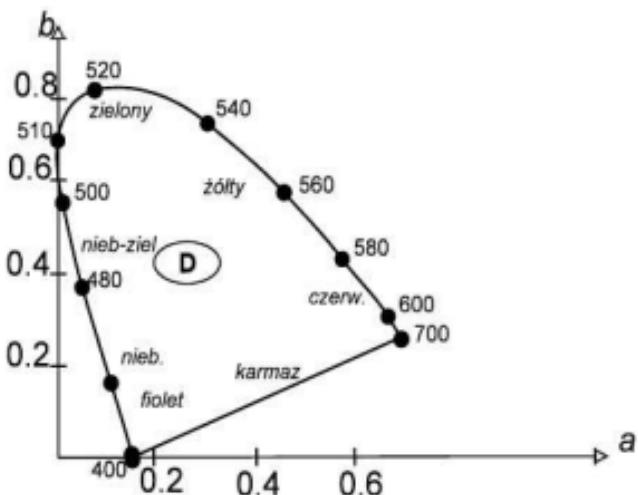
są to **współrzędne trójchromatyczne** tej barwy.

Widać, że zawsze: **a + b + c = 1**

tzn. **dowolne dwie** współrzędne wystarczą do określenia barwy.

Diagram chromatyczności CIE

wykres współrzędnych a i b wszystkich barw widzialnych; obszar D - barwa biała

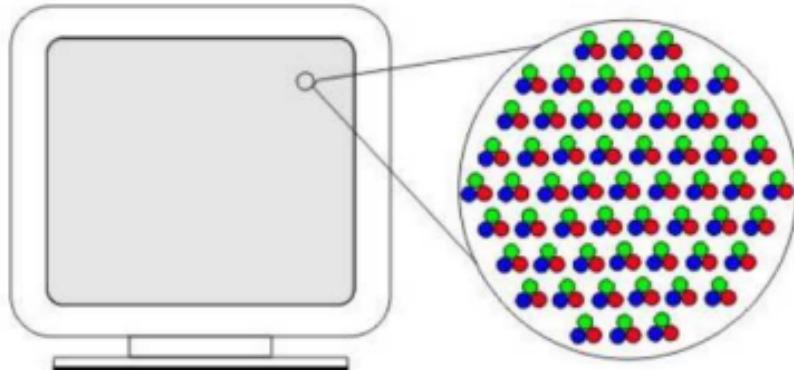


Cel

otrzymanie obiektywnych, jednoznacznych charakterystyk barw.

Mieszanie addytywne (synteza barwy)

- Równoczesne działanie trzech światel podstawowych (modele: trzy projektorów lub komponowanie z trzech małych punktów)



- Naprzemienne rzutowanie bodźców barwnych z częstotliwością większą od progu migotania (synteza w oku odbiorcy)
- Jeżeli obraz cyfrowy ma głębię 24 bity, to każdemu pikselowi można przyporządkować jedną z 16 777 216 barw

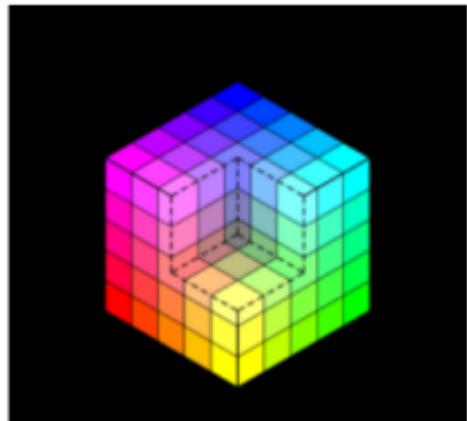
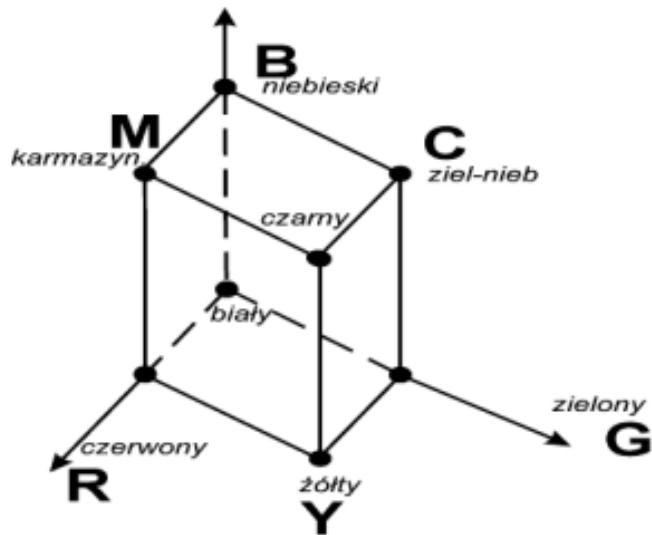
Model RGB

Opis piksela:

3 bajty → 16777216 barw

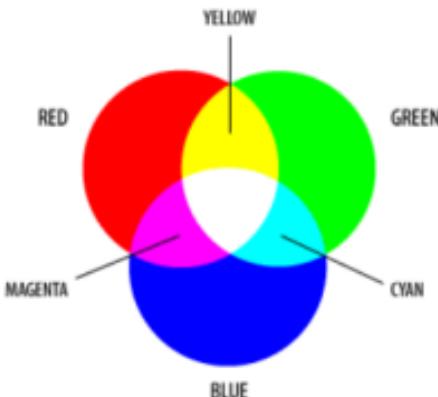
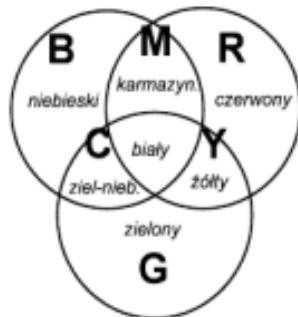
3 bity → $2^3 = 8$ barw

1 piksel na ekranie monitora składa się **fizycznie** z 3 elementów (RGB)



Barwy addytywne

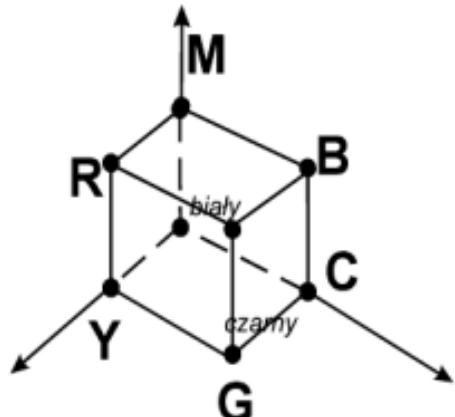
Model RGB - opis addytywnego procesu otrzymywania barw w wyniku dodawania (mieszania) w różnych proporcjach trzech barw podstawowych.



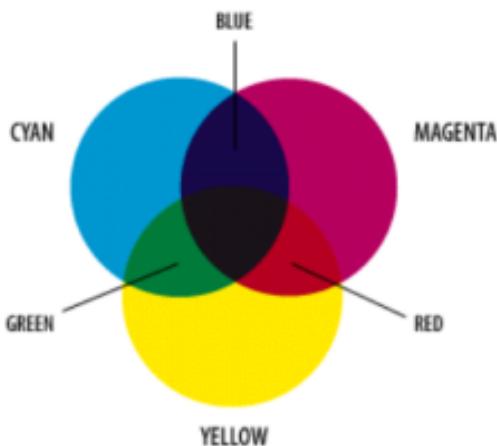
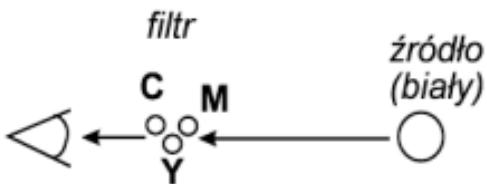
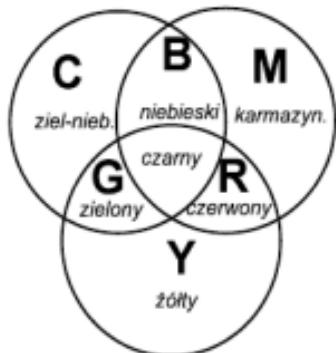
Model CMY

C - cyan (zielono-niebieski), M - magenta (karmazynowy), Y - yellow (żółty) (są to *dopełnienia* R, G, B).

CMY są podstawowymi barwami **subtraktywnymi**, ponieważ **usuważą** podstawowe barwy ze źródła białego. Np. przedmiot pokryty żółtą farbą pochłania składowe światła słonecznego o długościach dopowiadających barwie niebieskiej. Model CMY jest subtraktywny i służy do określenia kolorów obrazów na drukarkach i ploterach.



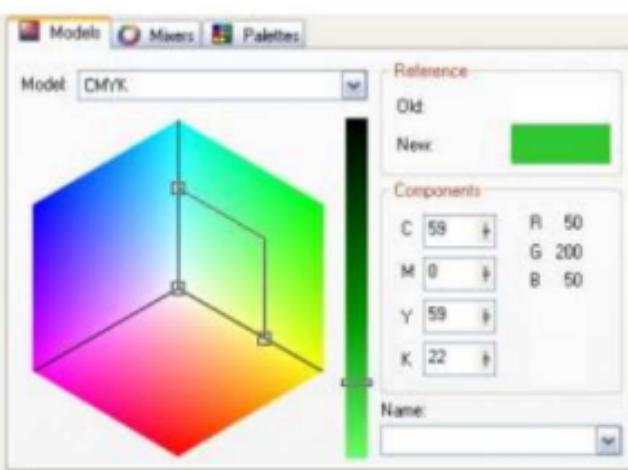
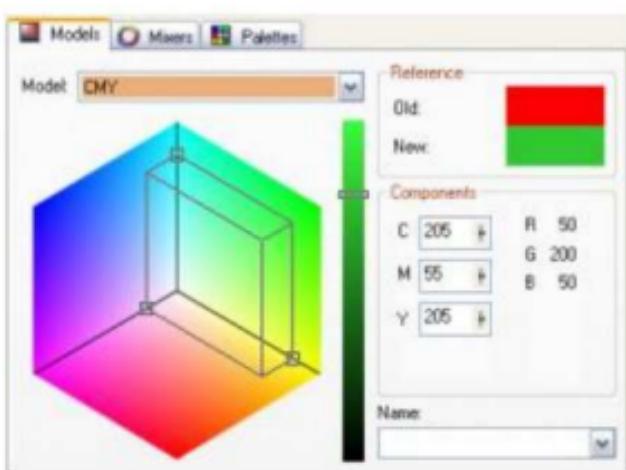
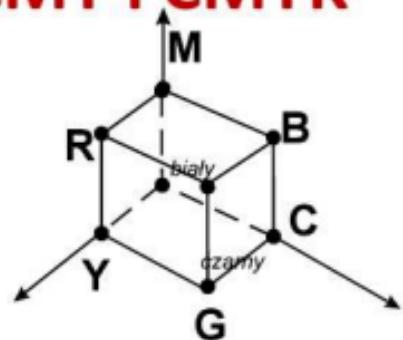
Barwy subtraktywne



Mieszanie substraktywne (filtrowanie barwy)

- Światło i trzy substancje posiadające właściwości selektywnej absorpcji barwy, zwane barwnikami (farbami) i działające jak filtry światła
- Możliwość mieszania roztworów barwników (farb) lub ich nakładania na siebie warstwami (fotografia barwna) albo układania punktowego w postaci mozaiki o rozmiarach poniżej progu rozróżnialności (lakiery, druk offsetowy kolorowy)
- Model CMYK
 - C (Cyan czyli zielono niebieski) = $1 - R$
 - M (Magenta czyli purpurowy) = $1 - G$
 - Y (Yellow czyli żółty) = $1 - B$
 - B (Black czyli czarny) = $\min(C, M, Y)$

Modele substraktywne CMY i CMYK



Modele percepcyjne

MODEL HSV POZWALA NA ŁATWE OPISANIE WYMAGANEJ BARWY (czynności malarza przy doborze farby)

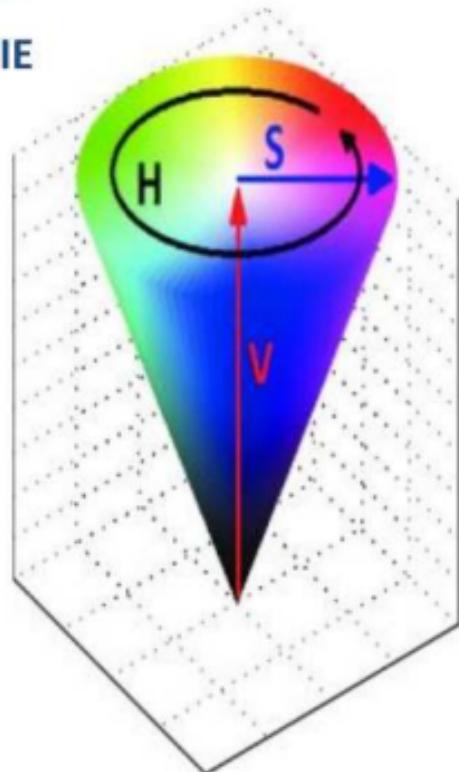
podanie β przy $V=S=1$

-ściemnianie (zmnieszanie V) i/lub rozjaśnianie (zmnieszanie S)

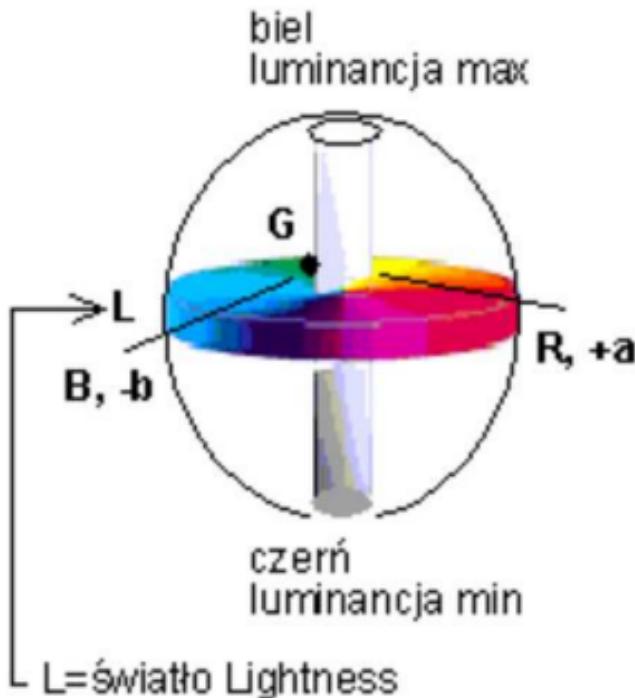
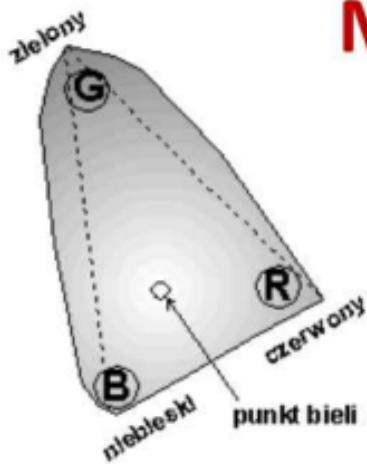
Do wyświetlenia na monitorze barw zdefiniowanych w HSV potrzebne są współrzędne RGB

Dlatego stosuje się:

Algorytm transformacji (przekształcenia) modelu HSV na model RGB



Model CIELAB



Etapy proces przetwarzania i rozpoznawania

Komputerowa analiza obrazów

Podejście klasyczne

- Segmentacji obrazu
- Pomiaru i lokalizacji obiektów na obrazie
- Zrozumienia treści niesionych przez obraz w sensie np. klasyfikacji obiektów, policzenia ich liczby lub określenia ich wzajemnych relacji

Podejście z użyciem głębokich sieci neuronowych



Cel segmentacji

Przygotowanie obrazu do etapu właściwego rozpoznawania obiektów, określenia relacji przestrzennych pomiędzy nimi.

Segmentacja stanowi poziom pośredni pomiędzy poziomem *wstępnego przetwarzania* a poziomem *analizy obrazu*.

Segmentacja

- Segmentacja to podział obrazu na **rozłączne** (nienakładające się) fragmenty.
- Segmentacja jest **powiązana z semantyką** (znaczeniem i rozumieniem) obrazu.

bywa rozumiana dwojako:

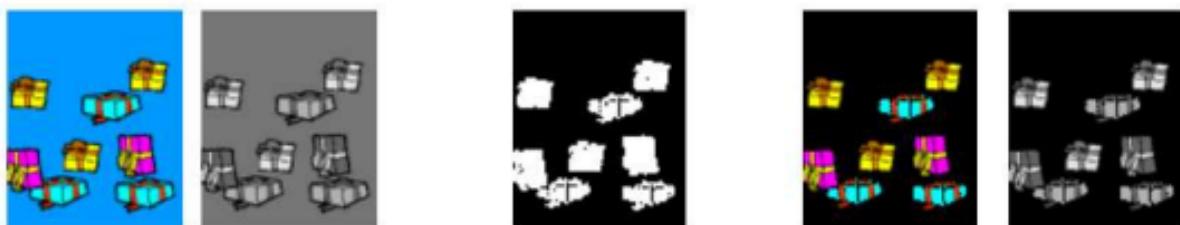
- Jako podział na jednorodne rejony, które składają się na znaną hierarchię lub strukturę
- Jako podział na to, co nas interesuje z punktu widzenia celu przetwarzania, pozostałe nieinteresujące obiekty i tło



Segmentacja

Wyodrębnienie spośród wybranych fragmentów tych, które stanowią obiekt zainteresowania ze względu na cel analizy obrazu.

Maska/Mapa



Najbardziej skomplikowane algorytmy

Klasyfikacja metod segmentacji

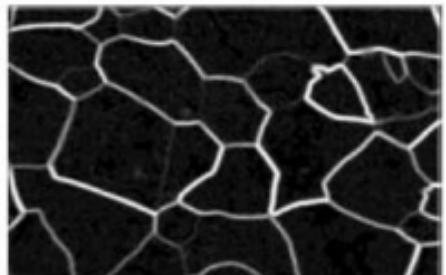
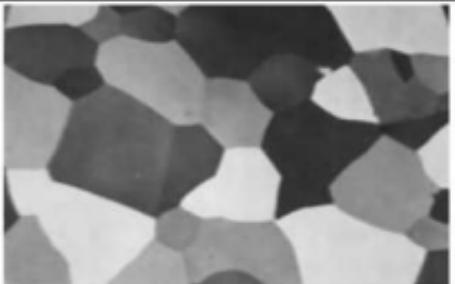
Segmentacja może być zarówno operacją kontekstową jak i niekontekstową (punktową), ale najczęściej jest kombinacją metod kontekstowych i punktowych.

Jeśli metoda:

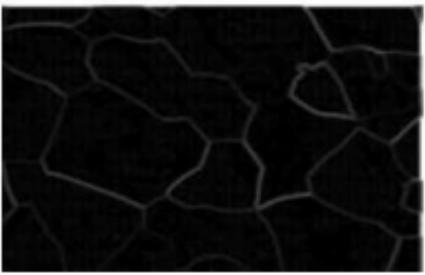
- Ignoruje zależności między pikslami i klasyfikuje je na podstawie globalnej cechy, np. wartości poziomu szarości – progowania, to jest metodą punktową niekontekstową
- Wykorzystuje zależności między pikslami, np.: podobieństwo wartości poziomu szarości – „dziel i łącz”, to jest metodą kontekstową



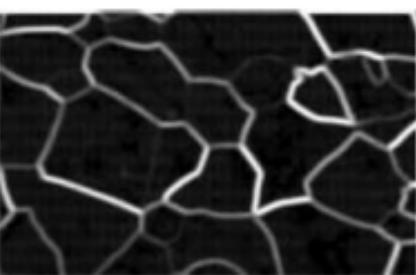
Etapy procesu segmentacji



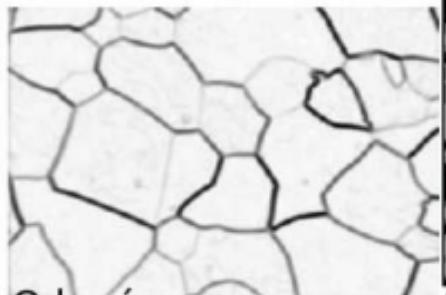
Operator Sobela



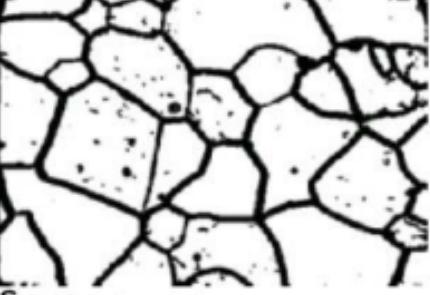
Operator Robertsa



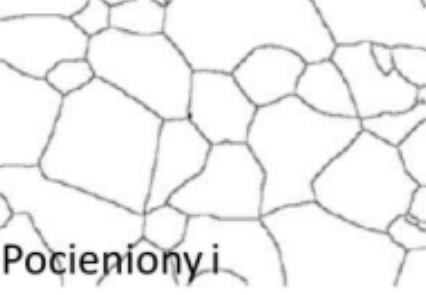
Wariancja 5x5



Odwrócony
operator Sobela



Sprobowany
odwrócony
operator Sobela



Pocięciony i
oczyszczony sprobowany
odwrócony operator Sobela

Rodzaje segmentacji

(około 1000 algorytmów segmentacji)

- Poszukiwanie nieciągłości czyli poszukiwanie krawędzi, a w konsekwencji wskazanie wnętrza obiektu (np. z wykorzystaniem histogramu dwuwymiarowego),
- Maksymalne obszary wykazujące podobieństwo w kolorze, odcieniu szarości, tekście (*k-means* lub *graph-cut*)
- Wododziały (ang. *watershed transform*),
- Podziały przeszukujące obszar (*drzewa czwórkowe*, *split and merge –region*)
- Dopasowywanie konturów np. metoda aktywnego konturu, (ang. *Active contour*).

Uwaga

W wyniku stosowania obszarowych metod segmentacji uzyskuje się **zawsze** obszary zamknięte (granice obszarów są ciągłe). Jest to zaleta w porównaniu np. z metodami *detekcji krawędzi*, które na ogół **nie zapewniają** ciągłości wykrytych krawędzi.

Techniki segmentacji

Progowanie

Warunki na dobre wyniki segmentacji :

- istnieje rozdzielność poziomów szarości lub kolorów obiektu i tła
- gdy wartość minimum lokalnego w histogramie jest jak najmniejsza (ang. *valley method*)

Podobieństwo tekstury

Tekstura reprezentuje, pewną relatywną jednorodność/jednolitość, odczuwaną wzrokowo przez odbiorcę lub udowadnianą jako matematyczna regularność dzięki analizie sygnału.

Progowanie - najprostszy algorytm segmentacji

- Progowanie z pojedynczym progiem segmentacji

$$q = \begin{cases} L_{\min} & \text{dla } p \leq p_1 \\ L_{\max} & \text{dla } p > p_1 \end{cases}$$

- Progowanie przedziałami

$$q = \begin{cases} L_{\max} & \text{dla } p_1 \leq p \leq p_2 \\ L_{\min} & \text{dla } p < p_1 \text{ lub } p > p_2 \end{cases}$$

- Progowanie z warunkiem spójności
lub warunkiem nałożonym na wielkość obiektu

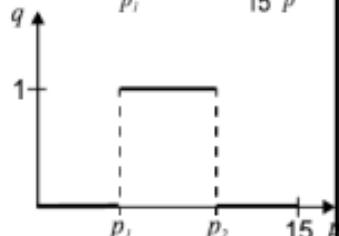
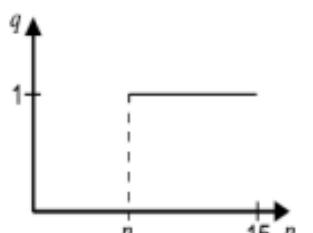
- Progowanie adaptacyjne**

- Progowanie rekurencyjne

- Progowanie hierarchiczne (piramidowe, skalowalne)

Progowania dokonujemy na obrazie szarościowym lub pojedynczej składowej koloru lub na kanale wspólnym, np. luminancji

Operacja punktowa,
bezkontekstowa

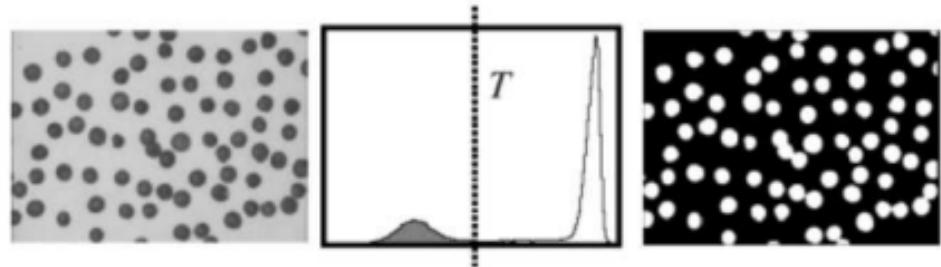
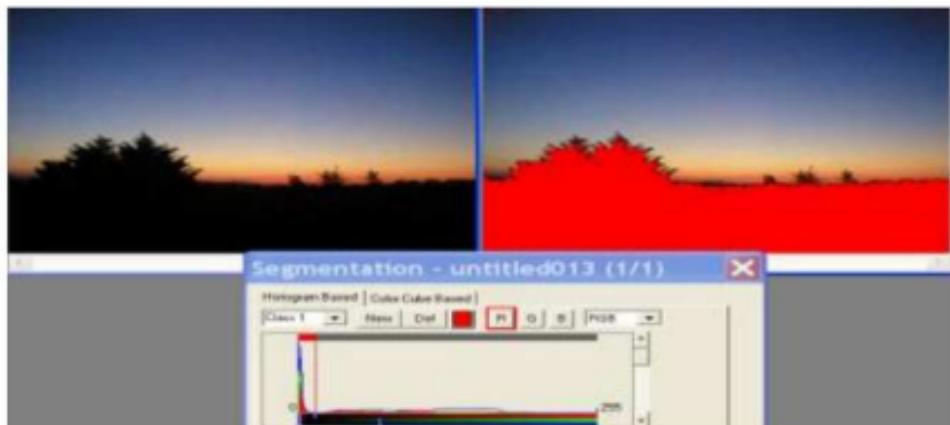


Przykład progowania

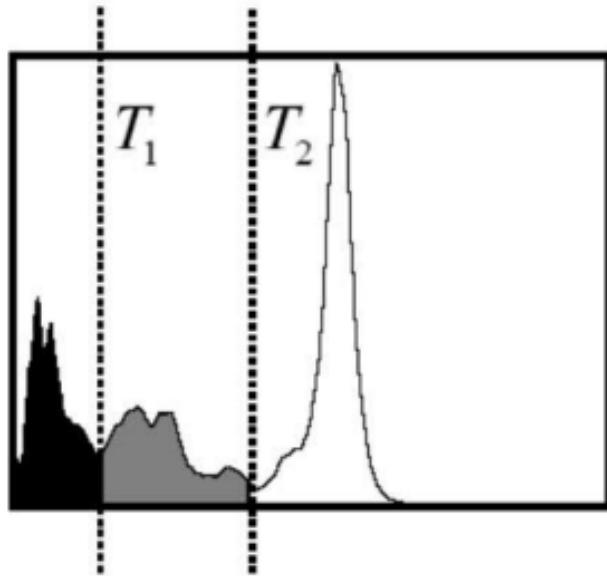
Progowanie dobrze segmentuje tylko wtedy, gdy:

- istnieje rozdzielność poziomów szarości lub kolorów obiektu i tła

- gdy „dolina” progu jest głęboka (najlepiej zerowa)



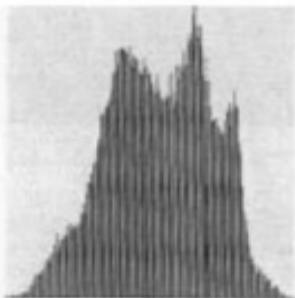
Progowanie przedziałami



Progowanie adaptacyjne



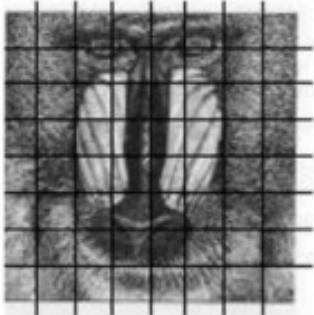
(A)



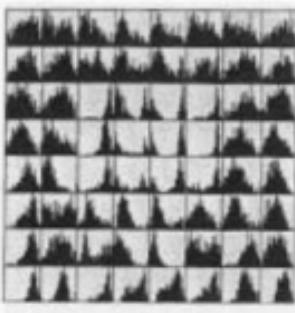
(C)



(E)



(B)

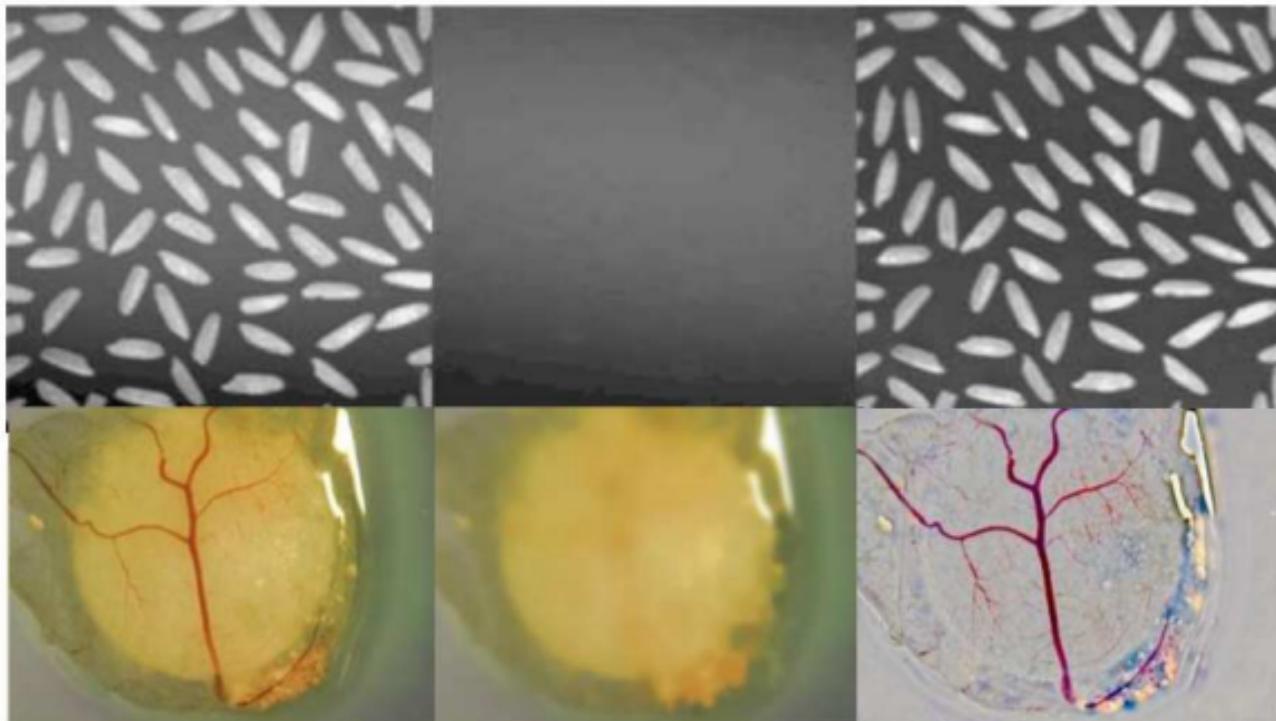


(D)



(F)

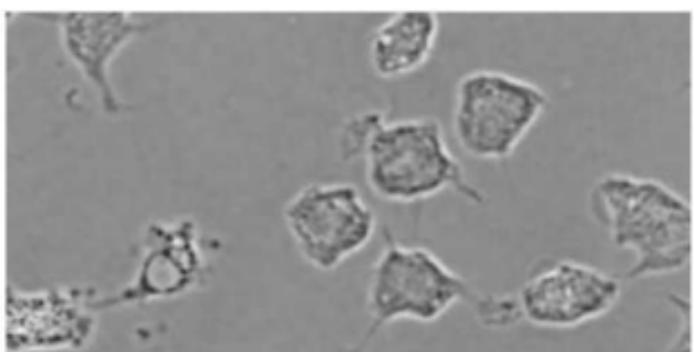
Wyrównanie nierównomierności tła



Progowanie

- Zalety
 - Prosta procedura zaimplementowana w każdym oprogramowaniu do manipulacji obrazami
 - Szybka realizacja algorytmu
- Wady
 - Trudności dla obrazów o niskim kontraście
 - Trudności przy nierównomiernym oświetleniu
 - Niejednoznaczności wynikające z braku rozdzielności poziomów szarości lub kolorów, występujących w tle i obiekcie

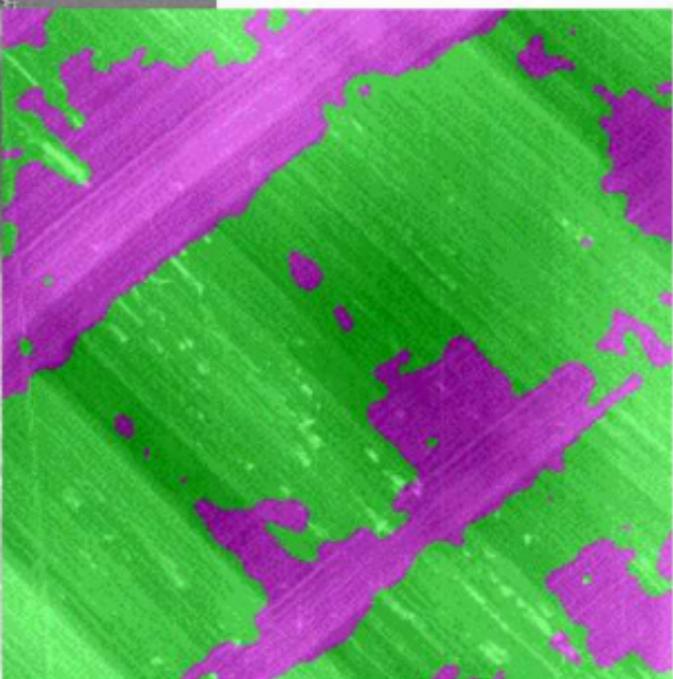
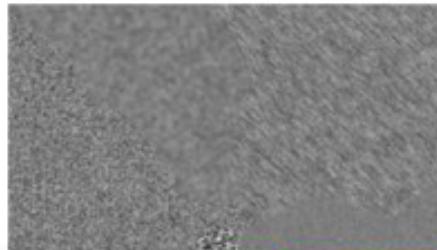
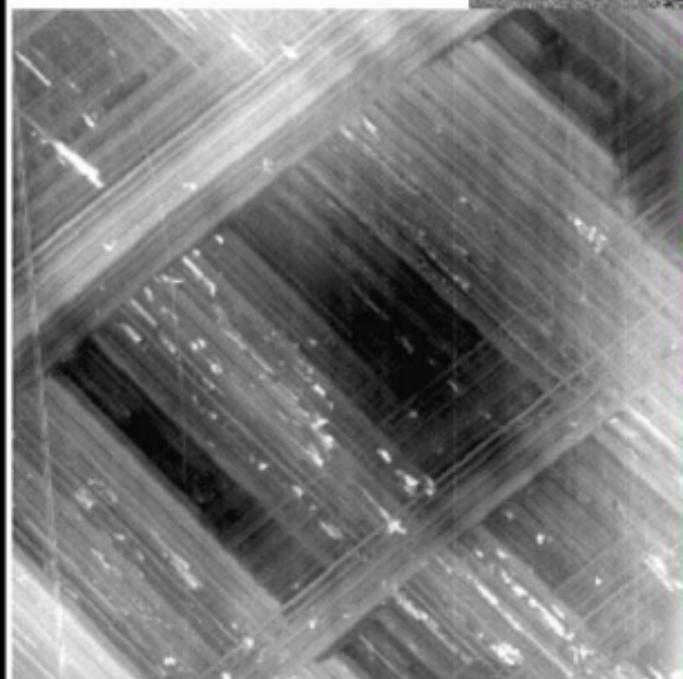
Tekstura



Jednorodność jest oparta na powtarzalności konstrukcyjnego elementu, wewnątrz którego istnieje pewna nierównomierność poziomów szarości (relacja między podelementami elementu konstrukcyjnego, czyli połączonymi grupami sąsiadujących ze sobą pikseli, jest stała) lub organizację lub uporządkowanie elementów w przestrzeni.

Analiza tekstury

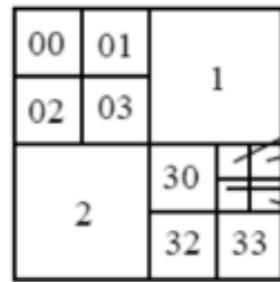
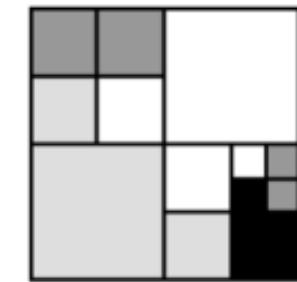
- oparta na regularności ocenianej metodami statystycznymi, na podstawie macierzy opisującej częstość występowania dwóch pikseli odległych od siebie o dystans d w kierunku θ (po angielsku **cooccurrence matrix**), zdefiniowaną przez Haralicka
Cechy: kierunkowość, ziarnistość, ...
- oparta na różnych zaawansowanych modelach matematycznej regularności (np. model powtarzalności/zależności poziomów szarości w różnych kierunkach obrazu, oparty na stochastycznych polach Markowa ang. **random Markov field** lub model fraktałny samopodobieństwa ang. **fractal model**)
- oparta na morfologii matematycznej, która używa różnych transformacji do porównywania struktur w obrazie do znanego elementu konstrukcyjnego tekstury



Techniki obszarowe segmentacji

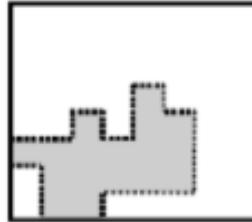
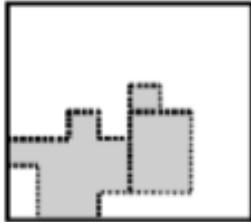
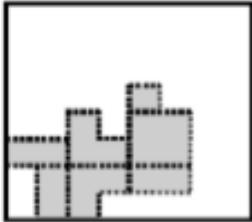
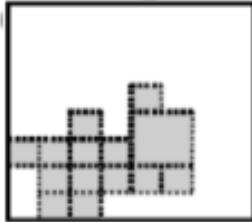
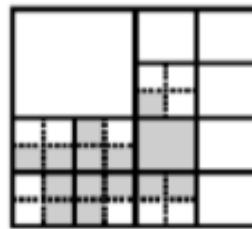
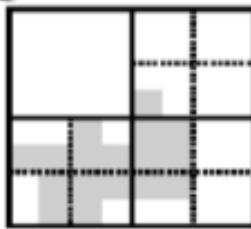
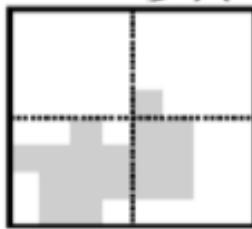
- segmentacja przez **podział obszaru** (*region - splitting*)
- segmentacja przez **rozrost obszaru** (*region - growing*)
- segmentacja dziel i łącz

Dziel iłącz

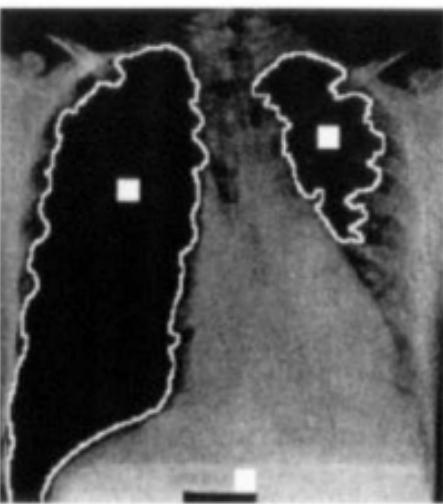
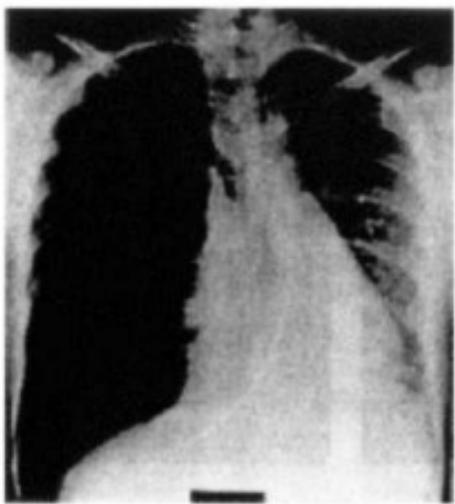


Zapis drzewa
czwórkowy

310
311
312
314



„Rozrost ziarna”



Segmentacja w oparciu o kolor



Obraz



a) Luminance



b) Saturation

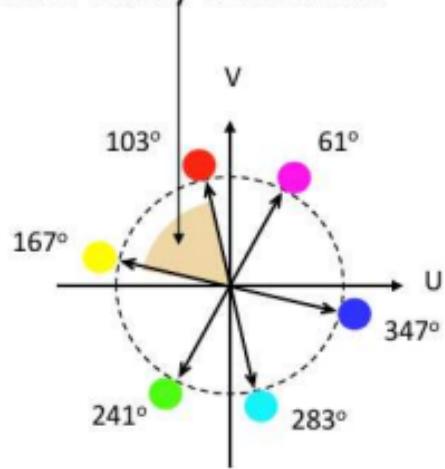


c) Hue



Poszukiwanie ludzi na kolorowych obrazach

Kolor skóry człowieka



Chrominancje w modelu YUV

Segmentacja w oparciu o poszukiwanie krawędzi

Typy krawędzi

Skok – stopień



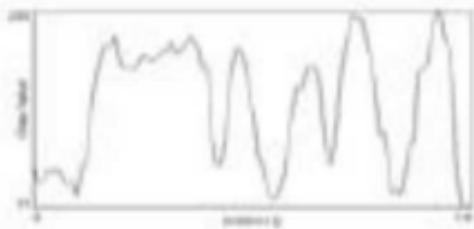
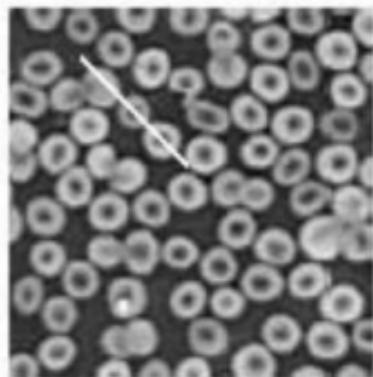
Rampa



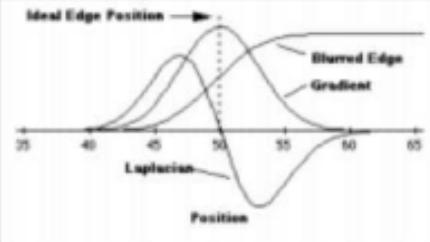
Dach



Linia



Poszukiwanie krawędzi



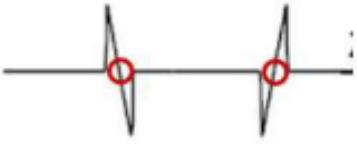
Krawędzie



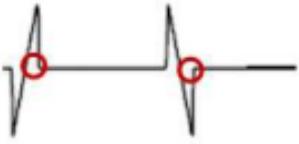
Pierwsza
pochodna
gradient



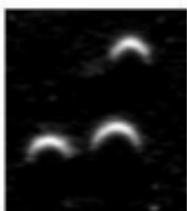
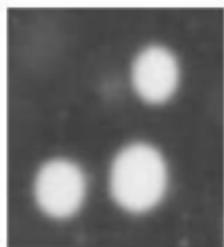
Druga
pochodna



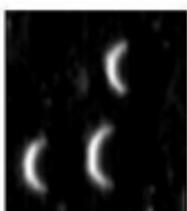
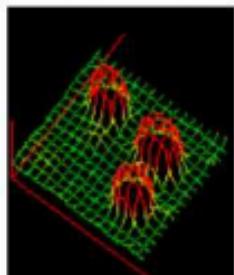
laplasjan



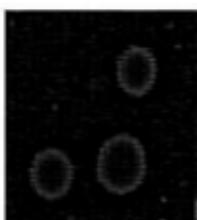
Krawędziowanie



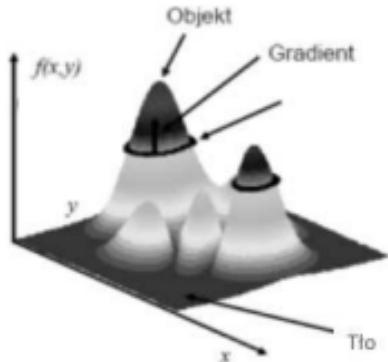
Gradient w
kierunku
poziomym



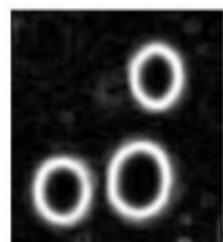
Gradient w
kierunku
pionowym



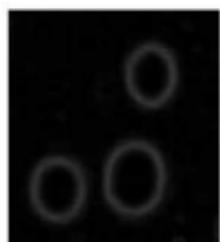
Laplasjan



Operator
Sobela



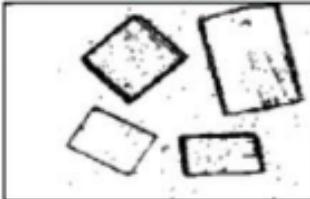
Operator
Roberta



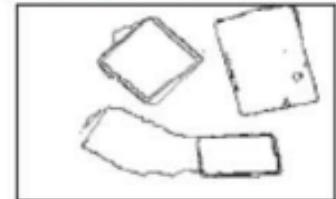
Porównanie krawędzi segmentowanych na podstawie różnych kanałów koloru



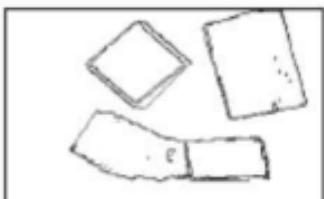
(a)



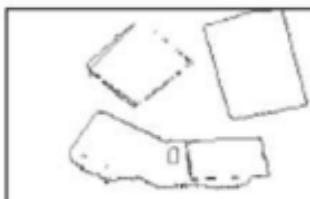
(b)



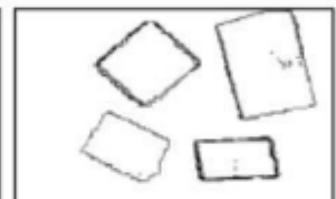
(c)



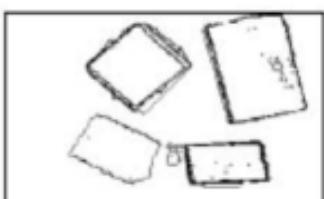
(d)



(e)



(f)

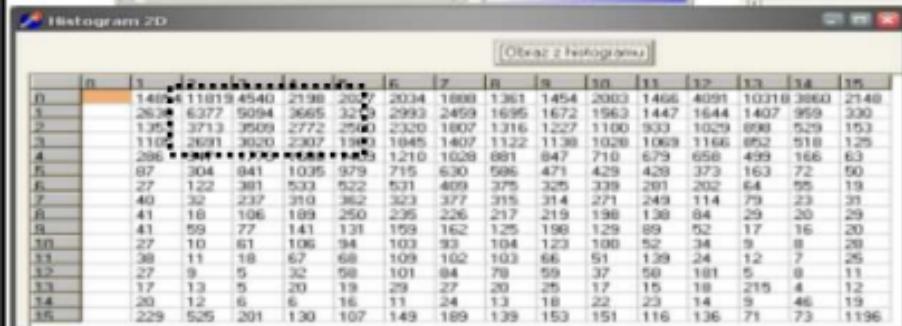
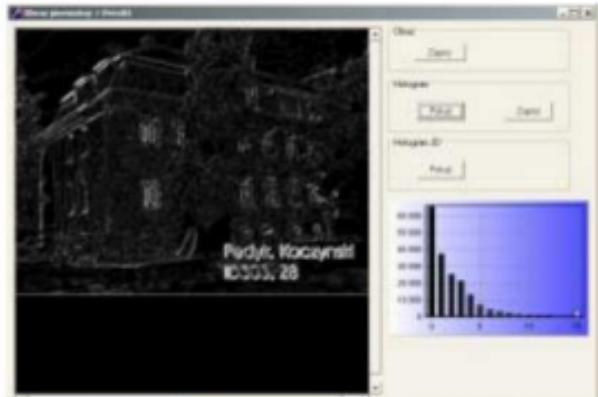


(g)

(a). Obraz; (b). Dla modelu RGB; (c). Dla kanałów U i V; (d). Dla modelu YUV; (e). Dla kanału I ; (f). Dla kanałów H i S ; (g). Dla modelu IHS

Histogram 2D

Wykorzystanie operacji krawędziujących do segmentacji



Obraz pierwotny

0	0	0	0	0	0	0	0	0	0
0	2	2	2	1	2	2	2	1	0
0	2	2	2	1	1	2	2	4	0
0	1	2	2	7	6	6	2	1	0
0	1	2	7	6	5	6	2	2	0
0	1	7	6	5	5	6	1	2	0
0	1	7	7	6	6	6	2	2	0
0	1	2	1	1	2	2	1	5	0
0	1	2	1	1	1	2	2	3	0
0	0	0	0	0	0	0	0	0	0

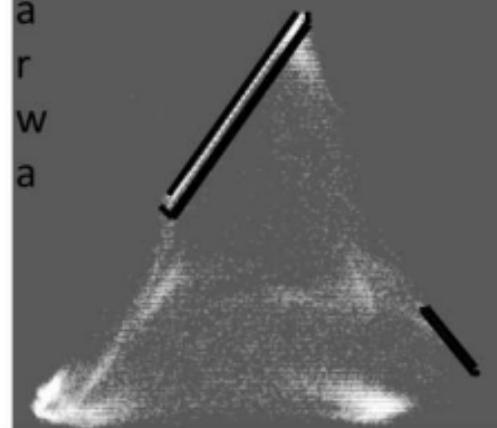
Obraz wynikowy (metoda gradientowa Sobel'a)

0	0	0	0	0	0	0	0	0	0
0	2	2	2	1	1	2	2	3	0
0	2	0	1	4	5	3	1	2	0
0	2	2	5	5	4	4	3	2	0
0	3	6	4	0	1	4	4	2	0
0	6	7	1	1	0	4	4	1	0
0	6	6	5	4	4	4	4	2	0
0	3	4	6	6	5	3	1	1	0
0	2	1	1	1	2	2	2	3	0
0	0	0	0	0	0	0	0	0	0

7	0	0	0	0	0	0	0	1
6	0	4	1	0	0	0	0	1
5	0	1	2	0	0	0	0	2
4	0	2	2	0	0	0	7	1
3	0	3	3	1	0	0	0	0
2	0	4	12	0	1	0	0	0
1	0	4	5	0	0	3	1	0
0	36	0	1	0	0	1	1	0
	0	1	2	3	4	5	6	7



B Luminamcja
a
r
w
a



Barwa



Luminancia

Yellow

Orange

Blue

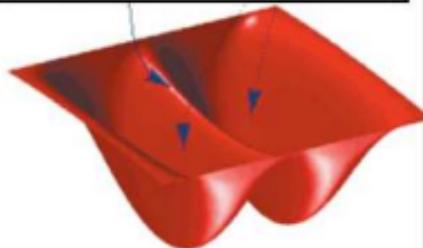
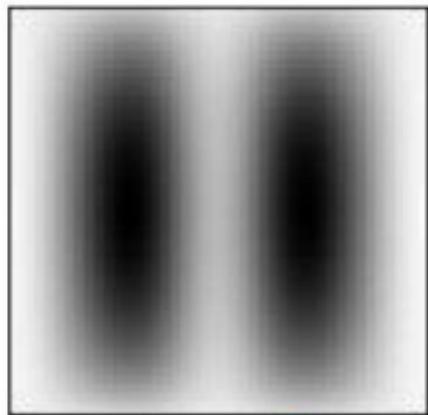
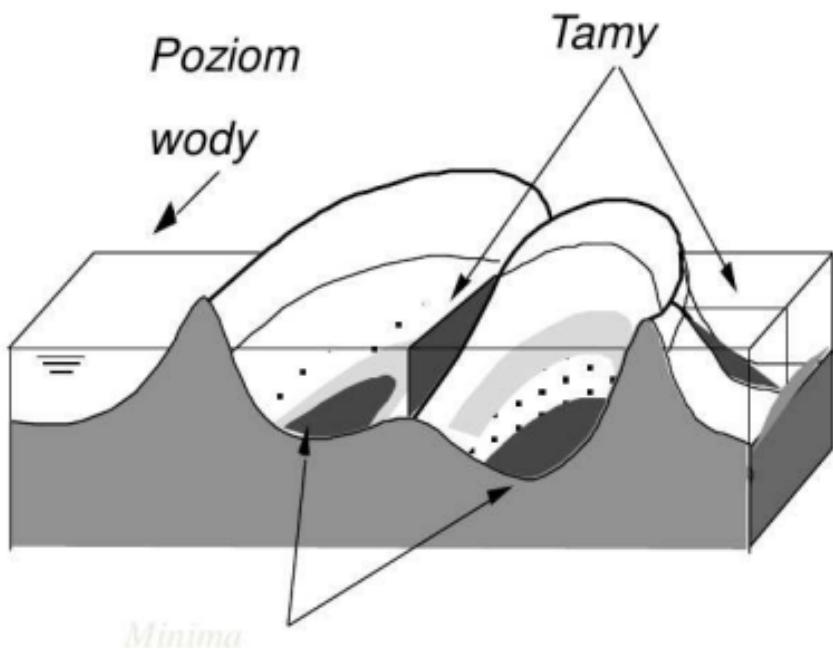
Green

Brown

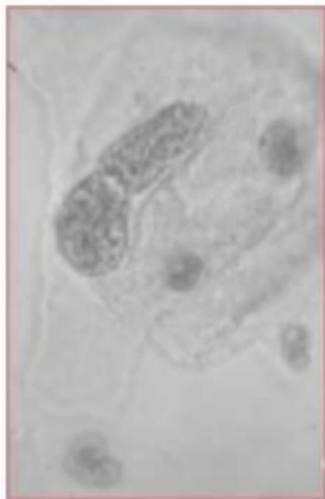
Red

Segmentacja metody zaawansowane

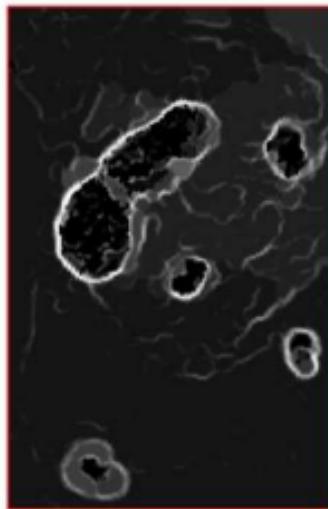
Segmentacja przez wododział



Wyniki wododziału na bazie gradientu Sobela



Obraz

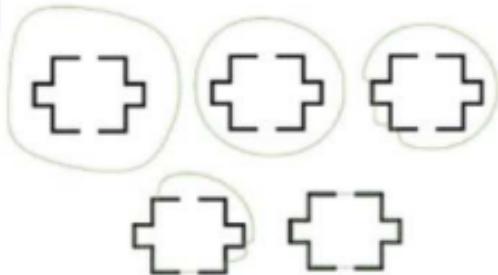


Jego gradient
operator Sobela

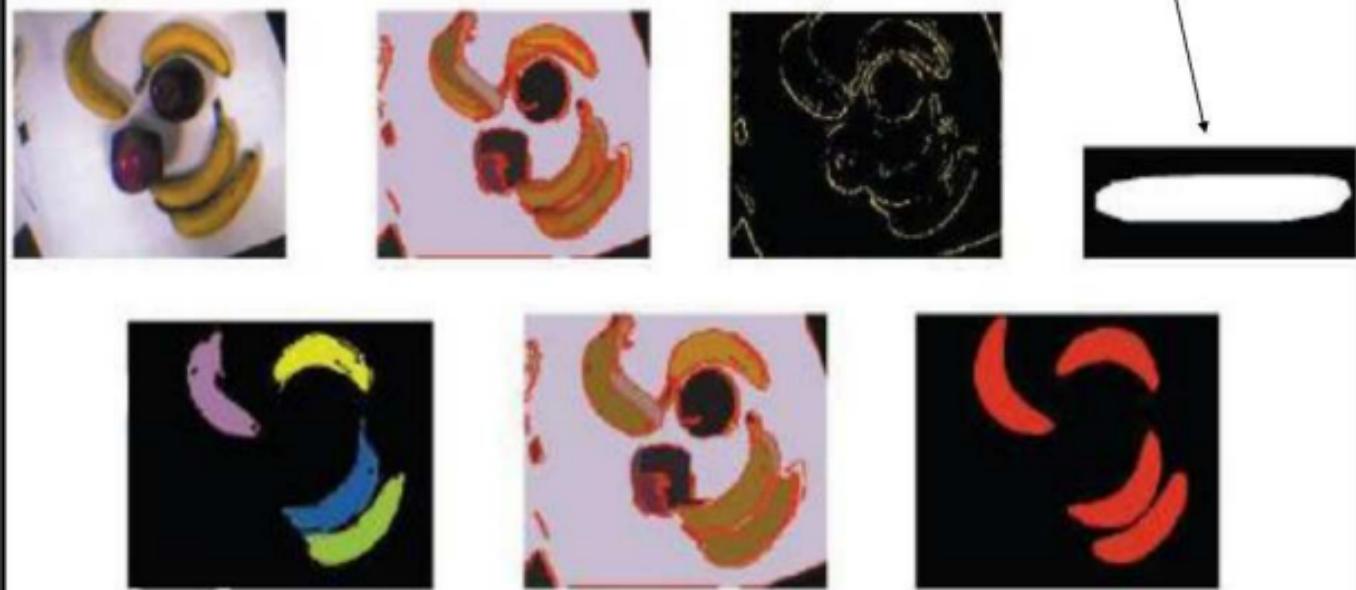


Wynik
segmentacji
przez wododział

Segmentacja metodami aktywnych konturów



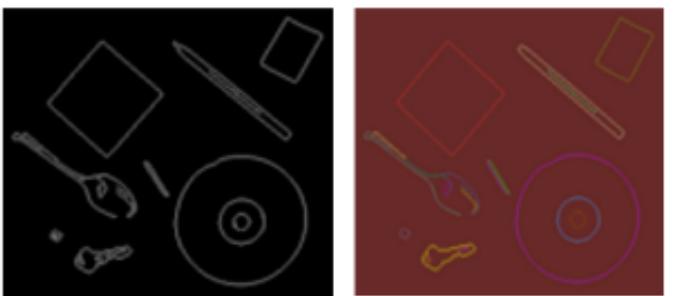
Segmentacja według modelu



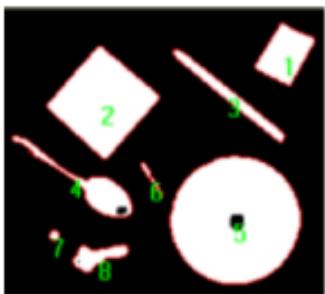
Etykietowanie obiektów

Za pomocą:

- Koloru
(pseudokoloru)



- Szarego odcienia



- Numeru



Analiza i rozpoznawanie obrazów

WYKŁAD 7

Dla studiów niestacjonarnych 2021/2022

Łukasz Roszkowiak

Jak ludzie rozpoznają obrazy?

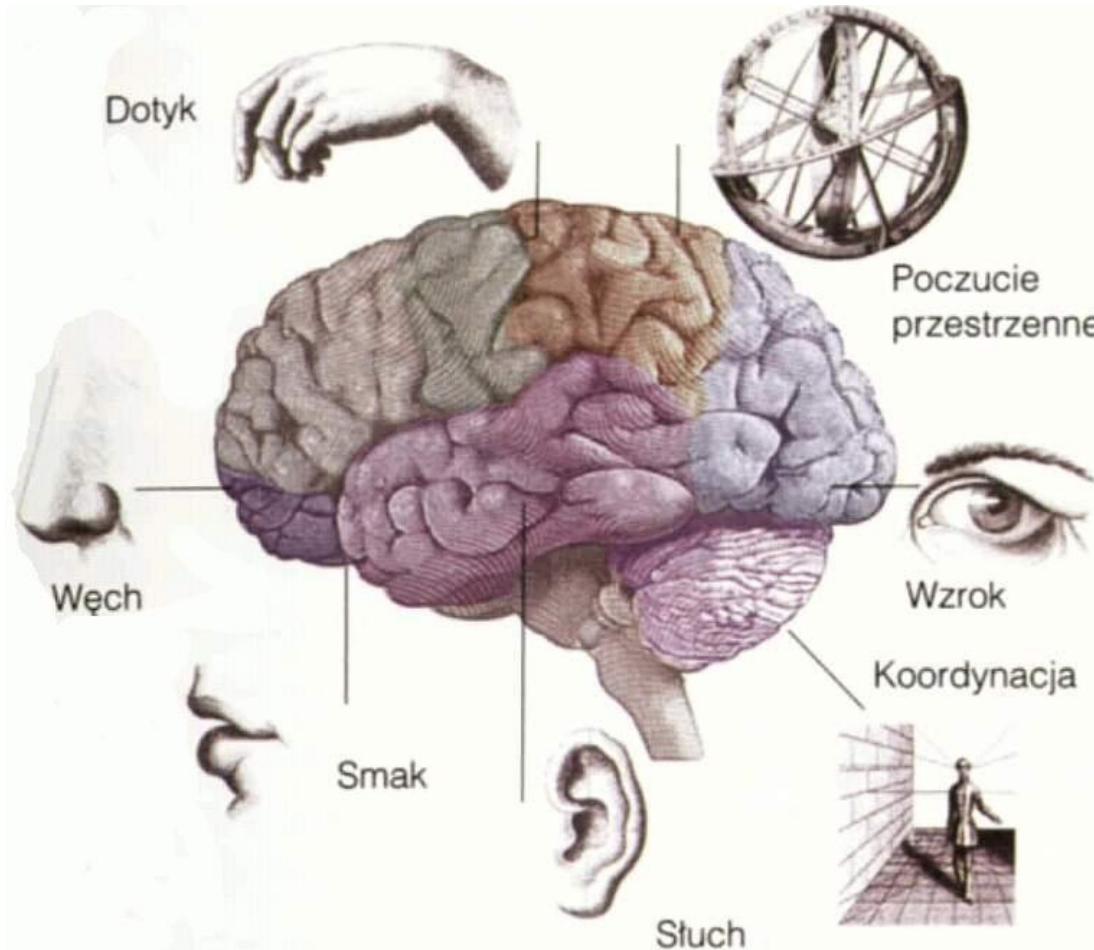
Rozpoznawanie bodźców

- **Psychofizyczny proces** zachodzący w mózgu człowieka pod wpływem **stymulacji** (np.: wzrokowej) lub nawet **jej wyobrażenia**, wywołujący **przyporządkowanie stymulacji do pewnej kategorii zjawisk, obiektów, itp..**

Ograniczamy się do bodźców wzrokowych

- W ten sposób rozpoznawanie to **interpretacja stymulacji wzrokowej** czyli **przyporządkowanie jej odpowiedniej kategorii obiektów i sytuacji** widzianych przez narząd wzroku (na podstawie poprzednich doświadczeń i znanych definicji)

Rozpoznawanie bodźców



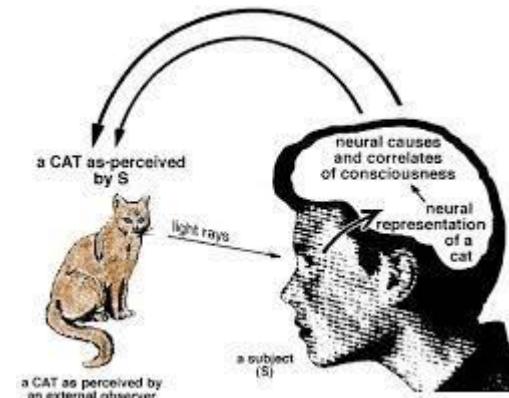
* Obraz: <https://www.is.umk.pl/~duch/Wyklady/img/03zmysly-kora.jpg>

Świadomość i wiedza

- Zdolność zdawania sobie sprawy z tego, że istnieją kategorie pojęciowe i że rozpoznanie oznacza **postrzeżenie** (reaktywność i percepcja) i **zakwalifikowanie do kategorii** czyli rozumienie
- **Niejednoznaczność obrazów**
- W postrzeganiu istnieje **progu świadomości**, czyli dolnej granicy wrażliwości na sygnały
- oraz wymagane są **doświadczenia i informacja o otaczającej rzeczywistości**

Świadomość i wiedza

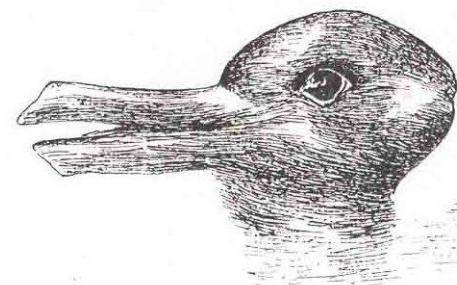
- **Zdolność zdawania sobie sprawy z tego, że istnieją kategorie pojęciowe i że rozpoznanie oznacza postrzeżenie (reaktywność i percepja) i zakwalifikowanie do kategorii czyli rozumienie**
- Niejednoznaczność obrazów
- W postrzeganiu istnieje progu świadomości, czyli dolnej granicy wrażliwości na sygnały
- Kontekst - wymagane są doświadczenia i informacja o otaczającej rzeczywistości



* Obraz: <https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSaQ9j1ldIYnyFpEesx8RCRjX9SKED4XBzFXw&usqp=CAU>

Świadomość i wiedza

- Zdolność zdawania sobie sprawy z tego, że istnieją kategorie pojęciowe i że rozpoznanie oznacza postrzeżenie (reaktywność i percepcaja) i zakwalifikowanie do kategorii czyli rozumienie
- **Niejednoznaczność obrazów**
- W postrzeganiu istnieje progu świadomości, czyli dolnej granicy wrażliwości na sygnały
- Kontekst - wymagane są doświadczenia i informacja o otaczającej rzeczywistości

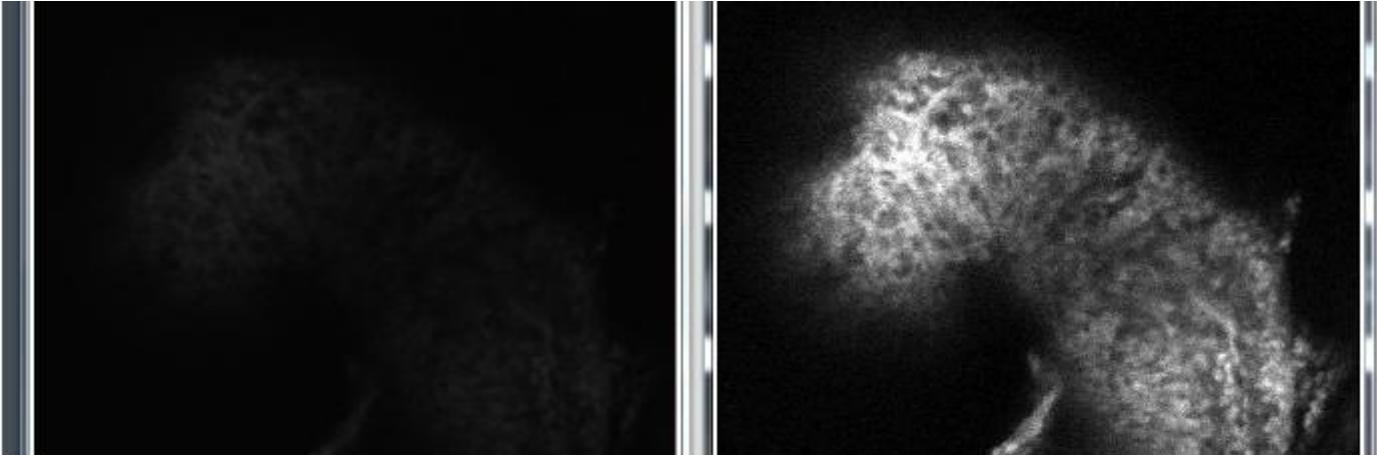


* Obraz:

<https://www.illusionsindex.org/images/illusions/duck-rabbit/duckrabbitmain.jpg>

Świadomość i wiedza

- Zdolność zauważania tego, że istnieje coś, co może być rozpoznane (reaktywność), zakwalifikowana i zrozumiana



- Niejednoznaczność
- **W postrzeganiu istnieje próg świadomości, czyli dolnej granicy wrażliwości na sygnały**
- Kontekst - wymagane są doświadczenia i informacja o otaczającej rzeczywistości

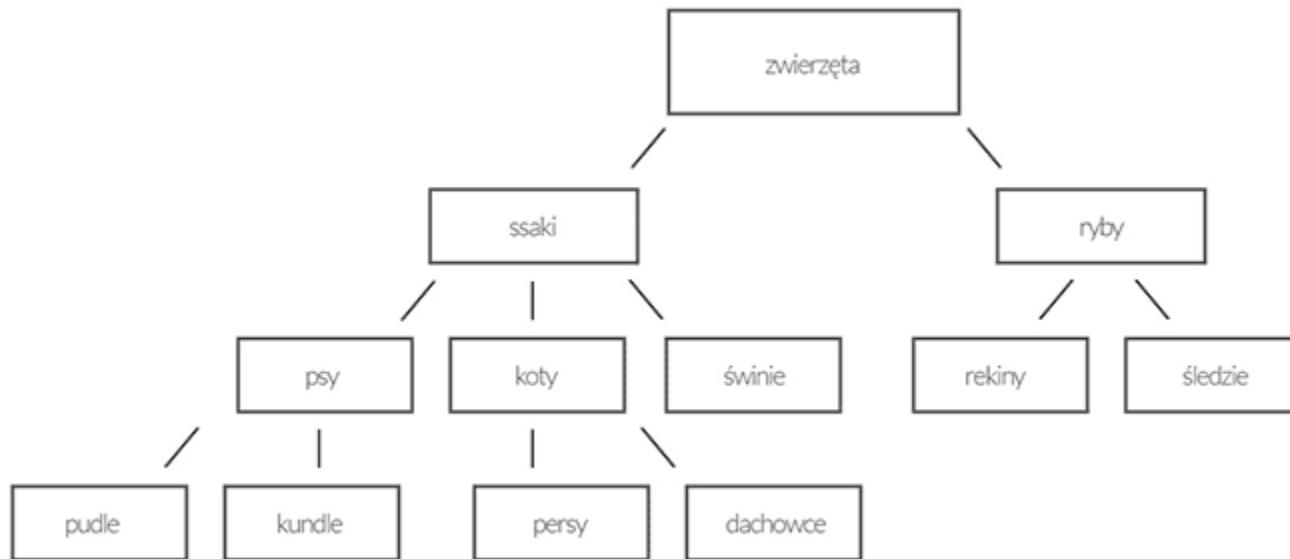
Świadomość i wiedza

- Zdolność zauważania tego, że istnieje coś, co może być rozpoznane (reaktywność) lub zakwalifikowane (rozumienie)
- Niejednoznaczność
- W postrzeganiu świadomość, doświadczenie, wrażliwości, mity, wyobrażenia
- **Kontekst - wymagane są doświadczenia i informacja o otaczającej rzeczywistości**



* Obraz: <https://www.rafalszrajnert.pl/wp-content/uploads/2020/12/percepcja-psychologia-1410x827.jpg>

Pamięć semantyczna



Rozpoznawanie

- Proces psychofizjologiczny zachodzący w mózgu na podstawie wrażenia zmysłowego lub jego wyobrażenia, przypisujący znaczenie procesom, które są „odbierane”.
- Rozpoznawanie wrażeń wzrokowych jest więc interpretacją znaczenia stymulacji wzrokowych przez zidentyfikowanie kategorii, do której należą obiekty stymulujące (odwołanie do wcześniejszych doświadczeń i wcześniej zdefiniowanych kategorii lub utworzenie nowej kategorii)

Tło gnostyczne = świadomość + wiedza

ontologie = wiedza i reguły wnioskowania

Ośrodki rozpoznawania

- Twarzy
- Znaków pisanych
- Położenia w przestrzeni
- Przedmiotów

Na podstawie niższych cech:

Linie i ich kierunki

Punkty charakterystyczne

Ruch

Kolor



Rozpoznanie

- Interpretacja stymulacji wzrokowej przez zidentyfikowanie **kategorii**, do której należą obiekty stymulujące

Impuls
stymulujący

Rozpoznanie

Kategoria

Teorie rozpoznawania

- Teoria wzorca
 - Przechowywanie nieskończonej liczby wzorców
 - Porównywanie (dokładność)
 - Klasyfikacja?



■ Teoria cech

- Dopasowywanie cech obrazu do cech przechowywanych w pamięci
- Dowody: „detektory cech” w korze wzrokowej (Hubel & Wiesel)

Porównanie percepcji obrazu

- Człowiek

- Duża szybkość
- Czasami błędna (iluzje)



- Łatwo modyfikowalne i rozszerzalna (rosnąca widza o świecie i umiejętności wnioskowania)

- Komputer

- Stosunkowo powolna dla skomplikowanych obrazów
- Niedoskonała bo ograniczona i sztywna tylko do pewnego ścisłe zdefiniowanego zakresu obrazów
- Nie uwzględnia innej wiedzy o świecie niż to założoną na etapie programowania

Rozpoznawanie obrazów

- Rozpoznawanie obrazów jest częścią dziedziny wiedzy, zwanej po angielsku *Pattern Recognition*, a po polsku rozpoznawaniem wzorca
- Jak jest realizowane przez komputery?

Komputerowe rozpoznawanie obrazów

Podejścia:

1. Klasyczne – oparte o cechy obiektu ułatwiające jednoznaczne rozpoznanie
 - Analiza obrazu zajmuje się ekstrakcją cech na potrzeby rozpoznawania obrazów
2. Oparte na sztucznej inteligencji - bazuje na cechach wyznaczonych automatycznie w procesie uczenia

Komputerowa analiza i rozpoznawanie obrazów



Jest to „sztuka” udzielania, **automatycznej** i mającej **matematyczne podstawy** odpowiedzi na pytanie:

**Co ten obraz przedstawia?
- o czym mówi ?**

Podejścia:

1. **Klasyczna** – oparta o cechy wyznaczone przez developera na podstawie rad ekspertów i badań naukowych
 - **Analiza obrazu** zajmuje się ekstrakcją cech na potrzeby rozpoznawania obrazów
 - **Rozpoznawanie obrazu** zajmuje się tworzeniem i weryfikacją reguł, na podstawie których udziela się odpowiedzi na powyższe pytanie oraz stosowaniem tych reguł w konkretnych zagadnieniach praktycznych
2. **Oparta na sztucznej inteligencji** - bazuje na cechach wyznaczonych automatycznie w procesie uczenia

Cel rozpoznawania

- Wspomaganie ludzkich decyzji za pomocą informacji obrazowej lub informacji ekstrahowanej z obrazów
- Uzależniony od zastosowania:
 - Zastąpienie człowieka
 - Weryfikacja działania manipulatorów i robotów
 - Selekcja dużych ilości danych (badania przesiewowe w diagnostyce medycznej)
 - Biometria

Komputerowe rozpoznawanie obrazów

- Naśladuje rozpoznawanie wykonywane przez człowieka choć wiedza o świecie zgromadzona w oprogramowaniu jest bardzo ograniczona
- Jest znacznie mniej efektywne
- Jest związane z zastosowaniami:
 - Zastąpieniem człowieka w trudnych warunkach (kontrola jakości mikroprocesorów), uciążliwych czynnościach (OCR)
 - Weryfikacja działania manipulatorów i robotów
 - Selekcją dużych ilości danych (badania skriningowe w diagnostyce medycznej)

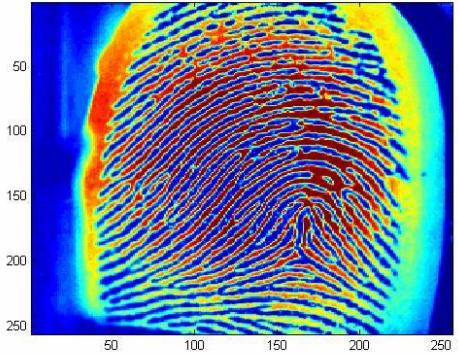
Rozpoznawanie obrazów

Cel:

Wspomaganie ludzkich decyzji za pomocą informacji obrazowej lub informacji ekstrahowanej z obrazów

Proces rozpoznawania jest wieloetapowy, zawiera dwa typy działań:

- ukierunkowane (detekcja dopasowania, analiza kształtu, pomiar wielkości lub odległości)
- nieukierunkowane (filtracja obrazu, zamiana na obraz monochromatyczny, wyodrębnianie krawędzi)



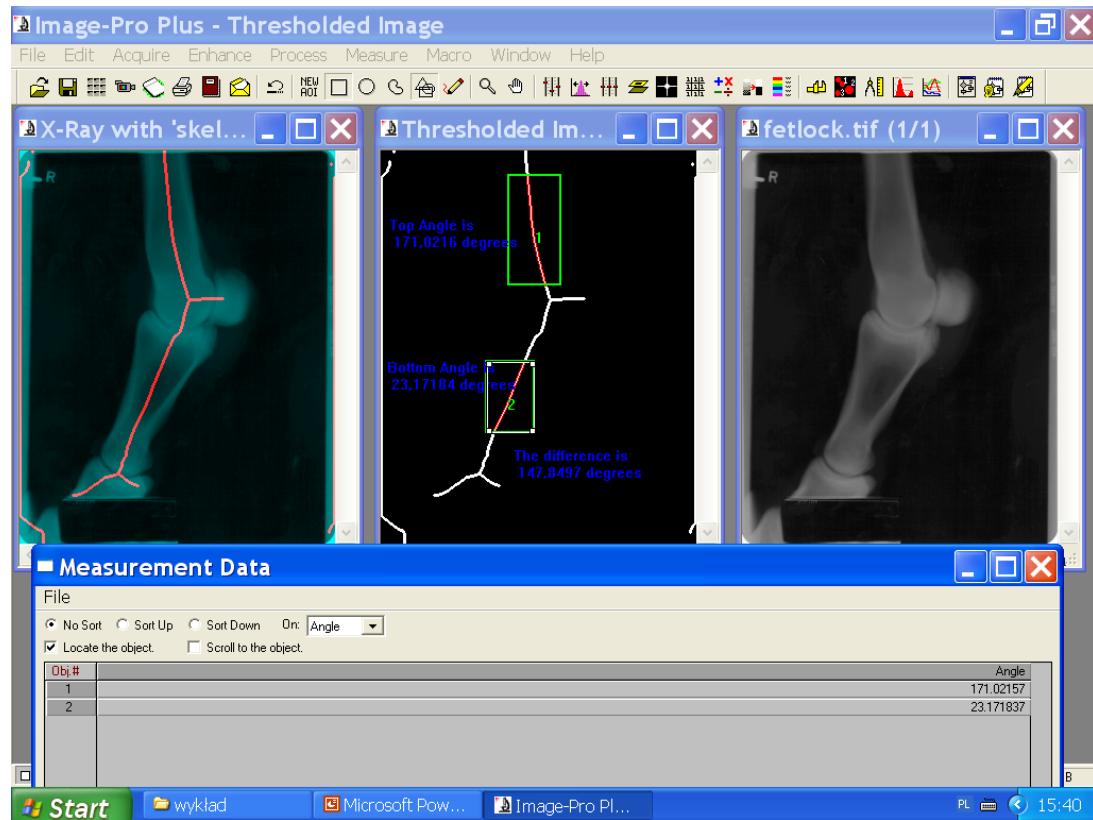
- Rozpoznawanie obrazów jest związane z innymi dziedzinami nauk komputerowych: uczeniem maszynowym UM (machine learning, ML), sztuczną inteligencją (artificial intelligence AI), komunikacją człowiek-komputer
- Zastosowania:
 - Bioidentyfikacja (oczy, uszy, odciski palców, głos)
 - Kontrola jakości produktów, kontrola samochodów na drogach (rozpoznawanie tablic rejestracyjnych), roboty i manipulatory
 - Badania przesiewowe (w diagnostyce medycznej)
 - Symulatory do nauki prowadzenia pojazdów (samolotów, pojazdów kosmicznych, samochodów wyścigowych, wieże kontrolnej lotów)
 - Rozpoznawanie twarzy
 - Marketing (Yamaha Motor)

Przykład (klasyczne)

Jak ocenić zdolność konia do biegania w wyścigach na podstawie zdjęcia rentgenowskiego jego stawu skokowego?

Podejście klasyczne

Analiza obrazów jest związana z ilościowym opisem danych zawartych na obrazie tzn. pomiarem rozmiaru, kształtu, koloru, ustaleniem stosunków pomiędzy zidentyfikowanymi elementami obrazu lub zbadaniem dyspersji elementów w przestrzeni.



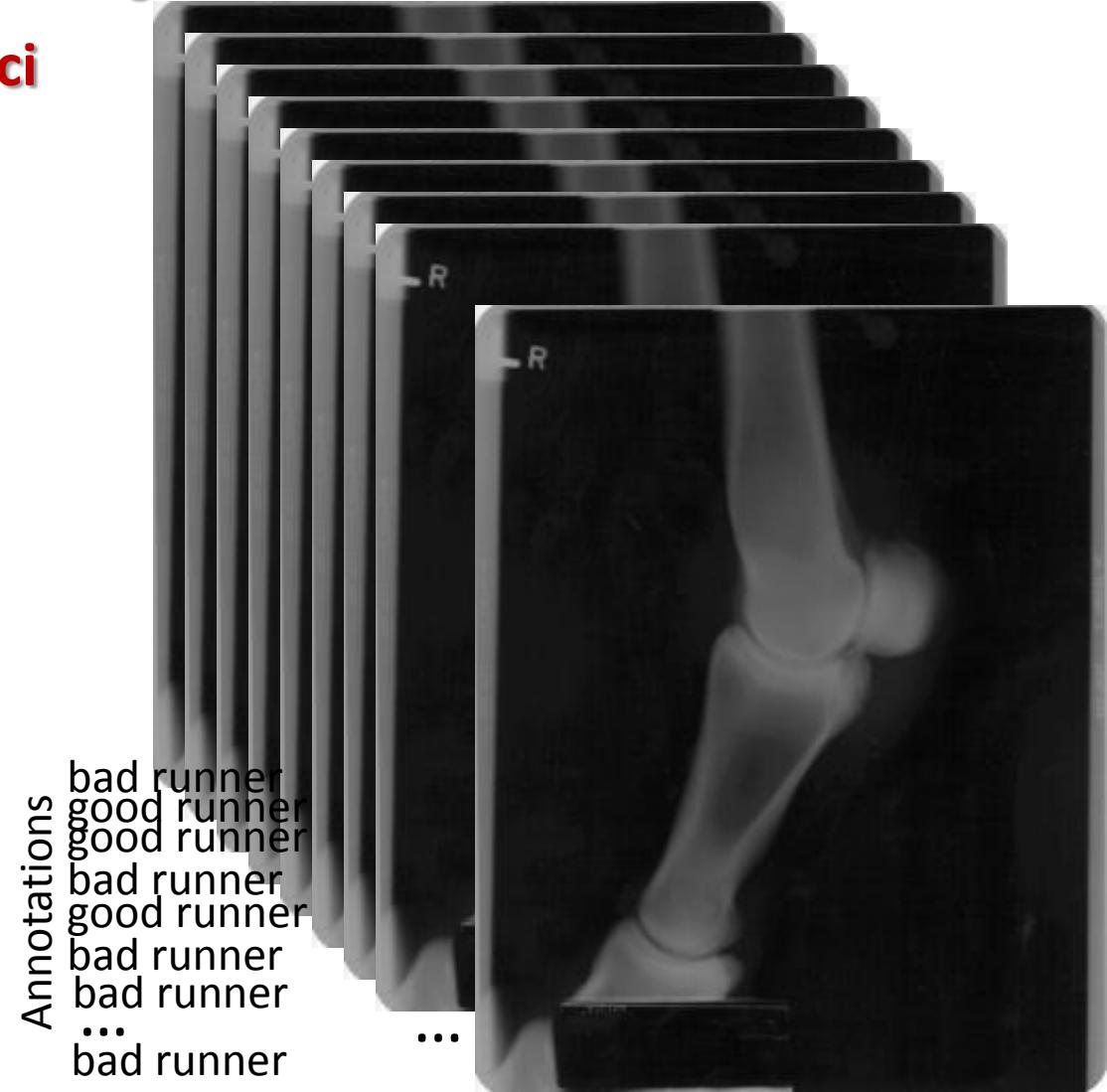
W procesie analizy na podstawie geometrycznych cech przestrzeni i obiektów na obrazie, ekstrahujemy „znaczenie/zawartość informacyjna” obrazu.

Przykład (AI)

Jak ocenić zdolność konia do biegania w wyścigach na podstawie zdjęcia rentgenowskiego jego stawu skokowego?

Podejście na podstawie Sieci Neuronowych (neural networks NN)

Mając duży zestaw przykładów zdjęć rentgenowskich kostek koni ze znajomością zdolności konia do biegania (w formie adnotacji), możliwe jest prezentowanie wszystkich danych w sieci neuronowej (NN) i uruchamianie narzędzi w celu zoptymalizowania odpowiedzi NN na pytanie.

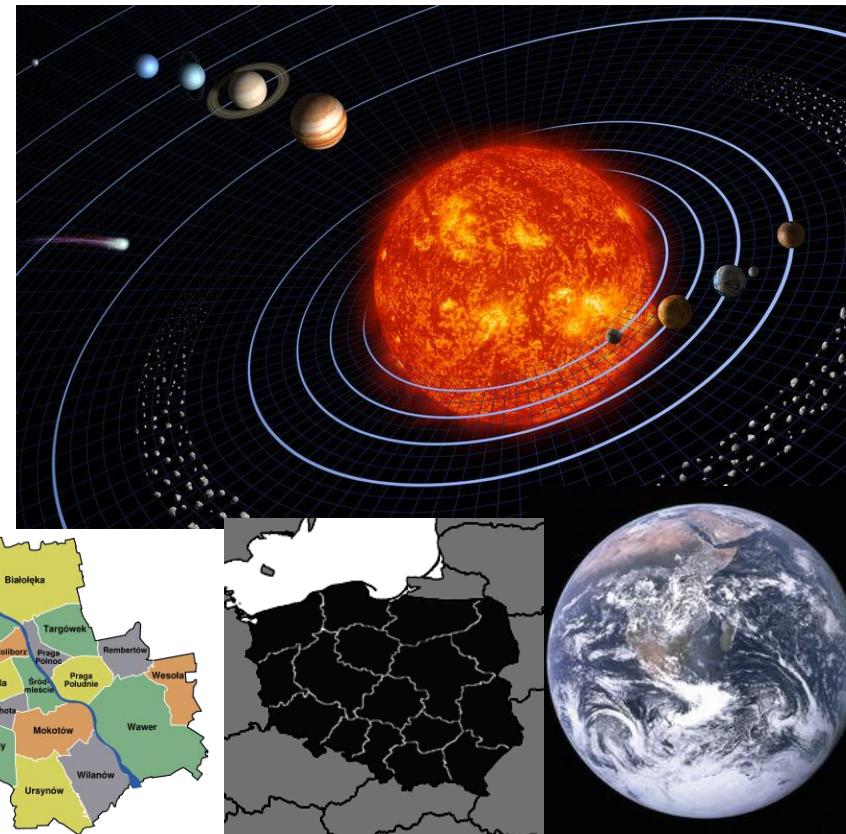


Podsumowując

Rozpoznawanie obrazów jest częścią dziedziny wiedzy, zwanej po angielsku **Pattern Recognition**, czyli po polsku rozpoznawaniem postaci lub wzorca

Świat może być obserwowany jako zorganizowany/zbudowany ze struktur i jego zasadniczą cechą jest organizacja i relacje między jego elementami, a nie ich wewnętrzna natura

(mogą być scharakteryzowane jako sposób ułożenia/relacje elementów, z których się składają)



Struktury mogą być opisane w sposób:

- ilościowy: przez pomiar długości, pola powierzchni, tekstury, itp
- jakościowy: przez opis relacji między elementami, itp

Podsumowanie – komputerowego rozpoznawania obrazów

- Złożony i wieloetapowy proces
- Naśladuje rozpoznawanie wykonywane przez człowieka choć wiedza o świecie zgromadzona w oprogramowaniu jest bardzo ograniczona
- Podejście klasyczne
- Podejście sztucznej inteligencji

Klasyczne rozpoznawanie obrazów

Składa się z dwóch etapów:

analizy

i

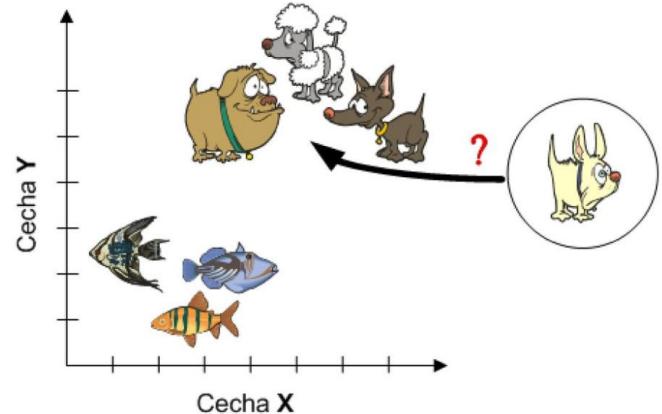
właściwego rozpoznawania

Analiza obrazu

realizacja odwzorowania:

$$B : D \rightarrow X$$

- D - przestrzeń obrazów,
- X - przestrzeń wektorów cech
- B – odwzorowanie



1. wyznaczenie **cech obiektów** (wyodrębnionych uprzednio w procesie segmentacji) przydatnych w procesie właściwego rozpoznawania;
2. cechy charakteryzujące kształt; współczynniki **niezmienne** względem typowych przekształceń obrazów (obroty, przesunięcia, zmiany skali)
 - współczynniki kształtu i momenty geometryczne
 - parametry opisu populacji obiektów
 - istnienie lub brak istotnych elementów/struktur/obiektów .

Rozpoznanie obrazu

– realizacja odwzorowań

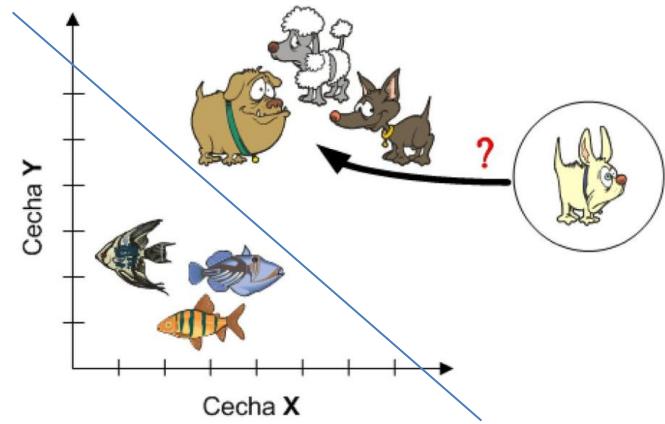
$$C : X \rightarrow R^L \quad F : R^L \rightarrow I$$

C - ustalenie *miary podobieństwa (dopasowania, odległości)* nieznanego obiektu

opisanego wektorem cech $x \in X$ do każdej z klas L

F - ustalenie ostatecznej decyzji o przynależności obiektu d opisanego

wektorem cech x do klasy $i \in I$ dla której miara podobieństwa jest maksymalna.



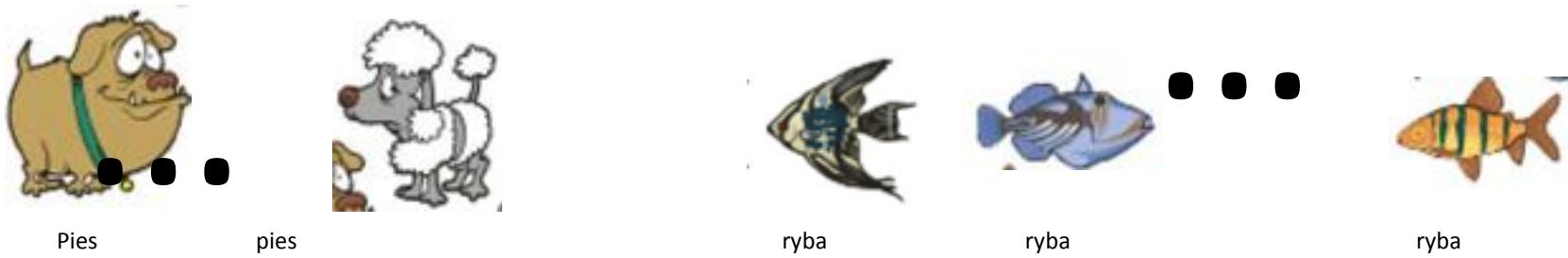
Efekt rozpoznania

automatyczna identyfikacja klasy, do której można zaliczyć nieznany obiekt (np. obraz).

Rozpoznanie jest procesem uogólnienia: przedstawiamy zbiór uczący, czyli przykłady znanych rozwiązań przynależności do klasy, a on w procesie rozpoznawania klasyfikuje również nieznane obiekty

Jak powstaje klasyczny system rozpoznawania obrazów?

1. Istnieje potrzeba dzielenia pewnych obiektów przedstawionych na cyfrowych zdjęciach na zadaną liczbę klas
2. Wymagany jest **dostępu do zbioru obrazów uczących** przedstawiających obiekty co do których nie ma wątpliwości do której klasy należą (mamy decyzje eksperta).

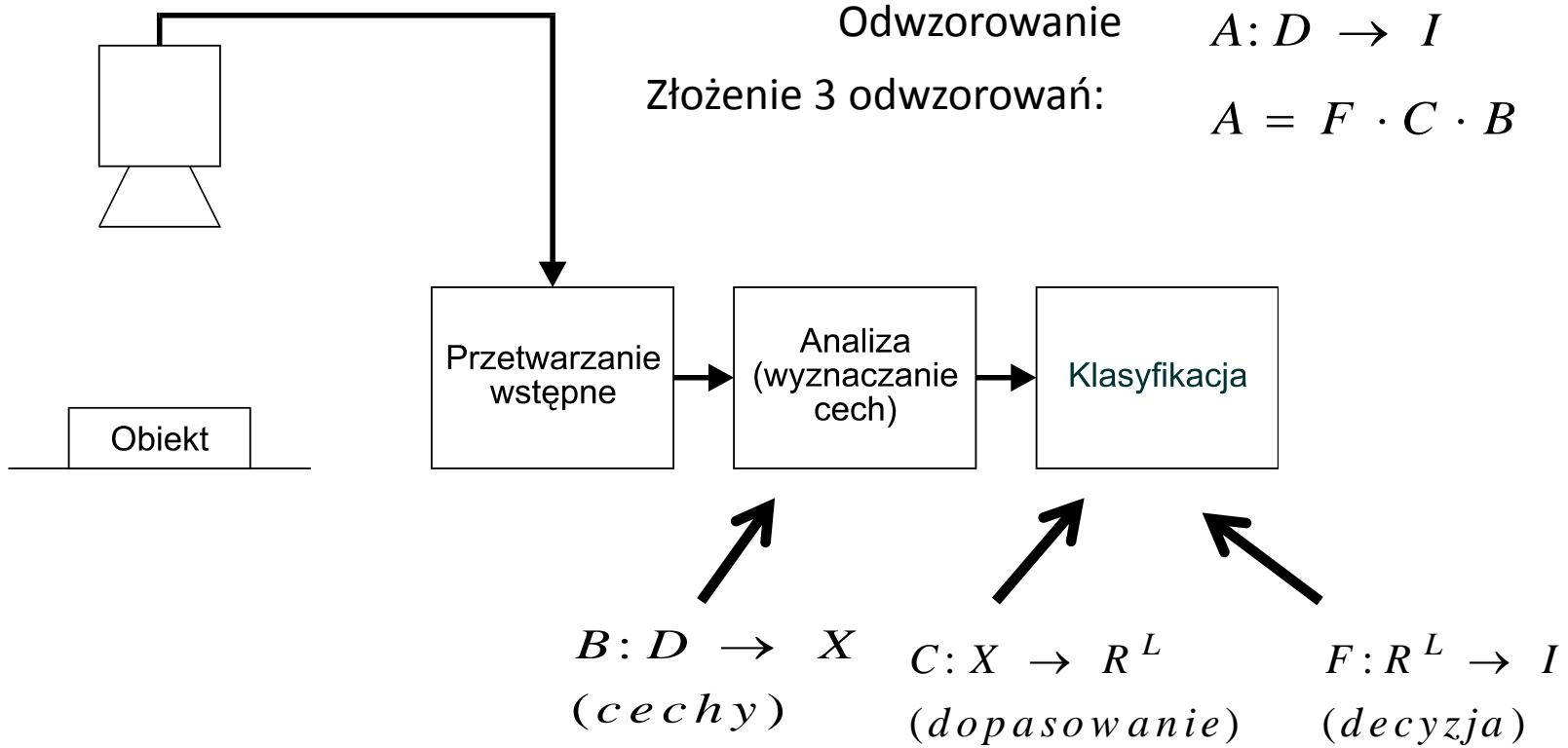


Etap wstępny: wybieramy cechy i sprawdzamy ich możliwości dyskryminacyjne

przygotowanie systemu (GUI, BD, itp.)

Etap zasadniczy: wykorzystanie systemu

ANALIZA I ROZPOZNANIE OBRAZU w klasycznym ujęciu



D - przestrzeń obrazów,

X - przestrzeń cech,

Cⁱ(x) - funkcja przynależności (miara dopasowania x do i -tej klasy),

R^L - L liczb rzeczywistych,

I - zbiór indeksów klas. tzn. $i \in I$

Matematyczne podstawy

Metodologia rozpoznawania w klasycznym podejściu

- Deterministyczna (dokładna znajomość wzorców – np. liter)
- Stochastyczna (nie istnieje jednoznaczny wzorzec; niejednoznaczność powoduje, że zakłada się pewien poziom błędu w klasyfikacji)
- Korelacyjna
- Lingwistyczna

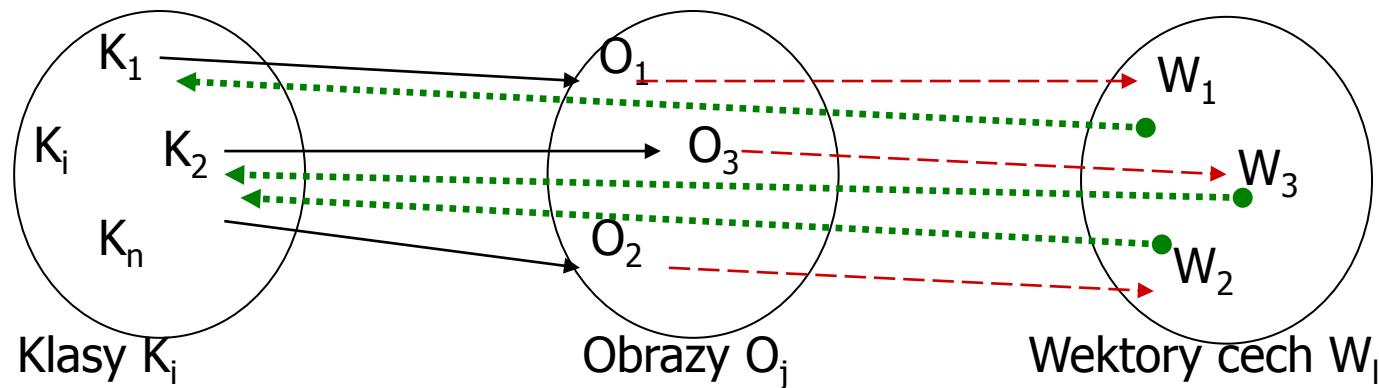
Case study:

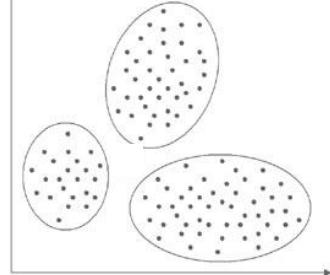
- Porównywanie z wzorcem
- k-NN (k-najbliższych sąsiadów)

Etapy procesu rozpoznawania

Faza wstępna

Ekstrakcja cech charakterystycznych dla danego obiektu w sensie celu rozpoznawania, konstrukcja klasyfikatora na podstawie obrazów za zbioru uczącego (z lub bez nauczyciela)





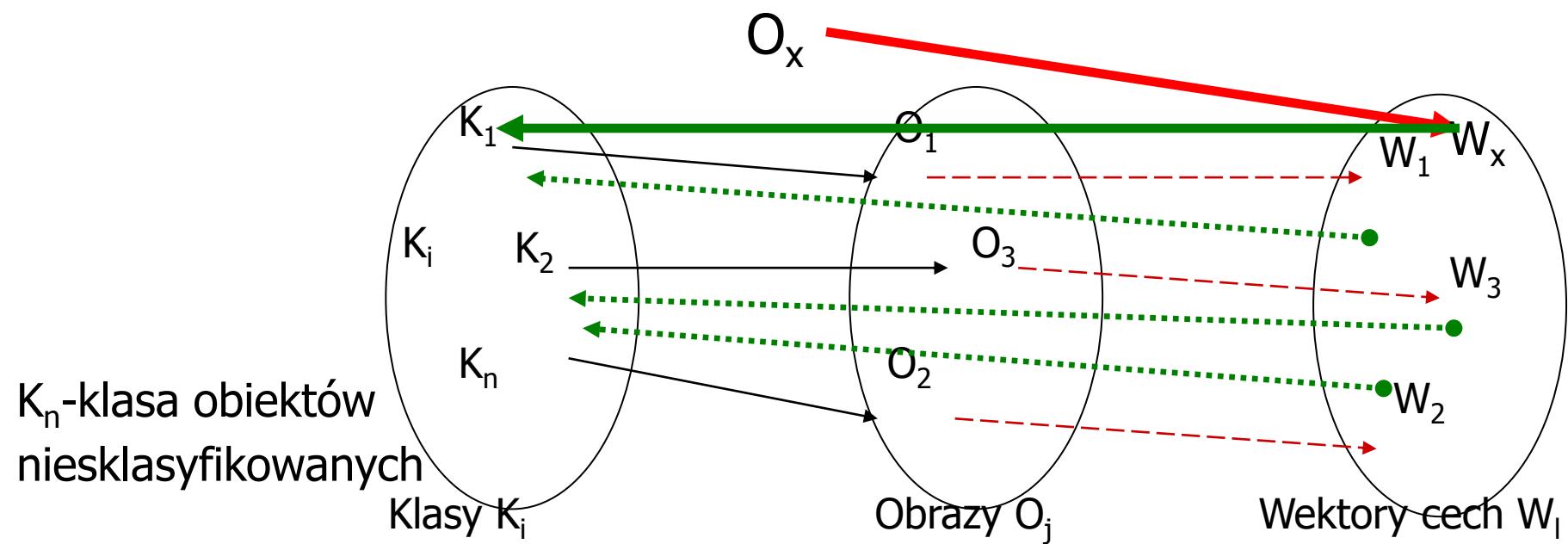
Dobór cech i konstrukcja klasyfikatora - faza wstępna

- Pierwszy, najtrudniejszy etap rozpoznawania - **wybór cech dla klasyfikatora** - wymaga doświadczenia, wyczucia i pracy na pograniczu heurezy
- Drugi etap konstrukcji klasyfikatora jest związany z wyborem jednego z formalnie **(matematycznie) opracowanego modelu rozpoznawania**

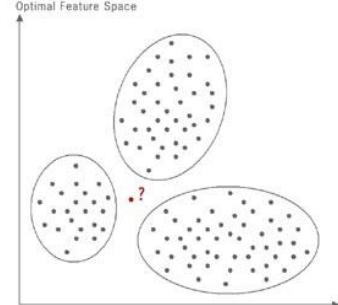
Etapy procesu rozpoznawania

Faza zasadnicza

Klasyfikacja nowych obiektów/obrazów, dokonywana na podstawie tych cech



Właściwe rozpoznawanie obrazów



Jeśli dobrze dobrzemy cechy do zadanego podziału na klasy i dobrze skonstruujemy klasyfikator, to:

- W ramach analizy obrazów dokonujemy **ekstrakcji wybranych cech**
- Ustalamy **podobieństwo (dopasowanie)** nieznanego obiektu, opisanego za pomocą wektora cech, do każdej z klas
- Podjęcie ostatecznej **decyzji o przynależności obiektu do jednej klasy**, dla której miara podobieństwa jest maksymalna.

Analiza obrazu

redukcja obrazu do **punktu** w n -wymiarowej przestrzeni lub do **wektora cech \underline{x}** w n -wymiarowej **przestrzeni cech X** :

$$\underline{x} = \begin{bmatrix} x_1 \\ | \\ \cdot \\ | \\ \cdot \\ | \\ x_n \end{bmatrix}; \quad x \in X$$

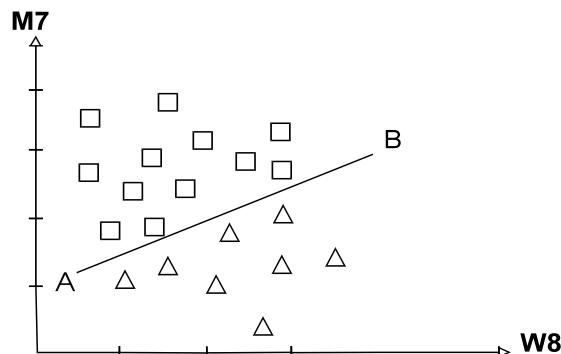
Przykład

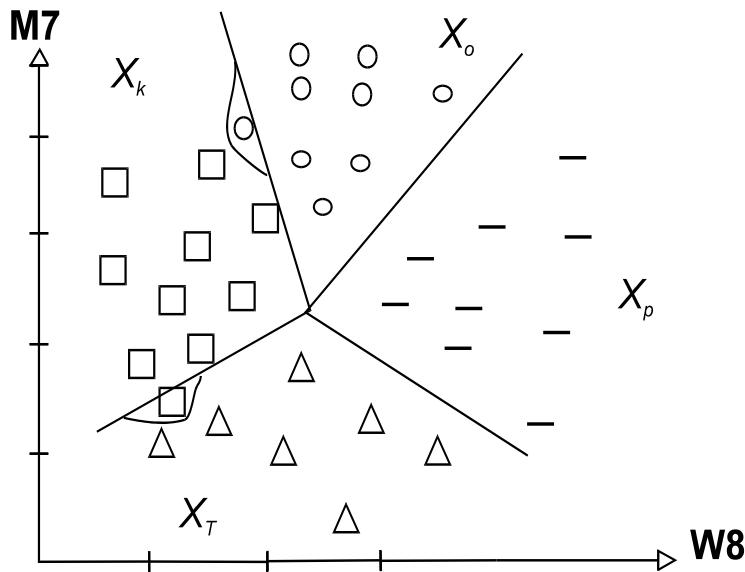
wektor w 2-wymiarowej przestrzeni cech.

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

gdzie

x_1 - wartość współczynnika kształtu (np.: W8)
 x_2 - wartość momentu np. M7





$$\underline{x}_k = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{kwadrat}$$

$$\underline{x}_T = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{trójkąt}$$

$$\underline{x}_p = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{prostokąt}$$

$$\underline{x}_o = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{okrąg}$$

- Idealny podział: taki, że wszystkie wektory cech znajdują się w odpowiadających im obszarach.
- Jeśli jest to niemożliwe - podział **minimalizujący prawdopodobieństwo błędu** (błędnej decyzji), lub podział **minimalizujący błąd średni**.
- -----
- **Algorytmy rozpoznawania** przynależności obiektów do pewnych klas wykorzystują **ciąg uczący** złożony z obiektów, dla których **znana jest** prawidłowa klasyfikacja
- Efektywność rozpoznania zależy od właściwego **doboru cech obiektów** w procesie analizy.
- Miary tej efektywności: wartości **prawdopodobieństwa błędu** (błędnej decyzji) lub **błędu średniego** powstające w trakcie podziału przestrzeni cech.

Recepcja i struktura przestrzeni cech

$B : D \rightarrow X$ zamiana obiektów $d \in D$ w punkty przestrzeni cech, recepcja (przyjmowanie) obrazów do X , czyli do przestrzeni cech.

Elementami przestrzeni cech X są wektory o n współrzędnych (składowych):

$$\underline{x} = [x_1, x_2, \dots, x_n]^T \in X$$

Składowe x_ν tych wektorów - liczby $x_\nu \in R$ z których każda określa **pewną** miarę danej cechy;

stąd:

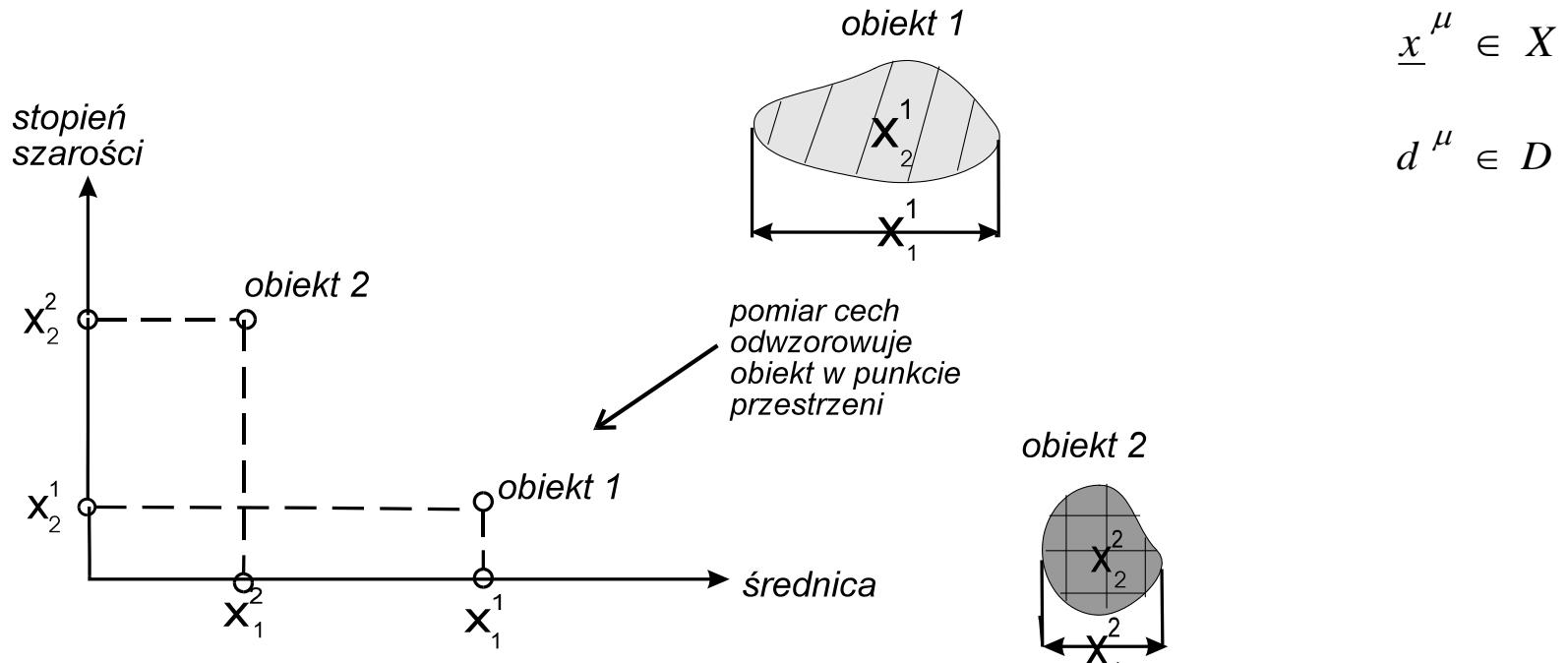
X - n -wymiarowa przestrzeń, np. Euklidesowa

czyli:

$$(X \subseteq R^n)$$

Przykład dla dwuwymiarowej (2D) przestrzeni cech

cechy obiektów są to **współrzędne przestrzeni X** ; odwzorowanie obiektu w punkt x_1^μ, x_2^μ charakteryzowany przez współrzędne



W tym przykładzie zdefiniowano cechy **ilościowe** (opisane za pomocą liczb rzeczywistych)

Obiekty $d \in D$ potencjalnie posiadają **nieskończenie wiele** cech

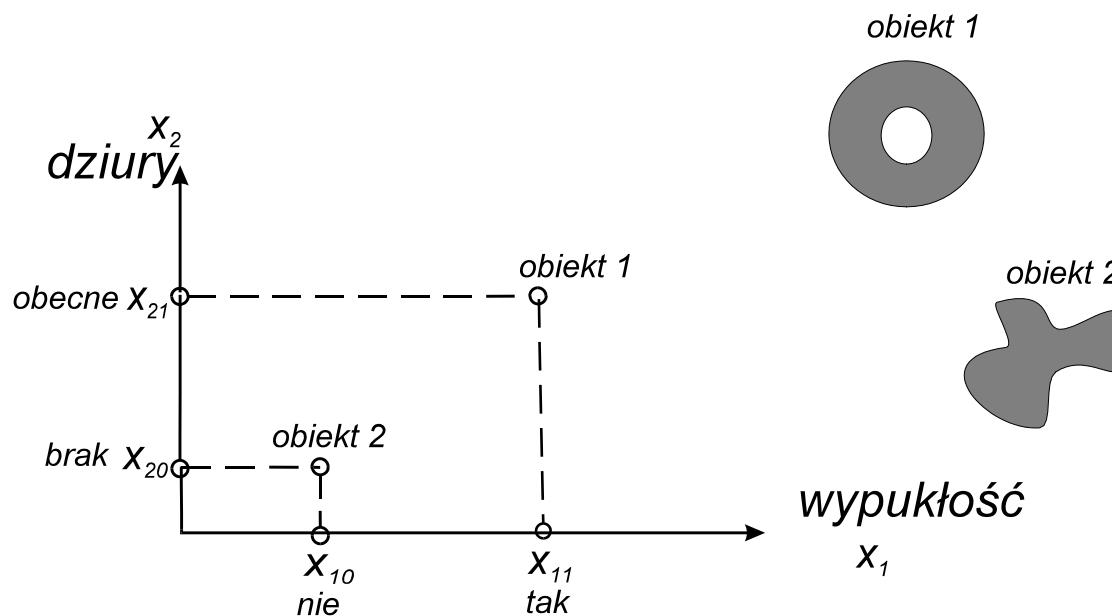
Zasada Brawermannna

Taki dobór cech x_v , aby w przestrzeni X punkty \underline{x} odpowiadające obiektom d należącym do jednej klasy ($d \in D_i$) grupowały się w postaci skupisk (ang. clusters) możliwie maksymalnie **zwartych** wewnętrznie i możliwie najbardziej **oddalonych** od podobnych skupisk dla innych klas.

Przykład
cechy **binarne**

$$\text{obiekt 1: } \underline{x}^1 = [x_1^1, x_2^1]^T$$

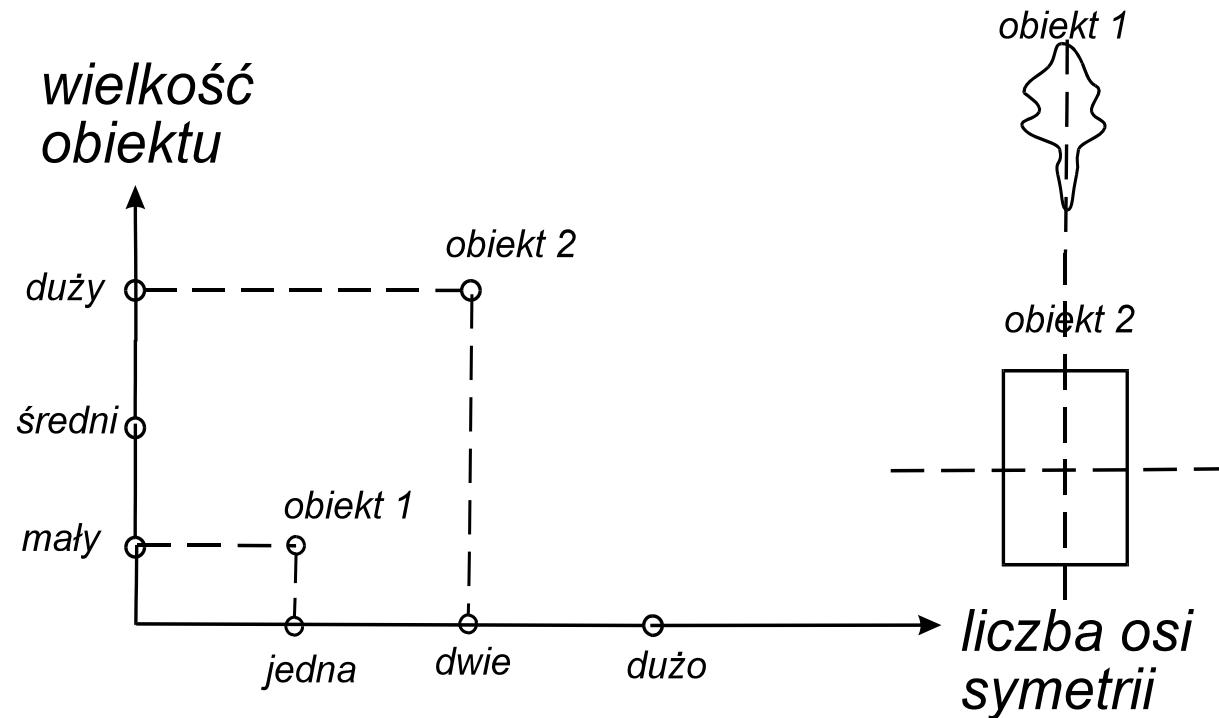
$$\text{obiekt 2: } \underline{x}^2 = [x_1^2, x_2^2]^T$$



Automatyczny wybór
(dobór) cech - **systemy
ekspertowe**

Przykład

Cechy jako **kody** opisujące właściwości obiektów



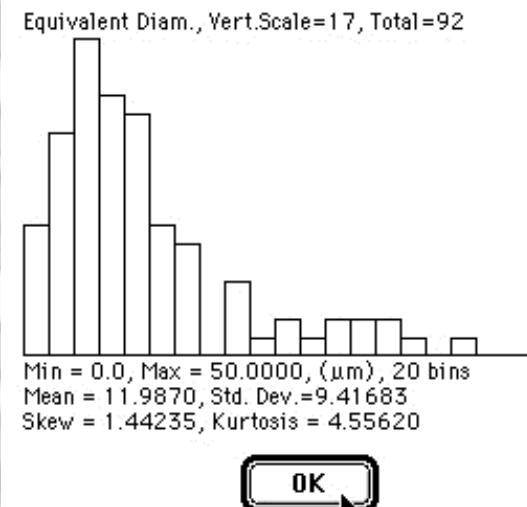
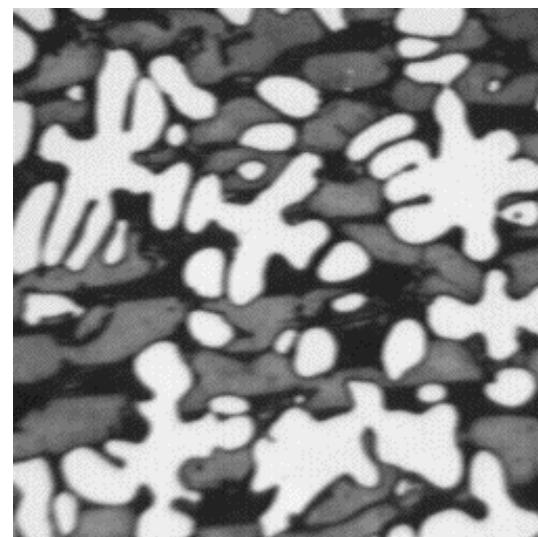
Analiza kształtu

Aby rozpoznać obraz potrzebna jest analiza

Analiza prowadzi do redukcji informacji opisującej obraz do informacji istotnej z punktu widzenia celu

- Ilościowe

- Intensywność
- Odległość
- Rozmiary i wielkości
(pole powierzchni, obwód)
- Rozmiar fraktalny
- Harmoniczne



- Jakościowe

- Istnienie wzorców/struktur i symboli
- Lokalizacja bezwzględna lub wzajemna wzorców/struktur

Charakterystyczne cechy obiektów

Analiza kształtu

wskaźniki prezentujące **niezmienność, inwariantność względem obrotów, przesunięć, zmiany skali**

Ilość obiektów (lista obiektów)

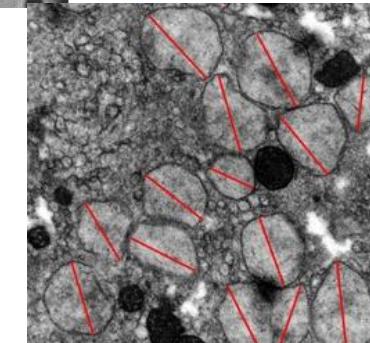
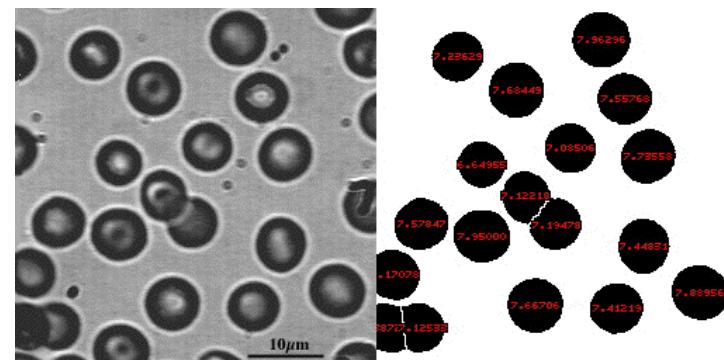
policzenie przez etykietowanie

Pole powierzchni

zliczenie pikseli w obiekcie

Osie i długości rzutów

Wzajemne położenia (drzewko opisu relacji przestrzennych)



Analiza kształtu

Współczynniki kształtu

Liczone na podstawie pola powierzchni S i obwodu L obiektu stanowią zgrubne przybliżenie kształtu

Momenty geometryczne

Pozwalają na lepsze rozróżnienie obiektów niż współczynniki kształtu, ale wymagają dłuższych obliczeń

Ani współczynnik kształtu ani moment nie mogą być użyte jako jedyna miara opisująca kształt obiektów (rozpoznanie byłoby niejednoznaczne)

Właściwości współczynników kształtu:

- zbliżone wartości W dla obiektów o zbliżonym kształcie pozwalają określić stopień podobieństwa nieznanego obiektu do poszczególnych znanych klas,
- identyczne kształty - identyczne wartości W .

Wady współczynników kształtu:

- duże zmiany skali mogą powodować, że współczynniki W dla różnych wielkości tego samego obiektu różnią się między sobą. Pojawia się wtedy możliwość błędnego zakwalifikowania do innej klasy, np. prostokąta do klasy „koło” lub odwrotnie.

Momenty geometryczne:

Pozwalają na lepsze rozróżnienie obiektów niż współczynniki kształtu, ale wymagają dłuższych obliczeń

Współczynniki kształtu

$$W1 = 2\sqrt{\frac{S}{\pi}}$$

Współczynniki cyrkularności (W1,W2):
W1 określa średnicę koła o równej **powierzchni badanego obiektu**

$$W2 = \frac{L}{\pi}$$

W2 określa średnicę koła o długości obwodu równej **długości obwodu badanego obiektu**
L - obwód obiektu, S - powierzchnia obiektu

$$W3 = \frac{L}{2\sqrt{S \cdot \pi}} - 1$$

Współczynnik Malinowskiej

Współczynniki W1,2,3 - szybkie obliczanie

$$W4 = \frac{S}{\sqrt{2\pi \iint_S (r^2) ds}}$$

Współczynnik Blaira-Blissa

(większa wrażliwość na zmiany kształtu);

r – odległość elementu pola ds od środka ciężkości obiektu

$$W5 = \frac{S^3}{\left(\iint_S l ds \right)^2}$$

Współczynnik Danielssona

l - minimalna odległość elementu ds od konturu obiektu

$$W6 = \sqrt{\frac{(\sum d)^2}{n \sum d^2 - 1}}$$

Współczynnik Haralicka

d - odległość pikseli konturu od jego środka ciężkości

n - liczba punktów konturu.

Współczynniki $W4,5,6$ - wolniejsze obliczanie niż $W1,2,3$

$$W7 = \frac{r_{\min}}{R_{\max}}$$

Współczynnik Lp1;

r_{\min} - minimalna odległość konturu od środka ciężkości

R_{\max} - maksymalna odległość konturu od środka ciężkości

$$W8 = \frac{L_{\max}}{L}$$

Współczynnik Lp2

W7, W8 określają cechy pośrednie

$$W9 = \frac{2\sqrt{\pi \cdot S}}{L}$$

Współczynnik Mz

(Malinowskiej uproszczony)

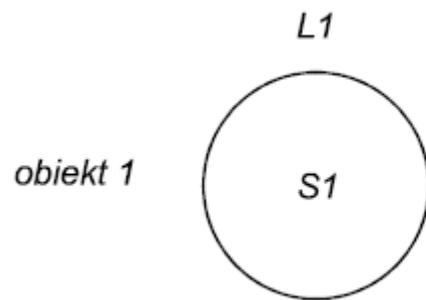
L_{\max} - maksymalny gabaryt obiektu

Przykład

wyznaczenie $W1$ i $W2$ dla wybranych obiektów

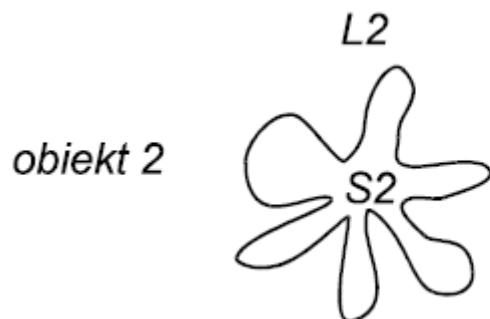
$1W1, 2W1$ - współczynnik $W1$ liczony odpowiednio dla obiektu 1 i 2

$1W2, 2W2$ - współczynnik $W2$ liczony odpowiednio dla obiektu 1 i 2



$$1W1 = 2\sqrt{\frac{S_1}{\pi}} = 2r_1 \quad 1W2 = \frac{L_1}{\pi} = 2r_1$$

$$S_1 \rangle S_2, \quad r_1 \rangle r_2 \quad L_2 \rangle L_1, \quad r_2 \rangle r_1$$



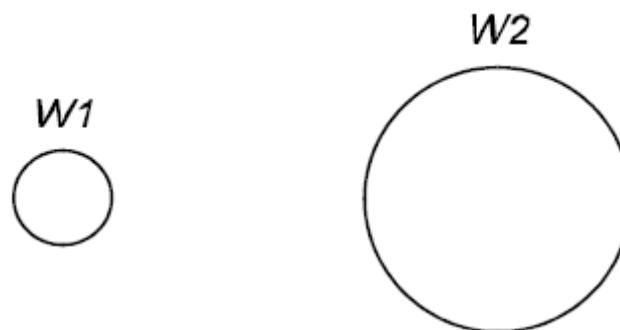
$$2W1 = 2\sqrt{\frac{S_2}{\pi}} = 2r_2 \quad 2W2 = \frac{L_2}{\pi} = 2r_2$$

$$1W1 \rangle 2W1, \quad 2W2 \rangle 1W2$$

W1 wyznacza średnicę koła
($2r_1$ lub $2r_2$) o **powierzchni**
równej powierzchni
analizowanego obiektu

W1 wyznacza średnicę koła
($2r_1$ lub $2r_2$) o **obwodzie**
równym powierzchni
analizowanego obiektu

Zobrazowanie $W1$ i $W2$
dla obiektu 2:



Problem:
zobrazować $W1$ i $W2$ dla obiektu 1

Przykład

Wyznaczenie $W7$ i $W8$ ($Lp1, Lp2$) dla obiektów 1 i 2

$$W7 = \frac{r_{\min}}{R_{\max}}$$

Obiekt 1: $W7 \cong 1, W8 < 1$

$$W8 = \frac{L_{\max}}{L}$$

Obiekt 2: $W7 < 1, W8 << 1$

Współczynniki kształtu $W1, \dots, W9$ stanowią **skalarną miarę kształtu** analizowanego obiektu.

Wyznaczone współczynniki kształtu służą do utworzenia **wektora cech**

Przykład

dla warunku, aby wartości metryki euklidesowej i ulicznej (Manhattan) były takie same.

Wektor cech obrazu 1:

$$\underline{x} = [x_1^1, \dots, x_2^1]^T = [W3, W8]^T = [2, 4]^T$$

Wektor cech obrazu 2:

$$x = [x_1^2, \dots, x_2^2]^T = [W3, W8]^T = [3, 4]^T$$

Metryka euklidesowa:

$$\rho_1(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{v=1}^n (x_v^\mu - x_v^\eta)^2}$$

$$\rho_1(\underline{x}^1, \underline{x}^2) = \sqrt{\sum_{v=1}^2 (x_v^1 - x_v^2)^2} = \sqrt{1+0} = 1$$

Metryka uliczna:

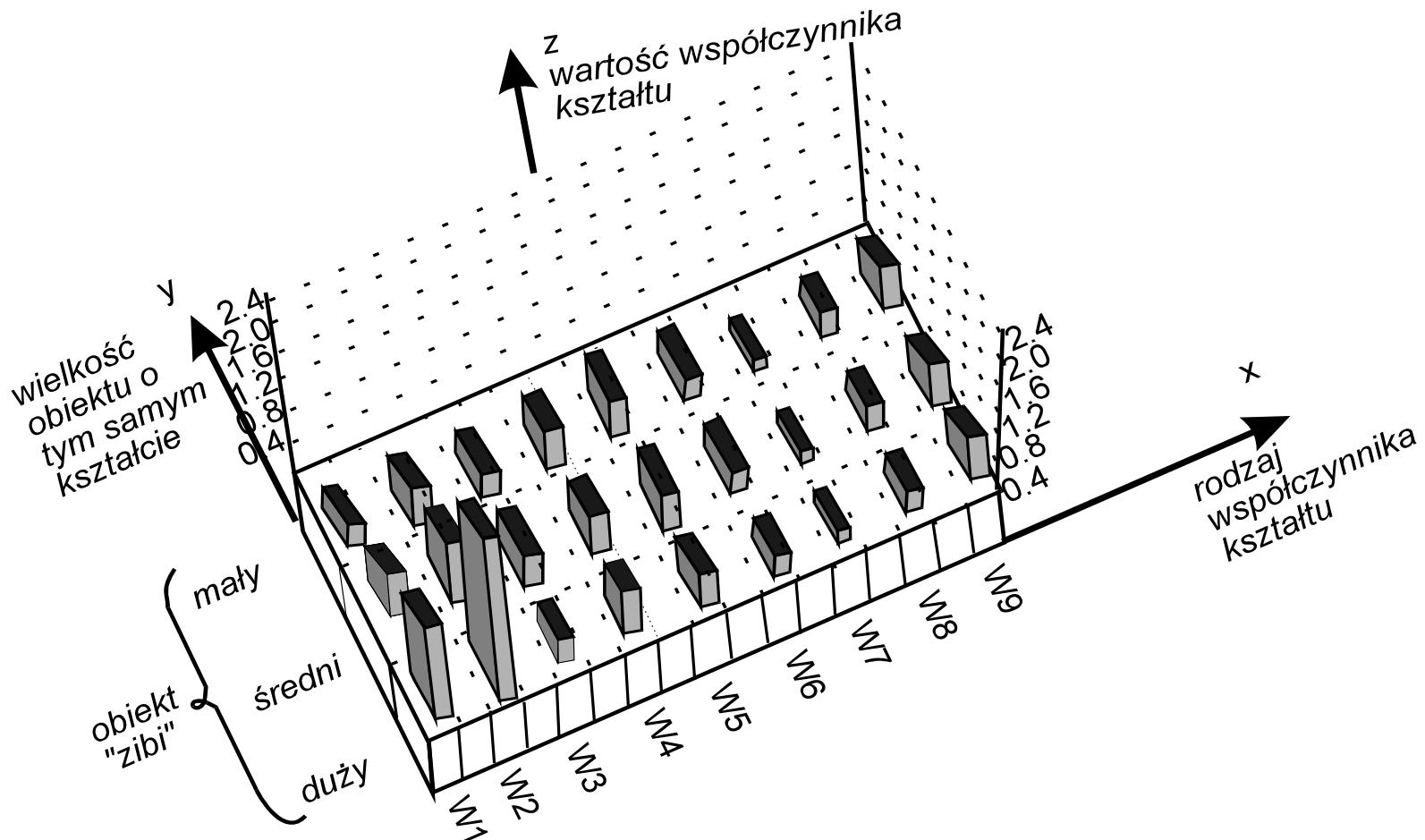
$$\rho_3(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n |x_v^\mu - x_v^\eta|$$

$$\rho_3(\underline{x}^1, \underline{x}^2) = \sum_{v=1}^2 |x_v^1 - x_v^2| = |1| + |0| = 1$$

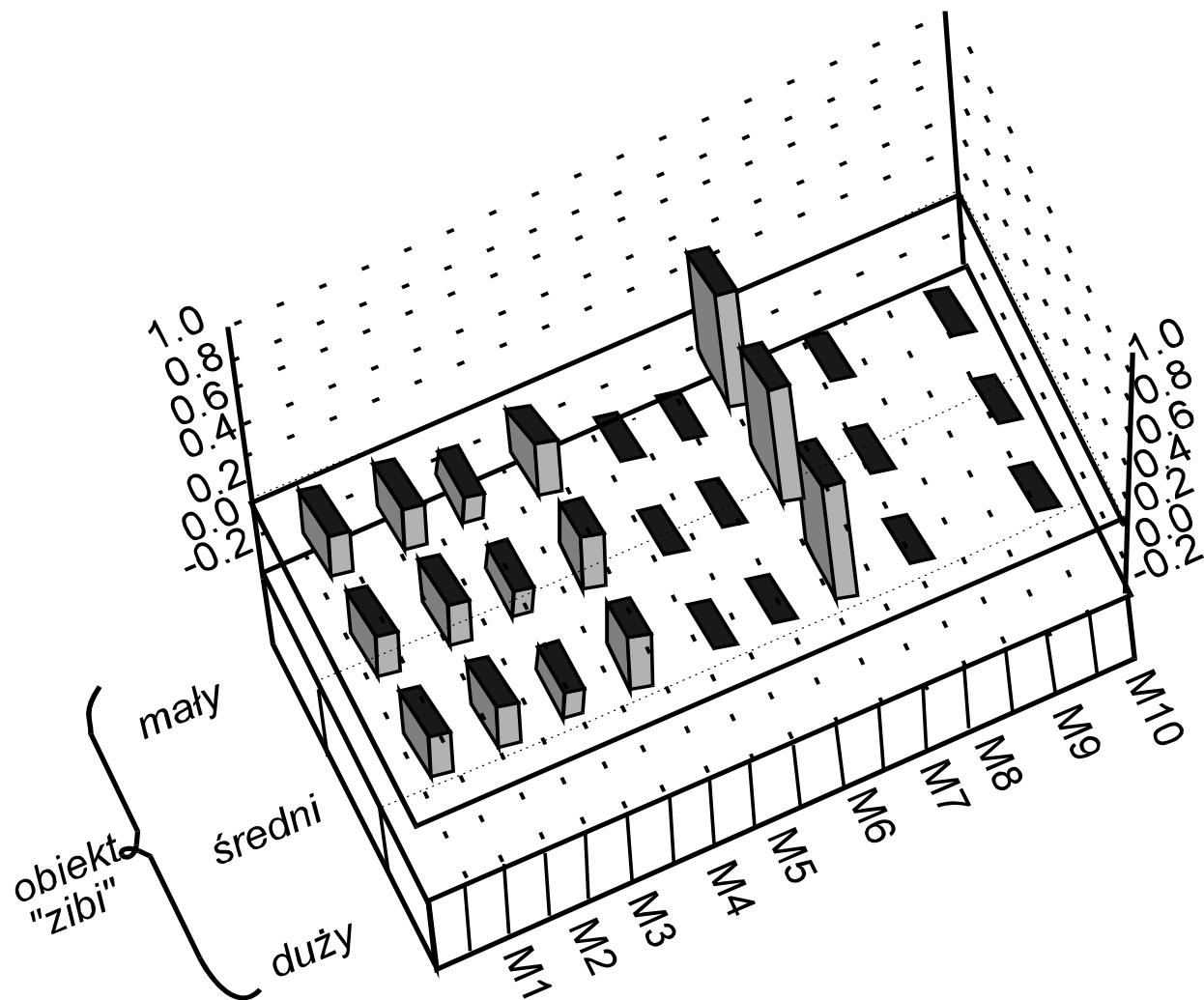
Stąd

$$\rho_1(\underline{x}^1, \underline{x}^2) = \rho_3(\underline{x}^1, \underline{x}^2)$$

Wykresy słupkowe (3D)



Wrażliwość współczynników kształtu na wielkość obiektu



Niewrażliwość momentów na zmiany wielkości obiektu

Momenty

$$\mu_{pq} = \int_A x^p y^q f(x, y) dx dy$$

gdzie p i q – rzędy momentów; A analizowana figura; f(x,y) intensywność pikseli o współrzędnych x i y;

- wyrażenia momentowe nie się zbyt wrażliwe na zmiany kształtów obiektów;
- wpływ dyskretyzacji na momenty daje błąd rzędu kilku %;
- błąd rośnie w miarę wzrostu rzędu momentów;
- zakres przyjmowanych wartości momentów: **10^{-22} – 10^{20}** ;
- w zależności od kształtu obiektów (dla określonej klasy) niektóre momenty przyjmują wartości zbyt małe dla istotności analizy
- momenty których wartości mogą być poniżej 10^{-9} , wtedy przy wyborze **wektora cech** można je pominąć;
- największą **inwariantność** wykazują momenty **M1 i M7**;
- istnieją szybkie algorytmy obliczania momentów

Współczynniki kształtu

- Współczynniki kształtu Wyróżnia się następujące parametry służące do opisu geometrii obiektów: a) pole powierzchni, którego pomiar sprowadza się do zliczenia pikseli należących do interesującego nas obszaru. Cecha ta jest czuła na błędy wynikłe z niewłaściwej binaryzacji, jednak z drugiej strony jest nieczuła na przesunięcie i obrót obiektu w polu widzenia. b) obwód, czyli długość brzegu obiektu. Pomiar tej cechy jest dość trudny z uwagi na konieczność przybliżania ciągłej linii dyskretną kombinacją punktów obrazu. W praktyce wykorzystuje się następujące sposoby pomiaru obwodu

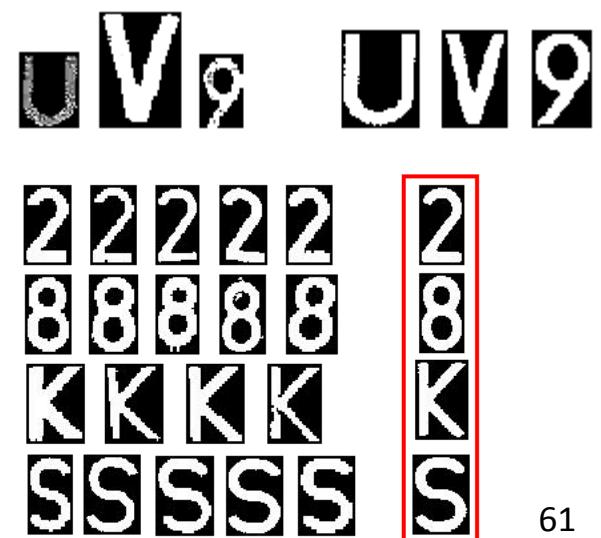
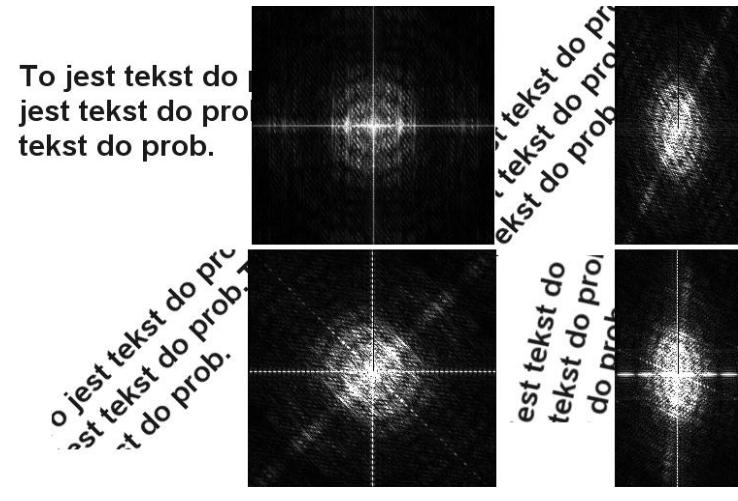
Porównanie współczynników kształtu i momentów

Współczynniki kształtu

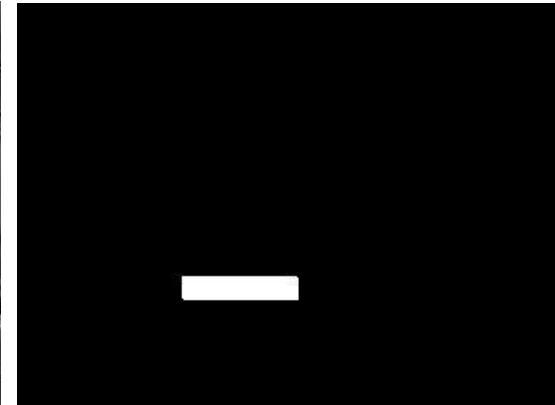
- wykazują większą czułość na zniekształcenia niż momenty;
- wpływ dyskretyzacji na współczynniki daje błąd rzędu kilku %;
- niektóre współczynniki (W_1, W_2) są silnie zależne od wielkości obiektu (zgodnie z ich definicją) i ich użyteczność jest zależna od stopnia normalizacji;
- zakres przyjmowanych wartości (z wyłączeniem W_1 i W_2) 0,01–100,0;
- wszystkie współczynniki mają zbliżoną wrażliwość na deformacje kształtów;
- czas obliczeń współczynnika Danielssona (W_5) jest kilkudziesiąt razy dłuższy od czasu obliczeń pozostałych współczynników

Etapy pracy systemu OCR

- Poziomowanie linii tekstu (FFT z obrazu tekstu i analiza kierunku linii)
- Wydzielenie znaków z linii tekstu i ich normalizacja do wielkości znaków wzorcowych
- Odległość od każdego ze znaków wzorcowych
- Klasyfikacji do klasy na podstawie maksymalizacji prawdopodobieństwa



Rozpoznawanie tablic rejestracyjnych samochodu



Wyszukanie obszaru tablicy rejestracyjnej



Wyodrębnienie znaków na tablicy

WZ 33088

W Z 3 3 0 8 8

Wydzielanie pojedynczych znaków
Rozpoznanie znaków graficznych

,

W

Z

3

3

0

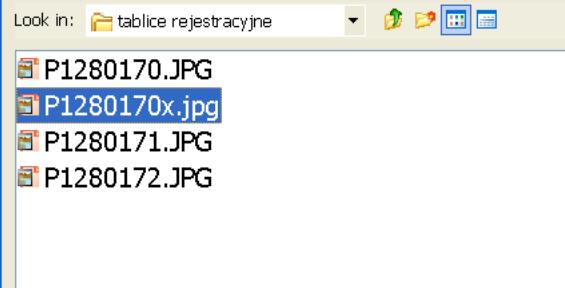
8

8

C:\WINDOWS\system32\cmd.exe

```
imageWidth: 264imageHeight: 71
x1: 13y1: 28
x2: 251y2: 43
tmp x=264tmp y=71
nowy x=268nowy y=75
GET_PLATE
Po labelowaniu mamy :8 labeli!!!!!!
Ilosc labeli: 8
label:1 ma wysokosc:43
label:2 ma wysokosc:32
label:3 ma wysokosc:43
label:4 ma wysokosc:43
label:5 ma wysokosc:43
label:6 ma wysokosc:42
label:7 ma wysokosc:42
label:8 ma wysokosc:27
Wykonam razy :1 !!!!!!!!
Po labelowaniu mamy :7 labeli!!!!!!
#####
Rozpoznawanie
Parametry rozpoznawania:
Odleglosc:HAMMING
```

Open File



Identyfikacja Pojazdów

File Wzorce Operacje Help



Tryb pracy programu

rozpoznawanie nauka

Parametry odszukania znaków

wysokosc znaku (H): kalibracja
wspolczynnik TH contrast ... H / c = 22
wspolczynnik TH (t): H / t = 9
wspolczynnik H closing (h... H * hc = 44
wspolczynnik V opening (... H / vo = 15
wspolczynnik H opening (... H * ho = 88
wspolczynnik V dilation (v... H / vd = 15
wspolczynnik H dilation (... H / hd = 22

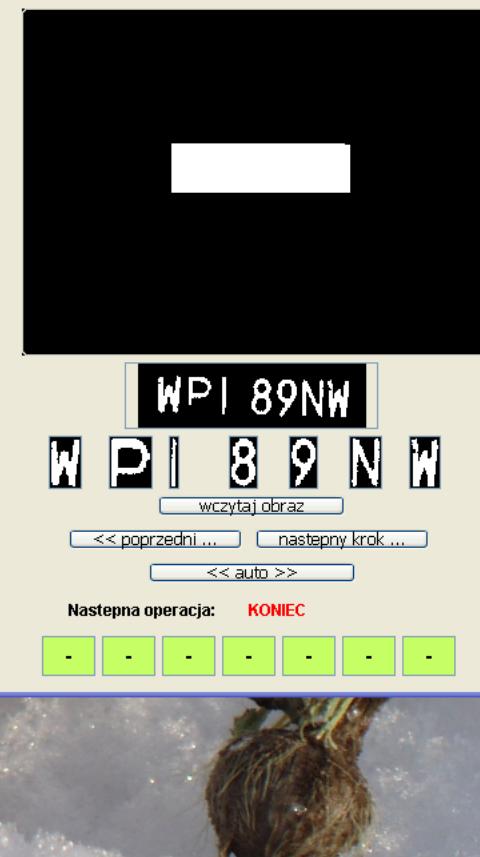
Parametry rozpoznania znaków

Miara odleglosci

Jaccard'a Yu...

srodek cięż... od pr...

uśredn... próg %



Start

3 Ekspl...

readme.tx...

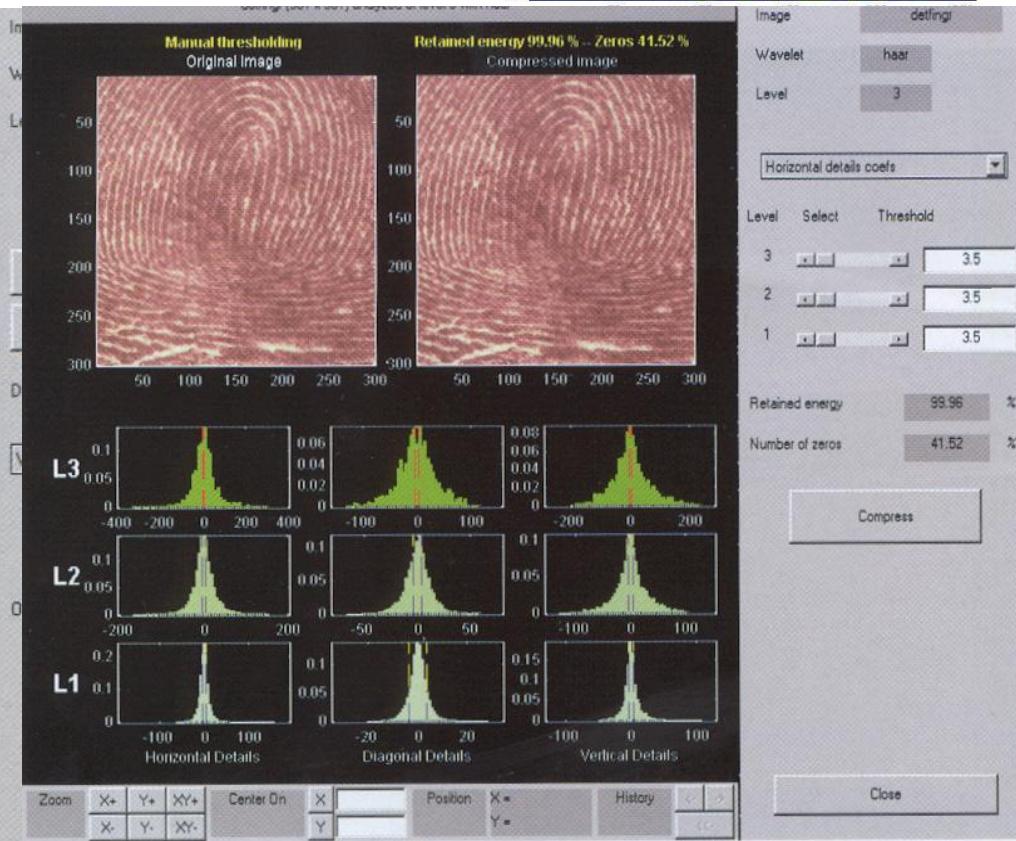
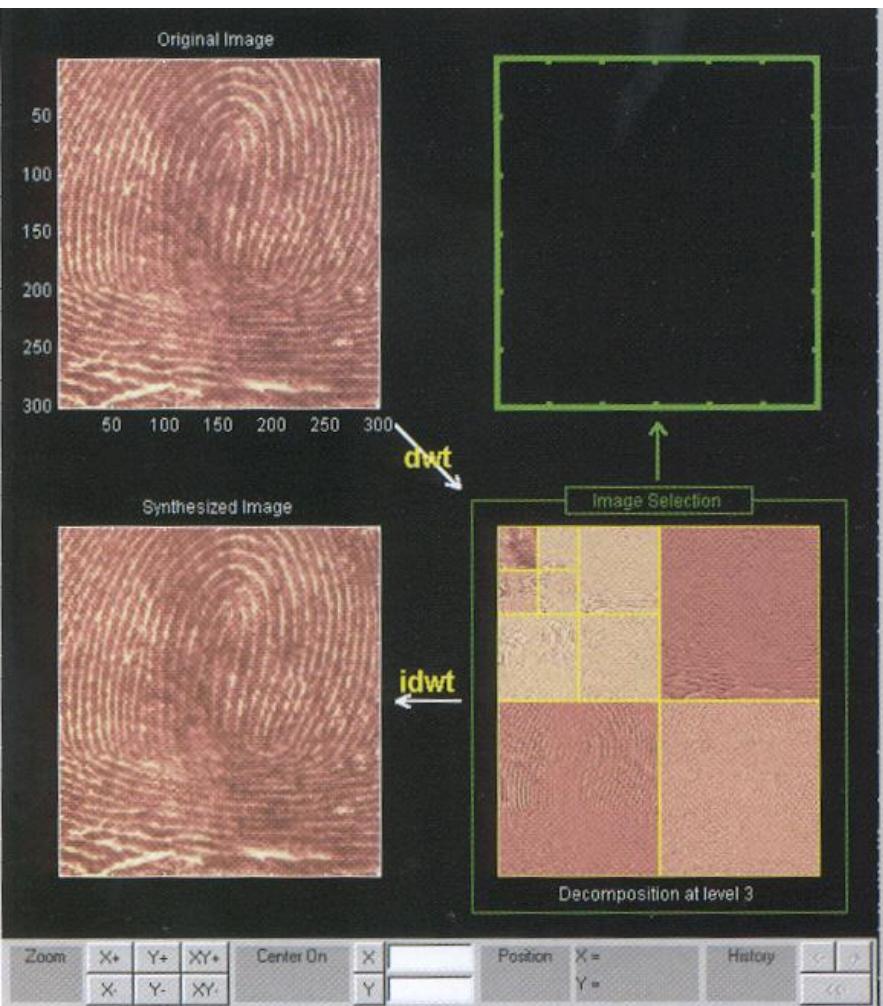
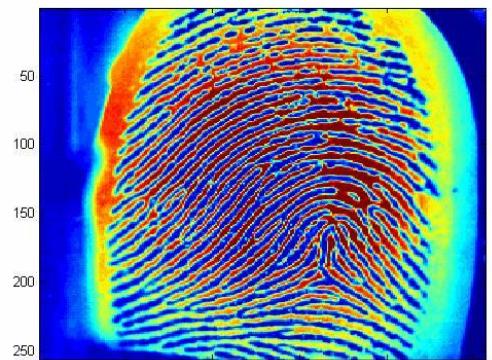
Corel PHO...

C:\WINDO...

Identyfika...

PL 18:21

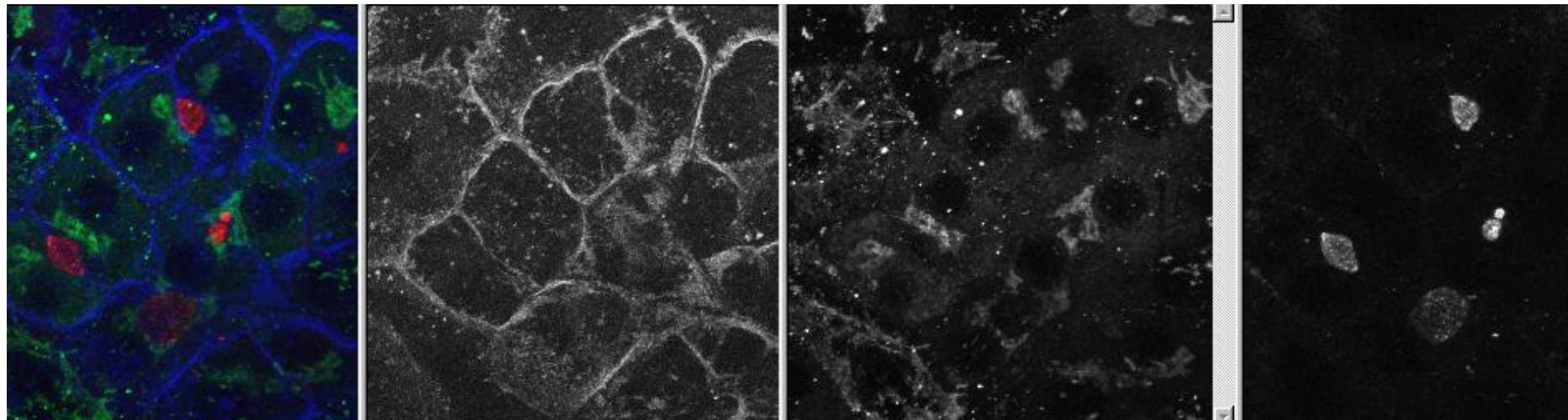
Rozpoznanie odcisków palców



Rys. 7.18. Odcisk palca oryginalny i zrekonstruowany po kompresji, w wyniku której, przy automatycznym ustawieniu progów, zostało wyeliminowanych około 42% współczynników falkowych przy zachowaniu 99,96% energii sygnału oryginalnego

k- najbliższych sąsiadów (k-NN)

Zastosowania metody k najbliższych sąsiadów k-NN

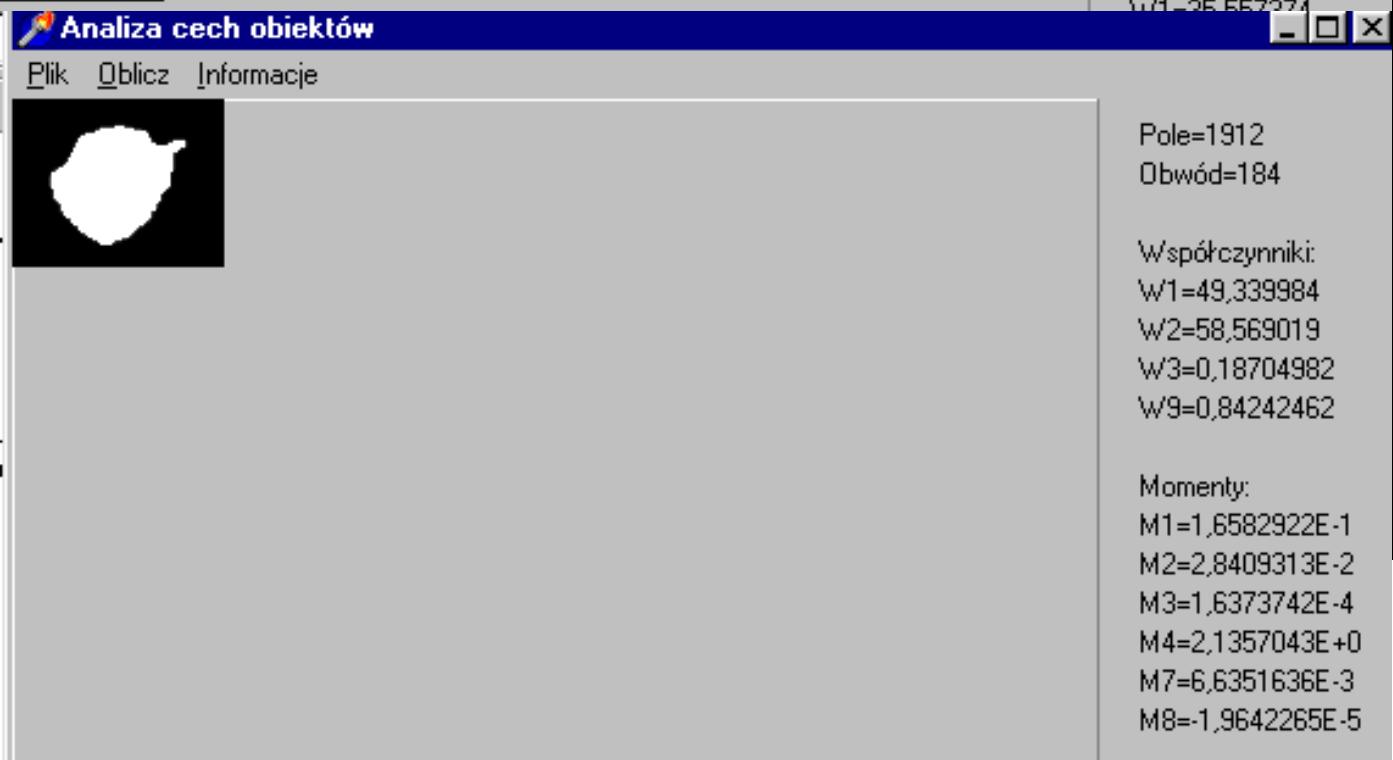




1

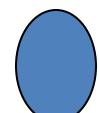
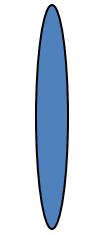


2



4

Współczynnik
kształtu



Optimal Feature Space



Pole w pikselach



?

Model rozpoznawania k-NN

K - najbliższych sąsiadów

Rozpoznawanie ze zbiorem uczącym

Mamy przestrzeń cech i w niej metrykę (Euklidesa, Czebyszewa, Manhattan, Minkowskiego, itp), na podstawie której możemy policzyć odległość między punktami tej przestrzeni.

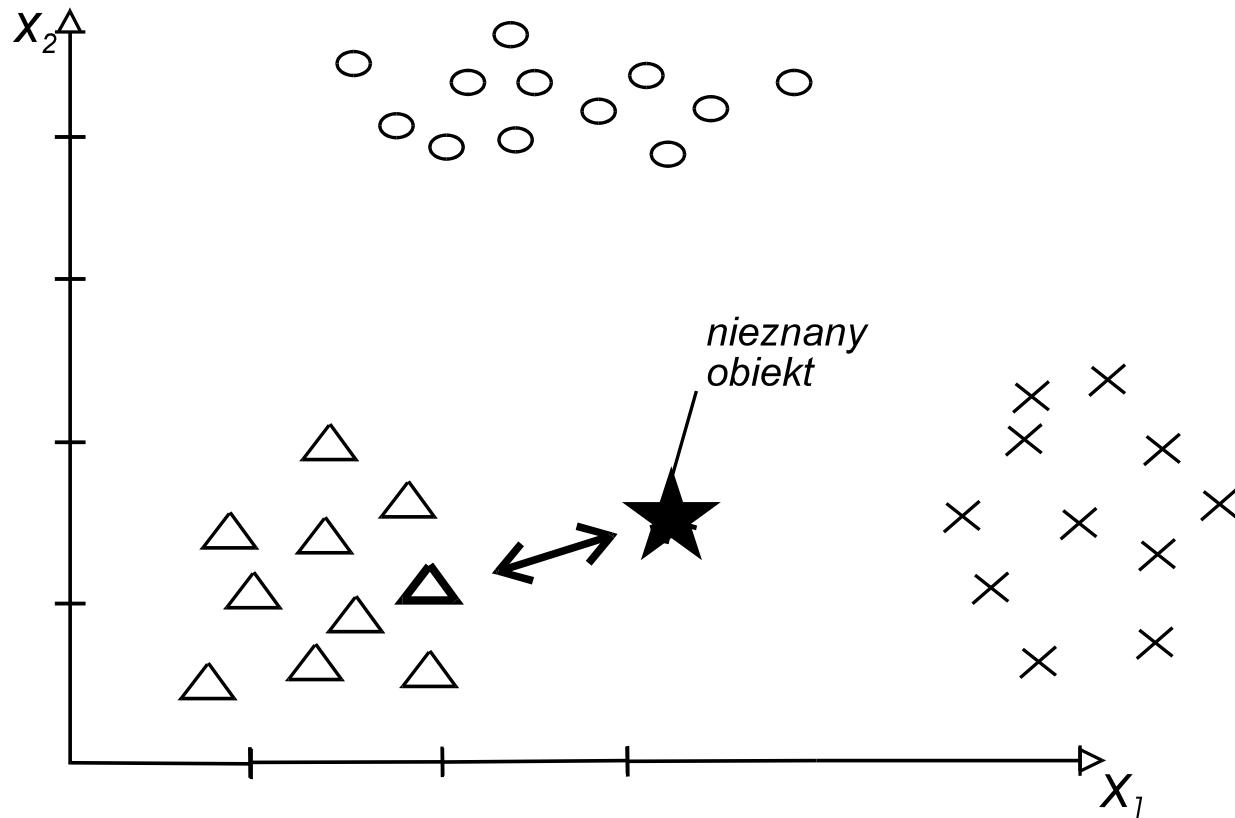
Mamy elementy zbioru uczącego, czyli obiekty o znanej klasyfikacji (dokonanej przez eksperta lub na podstawie innych metod identyfikacji klasy)

Dla nowego obiektu dokonujemy obliczenia jego odległości od wszystkich elementów zbioru uczącego i wybieramy spośród nich k-najbliższych sąsiadów

Obiekt/obraz klasyfikujemy do tej klasy, do której należy najwięcej jego najbliższych sąsiadów

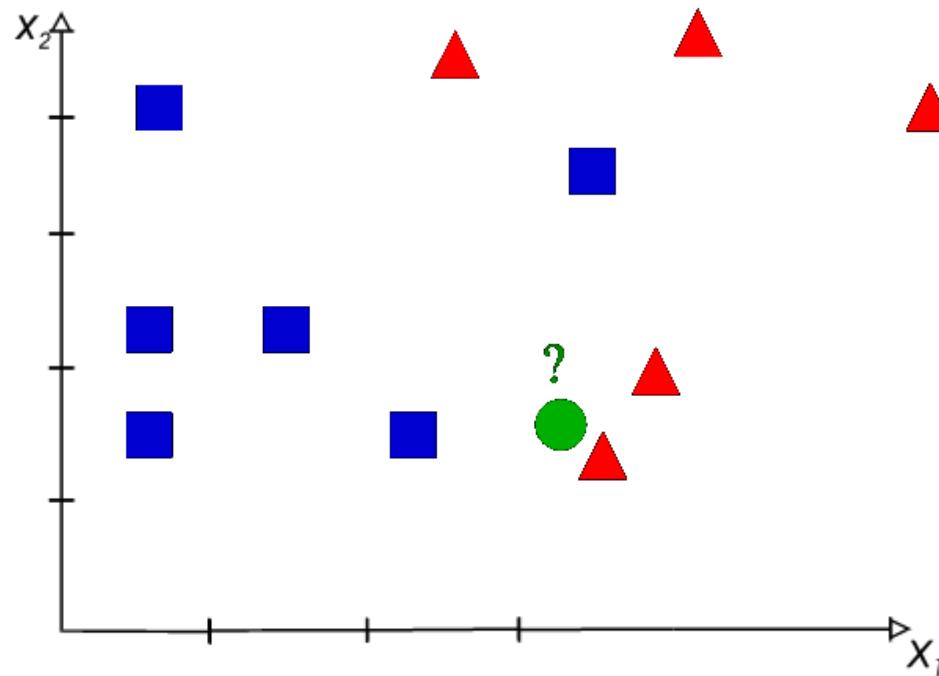
Zasada działania

Zasada postępowania w metodach **minimalno-odległościowych**: analizowany obiekt \star (wektor cech) należy do tej samej klasy, do której należy obiekt najbliższy w przestrzeni cech



Metoda najbliższego sąsiada

- **algorytm NN** - Nearest Neighbour,
czyli najbliższy sąsiad



kNN

Wady metody NN:

**Duża wrażliwość na błędy
ciągu uczącego**

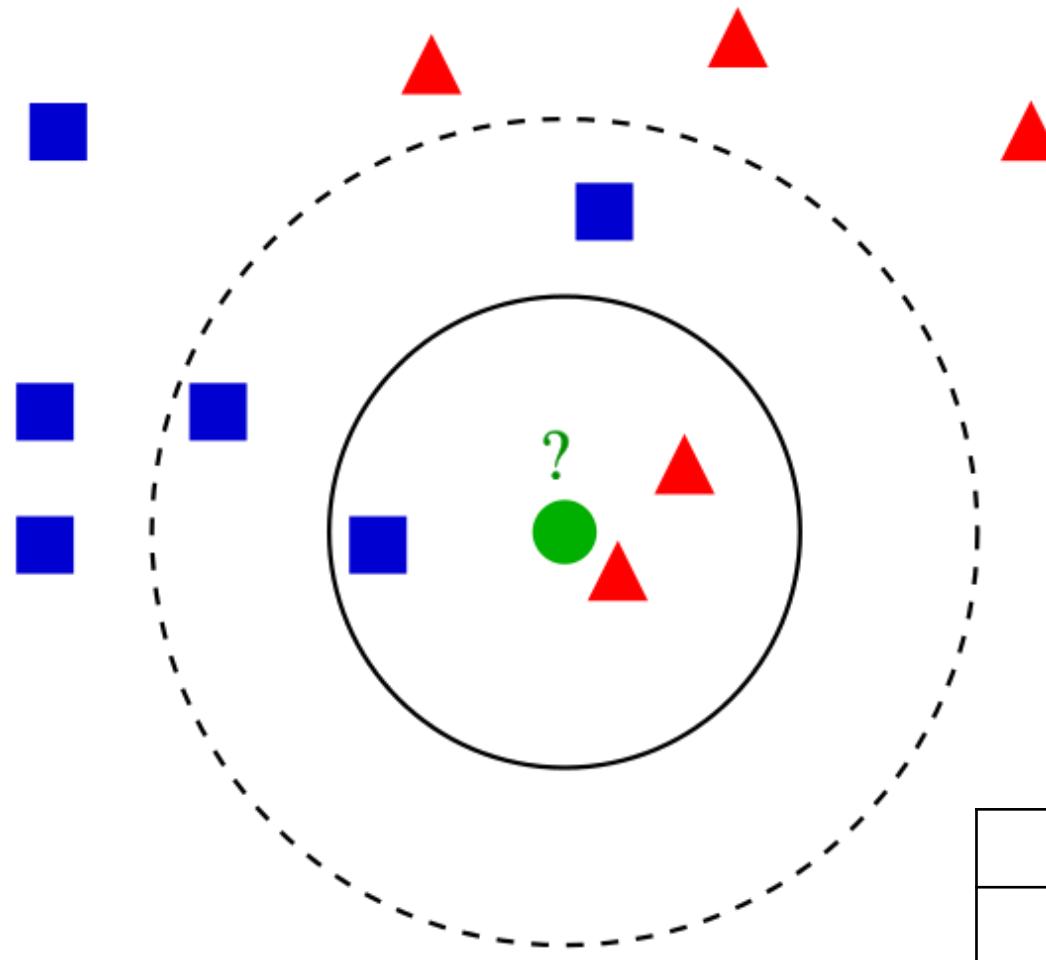
- jeśli błędnie określona zostanie przynależność chociaż jednego elementu ciągu uczącego to wówczas całe jego otoczenie będzie błędnie klasyfikowane

Niwelowanie problemu:

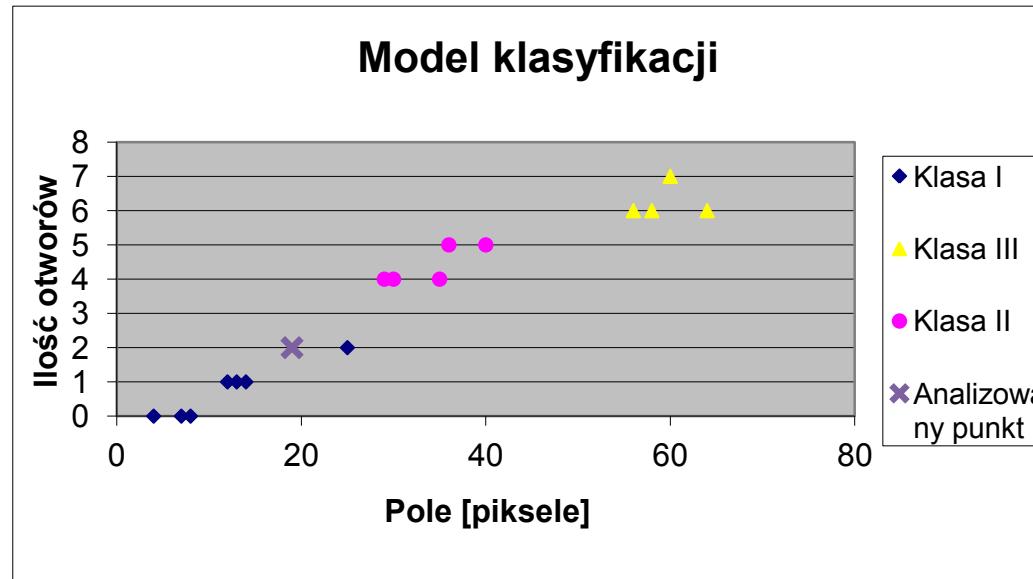
Wprowadzenie metody kNN

- obiekt jest klasyfikowany do tej klasy, do której należy większość z k najbliższych sąsiadów
- gdzie, k – to mała liczba całkowita, np. 3

Przykład kNN



Przykład wykonania



Analizowany punkt	
pole	otwory
19	2

nr	pole	otwory	manhattan
1	29	4	12
2	58	6	43
3	30	4	13
4	12	1	8
5	4	0	17
6	7	0	14
7	60	7	46
8	36	5	20
9	13	1	7
10	14	1	6
11	8	0	13
12	25	2	6
13	35	4	18
14	40	5	24
15	64	6	49
16	56	6	41

nr	pole	otwory	manhattan
10	14	1	6
12	25	2	6
9	13	1	7

3NN

sortowanie
→

nr	pole	otwory	manhattan
4	12	1	8
1	29	4	12
3	30	4	13
11	8	0	13
6	7	0	14
5	4	0	17
13	35	4	18
8	36	5	20
14	40	5	24
16	56	6	41
2	58	6	43
7	60	7	46
15	64	6	49

Obliczenia odległości

Euclidean: Take the square root of the sum of the squares of the differences of the coordinates.

For example, if $x = (\textcolor{green}{a}, \textcolor{red}{b})$ and $y = (\textcolor{green}{c}, \textcolor{red}{d})$, the Euclidean distance between x and y is

$$\sqrt{(\textcolor{green}{a} - \textcolor{red}{c})^2 + (\textcolor{red}{b} - \textcolor{red}{d})^2}.$$

Manhattan: Take the sum of the absolute values of the differences of the coordinates.

For example, if $x = (\textcolor{green}{a}, \textcolor{red}{b})$ and $y = (\textcolor{green}{c}, \textcolor{red}{d})$, the Manhattan distance between x and y is

$$|\textcolor{green}{a} - \textcolor{red}{c}| + |\textcolor{red}{b} - \textcolor{red}{d}|.$$

Chebyshev: Take the maximum value of the absolute values of the differences of the coordinates. For example, in two dimensions, if the points are $x=(a,b)$ and $y=(c,d)$, the Chebyshev distance between x and y is:

$$\max(|a-c|, |b-d|)$$

Obliczenia odległości cech

Metryka euklidesowa:

$$\rho_1(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{\nu=1}^n (x_\nu^\mu - x_\nu^\eta)^2}$$

$$\rho_1(\underline{x}^1, \underline{x}^2) = \sqrt{\sum_{\nu=1}^2 (x_\nu^1 - x_\nu^2)^2} = \sqrt{1+1} = \sqrt{2}$$

Metryka uliczna:

$$\rho_3(\underline{x}^\mu, \underline{x}^\eta) = \sum_{\nu=1}^n |x_\nu^\mu - x_\nu^\eta|$$

$$\rho_3(\underline{x}^1, \underline{x}^2) = \sum_{\nu=1}^2 |x_\nu^1 - x_\nu^2| = |1| + |1| = 2$$

Stąd

$$\rho_1(\underline{x}^1, \underline{x}^2) \neq \rho_3(\underline{x}^1, \underline{x}^2)$$

Sąsiedzi:

Najbliższy

Drugi pod względem
odległości

Następny pod
względem odległości

Kolejny

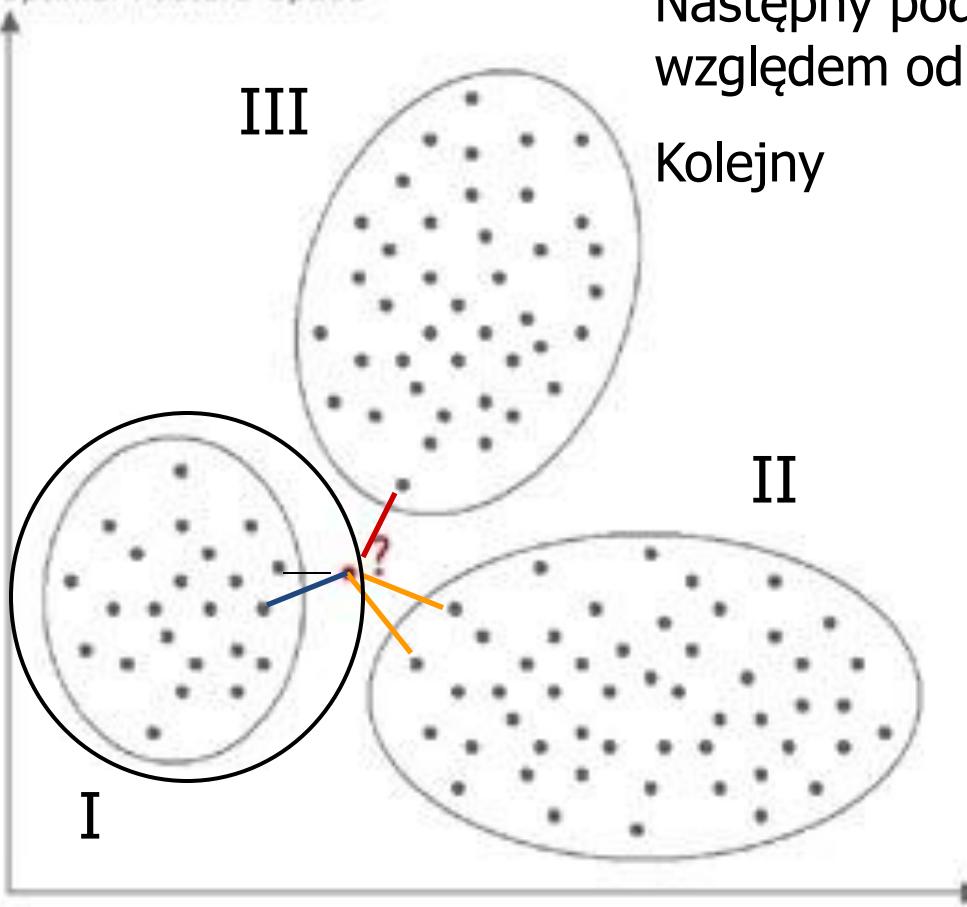
1-NN klasa I

3-NN klasa I

5-NN klasa I

Współczynnik
kształtu

Optimal Feature Space



Podsumowanie k-NN

1. Mamy przestrzeń cech
2. Mamy elementy zbioru uczącego, czyli obiekty o znanej klasyfikacji, i znanym położeniu w przestrzeni cech
3. Dla nowego obiektu dokonujemy obliczenia jego odległości, wg wybranej metryki (Euklidesa, Czebyszewa, Manhattan, Minkowskiego, itp.), od wszystkich elementów zbioru
4. Obiekt klasyfikujemy do tej klasy, do której należy większość jego k-najbliższych sąsiadów

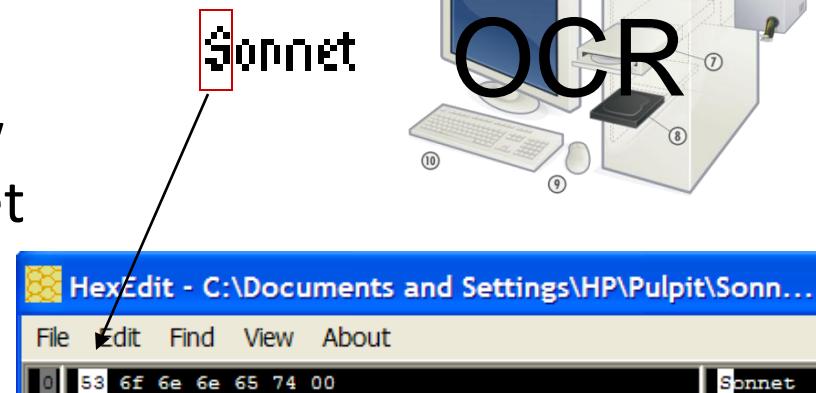
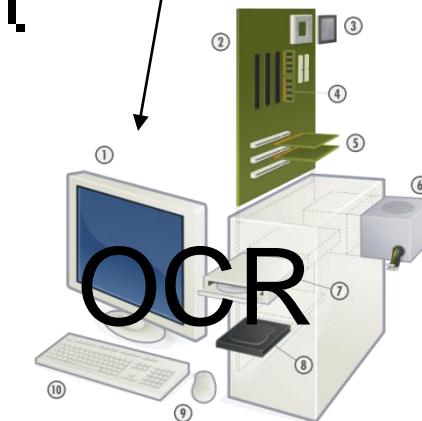
Zastosowanie metody porównania do wzorca

OCR (ang. Optical Character Recognition) - optyczne rozpoznawanie znaków.

Przetworzenie obrazu, zwykle mapy bitowej, np. ze skanera, zawierającej tekst pisany lub drukowany na edytowalny przez edytory komputerowe tekst (tekst kodowany) przez rozpoznawanie znaków pisarskich, kroju pisma, interlinii, a nawet układów tabelarycznych formularzy itp. oraz pól z grafiką.

Plik.bmp

Sonnet
Sonnet



Plik.txt lub .doc

Bez tytułu - Notatnik

Plik Edycja Format Widok Pomoc

Sonnet

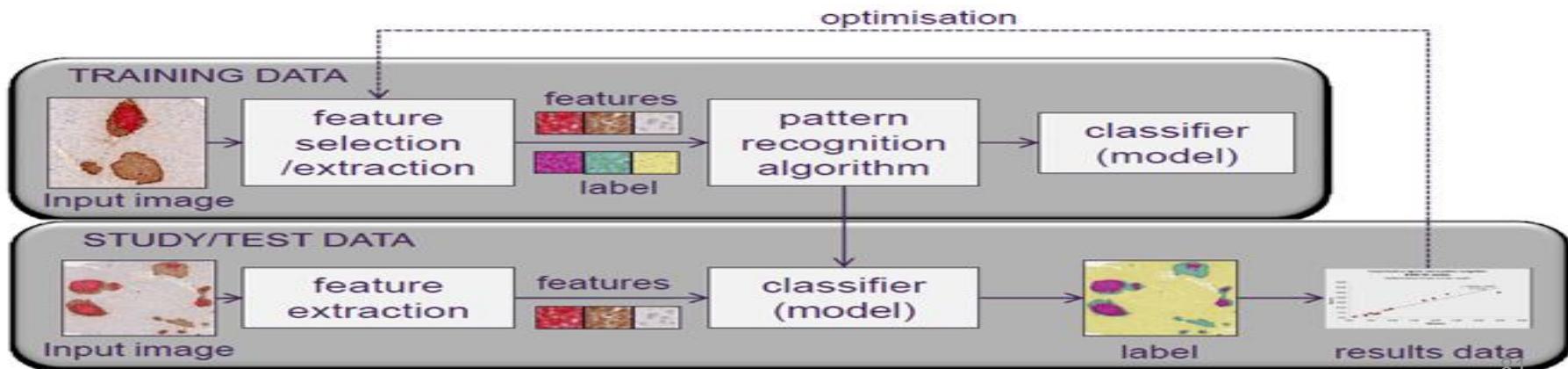
Kody ASCII

Dec	Hex	Znak	Skrót	Dec	Hex	Znak	Skrót	Dec	Hex	Znak	Skrót	Dec	Hex	Znak	Skrót
0	00	Null	NUL	32	20	Spacja		64	40	@		96	60	'	
1	01	Start Of Heading	SOH	33	21	!		65	41	A		97	61	a	
2	02	Start of Text	STX	34	22	"		66	42	B		98	62	b	
3	03	End of Text	ETX	35	23	#		67	43	C		99	63	c	
4	04	End of Transmission	EOT	36	24	\$		68	44	D		100	64	d	
5	05	Enquiry	ENQ	37	25	%		69	45	E		101	65	e	
6	06	Acknowledge	ACK	38	26	&		70	46	F		102	66	f	
7	07	Bell	BEL	39	27	'		71	47	G		103	67	g	
8	08	Backspace	BS	40	28	(72	48	H		104	68	h	
9	09	Horizontal Tab	HT	41	29)		73	49	I		105	69	i	
10	0A	Line Feed	LF	42	2A	*		74	4A	J		106	6A	j	
11	0B	Vertical Tab	VT	43	2B	+		75	4B	K		107	6B	k	
12	0C	Form Feed	FF	44	2C	,		76	4C	L		108	6C	l	
13	0D	Carriage Return	CR	45	2D	-		77	4D	M		109	6D	m	
14	0E	Shift Out	SO	46	2E	.		78	4E	N		110	6F	n	
15	0F	Shift In	SI	47	2F	/		79	4F	O		111	6F	o	
16	10	Data Link Escape	DLE	48	30	0		80	50	P		112	70	p	
17	11	Device Control 1 (XON)	DC1	49	31	1		81	51	Q		113	71	q	
18	12	Device Control 2	DC2	50	32	2		82	52	R		114	72	r	
19	13	Device Control 3 (XOFF)	DC3	51	33	3		83	53	S		115	73	s	
20	14	Device Control 4	DC4	52	34	4		84	54	T		116	74	t	
21	15	Negative Acknowledge	NAK	53	35	5		85	55	U		117	75	u	
22	16	Synchronous Idle	SYN	54	36	6		86	56	V		118	76	v	
23	17	End of Transmission Block	ETB	55	37	7		87	57	W		119	77	w	
24	18	Cancel	CAN	56	38	8		88	58	X		120	78	x	
25	19	End of Medium	EM	57	39	9		89	59	Y		121	79	y	
26	1A	Substitute	SUB	58	3A	:		90	5A	Z		122	7A	z	
27	1B	Escape	ESC	59	3B	:		91	5B	I		123	7B	{	
28	1C	File Separator	FS	60	3C	<		92	5C	\		124	7C		
29	1D	Group Separator	GS	61	3D	=		93	5D	J		125	7D	}	
30	1E	Record Separator	RS	62	3E	>		94	5E	^		126	7E	~	
31	1F	Unit Separator	US	63	3F	?		95	5F	_		127	7F	Delete	DEL

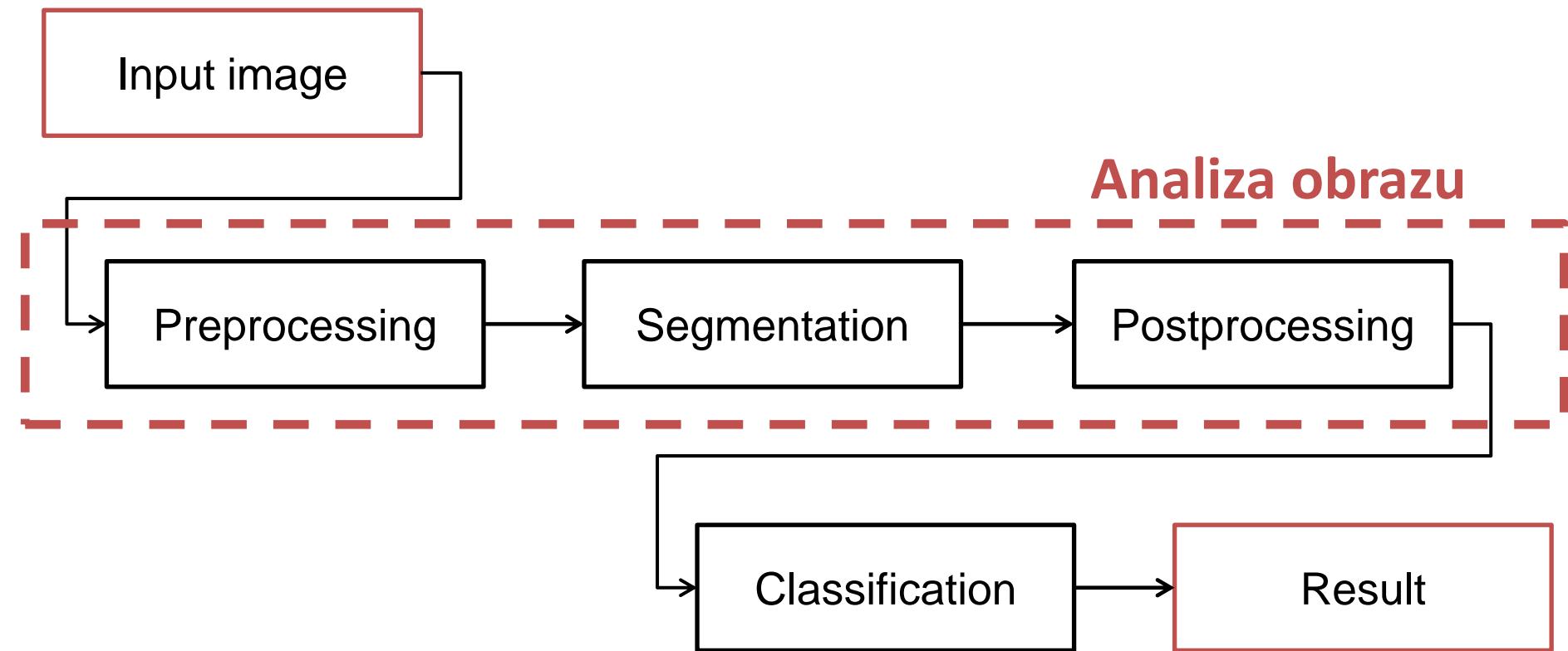
Podsumowanie metod klasycznych

Klasyczne algorytmy rozpoznawania obrazów zawsze **zależą** od skomplikowanego przetwarzania obrazu w celu poprawy początkowej jakości obrazu i umożliwienia **segmentacji / separacji obiektów**.

Jednak najważniejszy jest proces inżynierii cech, który jest kluczową klasyczną techniką rozpoznawania. Tradycyjne „ręcznie” **dobierane cechy i funkcje** w dużej mierze **opierają się na wiedzy specjalistycznej** w dziedzinie, które powinny zaproponować je inżynierom i przygotować adnotację do zestawu szkoleniowego.



Podsumowanie metod klasycznych



Koniec dzisiejszego wykładu

Rozpoznawanie w oparciu o algorytmy CNN. Steganografia

WYKŁAD 8
Dla studiów niestacjonarnych
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

Komputerowa analiza i rozpoznawanie obrazów



Jest to „sztuka” udzielania, **automatycznej** i mającej **matematyczne podstawy** odpowiedzi na pytanie:

Co ten obraz przedstawia (i o czym mówi) ?

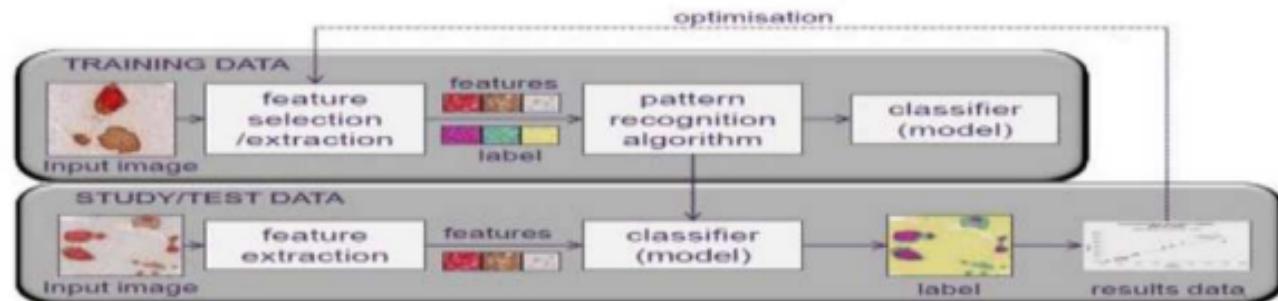
Podejścia:

1. **Klasyczna** – oparta o cechy dobranych przez deweloperów systemu na podstawie rad ekspertów i ich doświadczenia
 - **Analiza obrazu** - zajmuje się ekstrakcją cech obiektów wysegmentowanych z obrazów na potrzeby rozpoznawania obrazów
 - **Właściwe rozpoznawanie obrazu** - zajmuje się tworzeniem i weryfikacją reguł, na podstawie których udziela się odpowiedzi na powyższe pytanie oraz stosowaniem tych reguł w konkretnych zagadnieniach praktycznych
2. **Oparta na sztucznej inteligencji** - bazuje na cechach wyznaczonych automatycznie w procesie uczenia

Podsumowanie metod klasycznych

Klasyczne algorytmy rozpoznawania obrazów zawsze zależą od skomplikowanego przetwarzania obrazu w celu poprawy początkowej jakości obrazu i umożliwienia segmentacji / separacji obiektów.

Jednak najważniejszy jest proces inżynierii cech, który jest kluczowy w klasycznej technice rozpoznawania. Tradycyjne „ręcznie” dobierane cechy i funkcje w dużej mierze opierają się na wiedzy specjalistycznej w dziedzinie. Najczęściej są zaproponowane deweloperom systemu przez specjalistów, którzy przygotowują adnotację do zestawu szkoleniowego.



Rozpoznawanie obrazów metodami opartymi na głębokich konwolucyjnych sieciach neuronowych

Wymagają nauczenia sieci neuronowej rozwiązywania pewnego zadania stosując algorytmy głębokiego uczenia (ang. deep learning) należące do dziedziny zwanej sztuczną inteligencją

Sztuczna inteligencja

Sformułowanie „sztuczna inteligencja” (ang. Artificial Intelligence; AI) jest używana do nazwania takich komputerów/maszyn, które naśladują „funkcje poznawcze”, które ludzie kojarzą z ludzkim umysłem, takie jak „uczenie się” i „rozwiązywanie problemów”.

Powstała jako dyscyplina akademicka w 1955. Termin zaproponował McCarthy (z Uniwersytetu Stanford) na konferencji w Dartmouth

1. inteligencja realizowana w procesie technicznym, a nie naturalnym, biologicznym;
2. dziedzina badań naukowych informatyki i kognitywistyki czerpiąca także z osiągnięć psychologii, neurologii, matematyki i filozofii.

Definicja z 2019 r: „... zdolność systemu (maszyny, komputera dopisek) do prawidłowego interpretowania danych pochodzących z zewnętrznych źródeł, nauki na ich podstawie oraz wykorzystywania tej wiedzy, aby wykonywać określone zadania i osiągać cele poprzez elastyczne dostosowanie.”

Cytat z: Andreas Kaplan; Michael Haenlein (2019) Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence, Business Horizons, 62(1), 15-25

Algorytmy sztucznej inteligencji (AI)

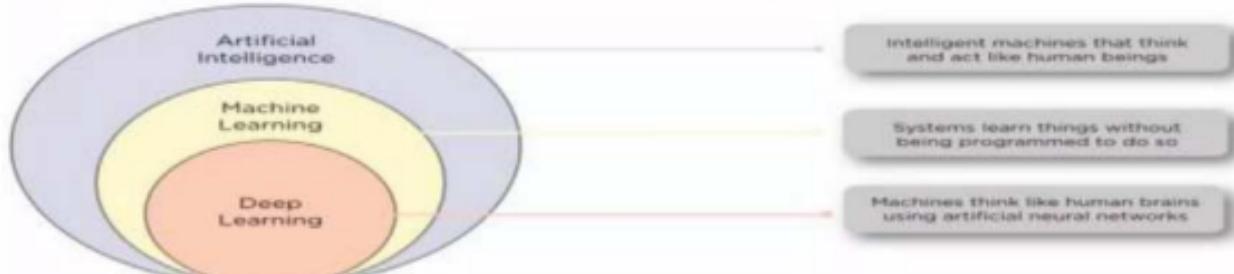
Sztuczna inteligencja jako technologia aktualnie dostępna potrafi:

- Używać naturalnego werbalnego języka, tłumaczyć z języka na język, ale nie potrafi prowadzić komunikacji pozawerbalnej; wyczuć intencji i nastroju mówiącego
- Uczyć się na podstawie danych, na próbach i błędach przez powtarzanie procesy uczenia, ale tylko w stosunku do konkretnego ściśle wyznaczonego zadania
- Otrzymuje bardzo dokładne rezultaty w specyficznych trudnych dla człowieka zadaniach obliczeniowych, w grach o ściśle określonych regułach, poszukiwaniu podobieństwa w danych and nie potrafi żadnej z tych czynności adoptować do nawet lekko zmienionych warunków zadania
- nie ma celu, woli, emocji, poczucia humoru, zawziętości czy wyrozumiałości i innych cech charakteru (indywiduonalności)

Zastosowanie AI w: wizji komputerowej, w rozpoznawaniu mowy, maszynowym tłumaczeniu, w filtrowaniu poczty, graniu w gry nawet trudne jak szachy czy Go . Ostania również odnosi wielkie sukcesy we wspomaganiu diagnostyki w medycynie.

Uczenie maszynowe czy uczenie się maszyn?

AI with Machine Learning and Deep Learning



- **Uczenie z nauczycielem, nadzorowane** (supervised learning) – na podstawie przeszłych danych o wejściu i wyjściu dostarczonych przez dewelopera systemu (np. od ekspertów)
- **Uczenie bez nauczyciela, nienadzorowane** (unsupervised learning) – ukryte zależności i organizacja danych, ich podobieństwa i zaburzenia, czy odchylenia od normy są identyfikowane przez system bez pomocy ekspertów w dziedzinie lub dewelopera systemu
- **Nie ma treningu tylko następuje uczenie się dzięki próbom i błędom** (reinforcement learning) i naturalnej pochodzącej ze środowiska karze, która pozwala podnieść efektywność uczenia się np.: uczenie chodu i gry np.: szachy
- **Uczenie częściowo nadzorowane** (weakly supervised), w którym na wstępie zastosowane jest uczenie z nauczycielem na tzn. danych etykietowanych, a następnie zasadnicza część uczenia to uczenie bez nauczyciela.

Przykłady możliwości systemów opartych na uczeniu maszynowym

- Prowadzenie samochodów i pilotowanie samolotów oraz robotów
- Rozpoznawanie mowy (2012);
- Ekstrahuje i aplikuje cechy stylu
(Deep Convolution Generative Adversarial Networks DCGAN)



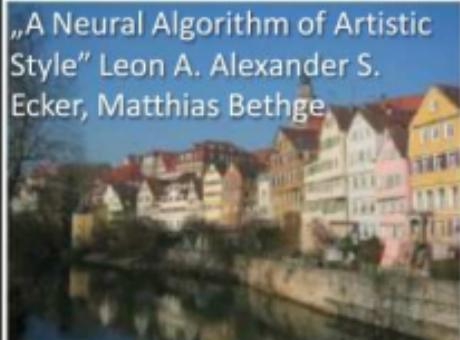
Nvidia

<https://www.youtube.com/watch?v=qhUvQiKec2U>



Boston

Dynamics:
<https://www.youtube.com/watch?v=e9QzlkP5ql>



„A Neural Algorithm of Artistic Style” Leon A. Alexander S. Ecker, Matthias Bethge



Van Gogh

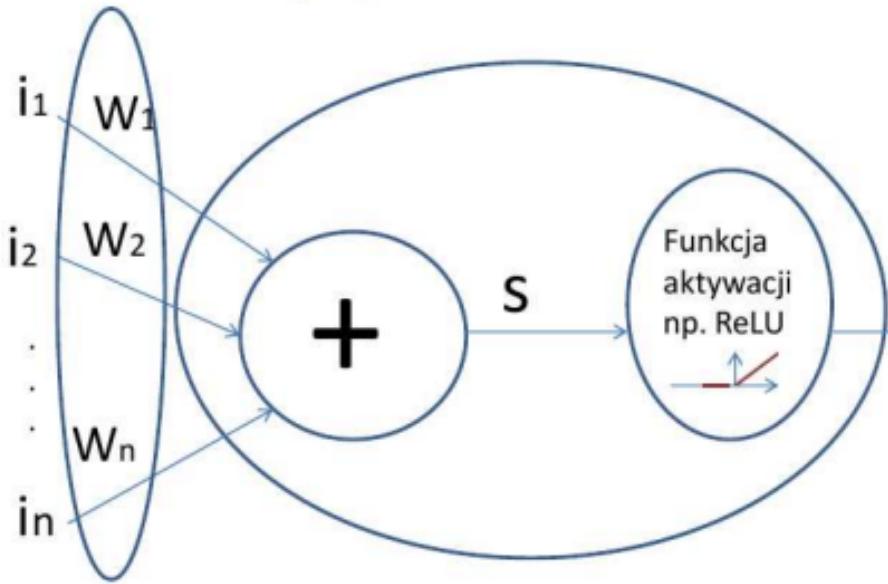


Munch

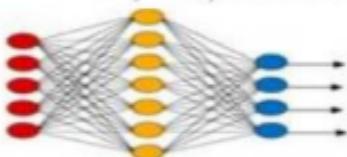


Jak są zbudowane sztuczne sieci neuronowe?

Sieci neurunowe



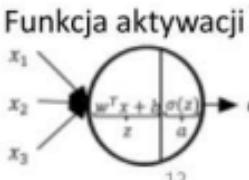
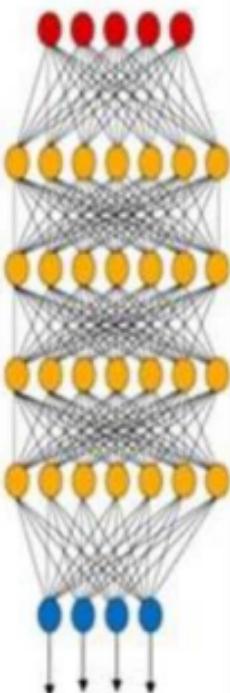
Elementem funkcjonalnym sieci neuronowej jest perceptron / neuron - matematyczna symulacja funkcji komórek nerwowych (odbiór i odesłanie przetworzonego sygnału). Sieć jest złożona z warstw neuronów połączonych w taki sposób, że wyjście warstwy wcześniejszej jest wejściem warstwy następującej po niej. Połączenia poszczególnych neuronów regulują przekazywany sygnał według dobieranych w procesie uczenia współczynników W



McCulloch and Walter Pitts
w 1943 roku

Głębokie sieci neuronowe

- Głębokie sieci neuronowe (NN) to sieci o więcej niż trzech warstwach.
- Sieci neuronowe uczą się z danych źródłowych i mapują „informacje/wiedze wciągniętą ze zbioru uczącego” do **współczynników sieci**.
- W zadaniach rozpoznawania sieci neuronowe potrafią w procesie optymalizacji automatycznie dobrać cechy konieczne do rozpoznawania pewnych klas obiektów w procesie treningu.
- Typowe sieci neuronowe w eksploatacji zachowują się jak „czarna skrzynka”.
- Sieci, w których mamy dostęp do informacji o „przyczynach” decyzji o rozpoznawaniu i klasyfikacji nazywają się po angielski **eXplanable NN; (XNN)**
- Okazuje się, że kombinacja cech pochodzących od ekspertów (używanych w klasycznych metodach rozpoznawania) z cechami zidentyfikowanymi na etapie uczenia się maszyny daje szansę nie tyle bardzo dobre wyniki, ale zrozumiałe dla użytkownika wyniki.



Sztuczne sieci neuronowe

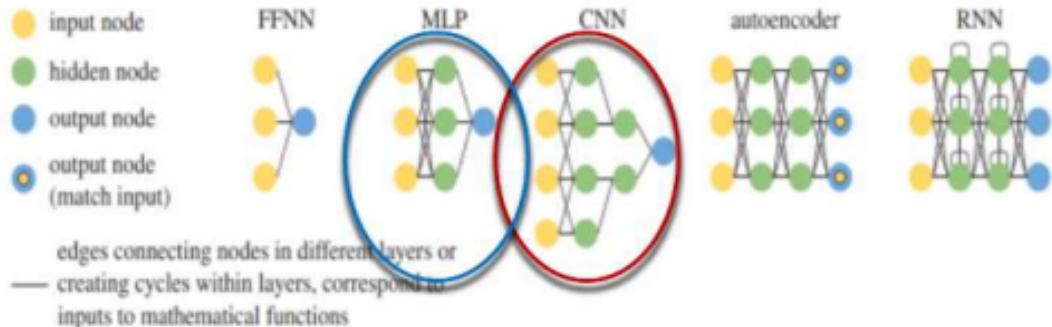


Figure 1. Neural networks come in many different forms. Left: A key for the various types of nodes used in neural networks. Simple FFNN: a feed-forward neural network in which inputs are connected via some function to an output node and the model is trained to produce some output for a set of inputs. MLP: the multi-layer perceptron is a feed-forward neural network in which there is at least one hidden layer between the input and output nodes. CNN: the convolutional neural network is a feed-forward neural network in which the inputs are grouped spatially into hidden nodes. In the case of this example, each input node is only connected to hidden nodes alongside their neighbouring input node. Autoencoder: a type of MLP in which the neural network is trained to produce an output that matches the input to the network. RNN: a deep recurrent neural network is used to allow the neural network to retain memory over time or sequential inputs. This figure was inspired by the Neural Network Zoo by Fjodor Van Veen.

Headline review



Cite this article: Ching T et al. 2018

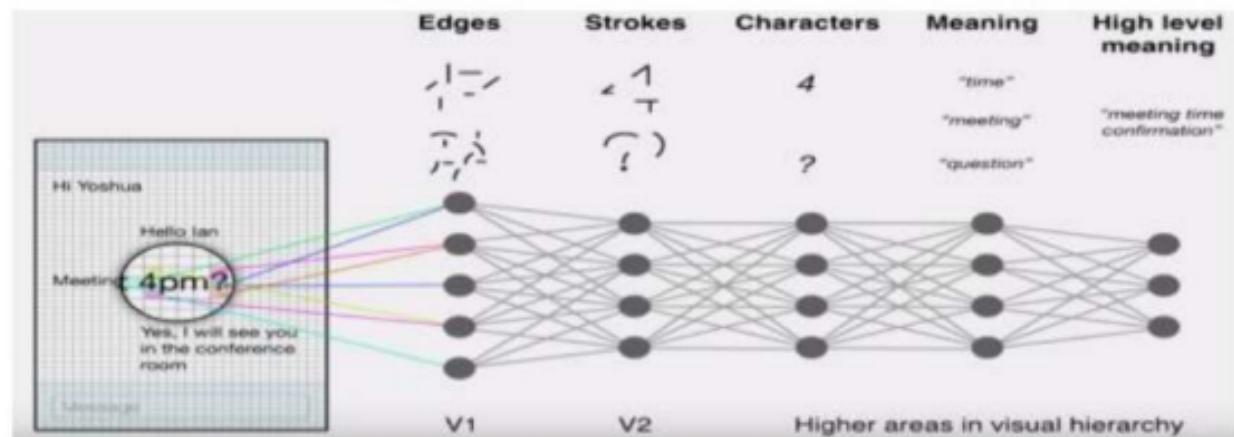
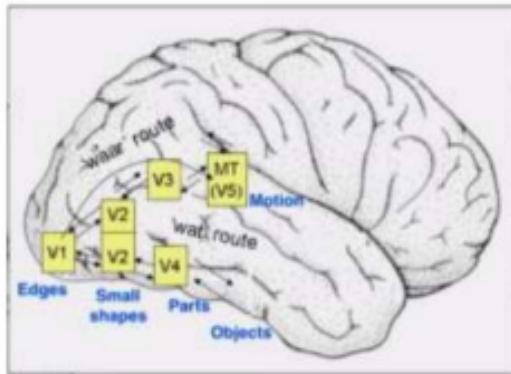
Opportunities and obstacles for deep learning

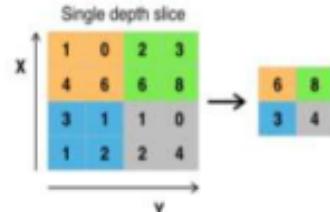
in biology and medicine. *J. R. Soc. Interface* **15**:

20170387.

<https://doi.org/10.1098/rsif.2017.0387>

Praca głębokiej konwolucyjnej sieci (CNN) imituje proces rozpoznawanie w ludzkim mózgu





Warstwy

Podstawowe rodzaje warstw i procesów z nimi związanych:

- **Klasyczna** – Fully Connected (FC) - dense
- **Konwolucyjna** - W odróżnieniu od warstwy typu fully-connected, rozmiar wyniku działania warstwy konwolucyjnej zależy od rozmiaru danych wejściowych, rozmiaru filtrów i funkcji zmieniającej rozmiar, oraz wartości kroku. Sieci tego typu dobrze działają tam gdzie jest duża redundancja informacji np.: przy analizie obrazów, dźwięków lub zbioru cech.
- **Pooling**. Pooling to proces zmniejszeniu przestrzeni cech/rozmiaru sieci wewnętrz konwolucyjnej sieci neuronowej przez uśrednianie (average pooling) lub przyjmowanie wartości maksymalnej (max pooling) z określonego otoczenia.
- **Padding**, czyli uzupełnianie odpowiednimi wartościami, do określonego rozmiaru. Od liczby dodanych zer (stosowany zero padding) zależy, czy mapa wyjściowa będzie miała rozmiar większy, mniejszy czy taki sam, w porównaniu do rozmiaru mapy wejściowej na danym poziomie-warstwie.

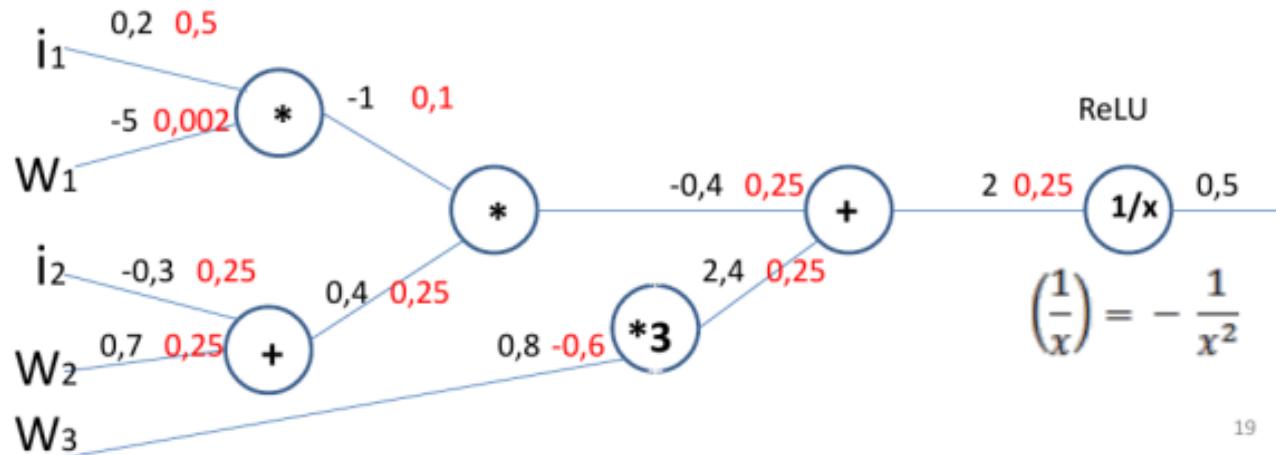
**Jakie matematyczne operacje
umożliwiają uczenie się sieci**

Funkcje aktywacji

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Tanh		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
Arctan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parametric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

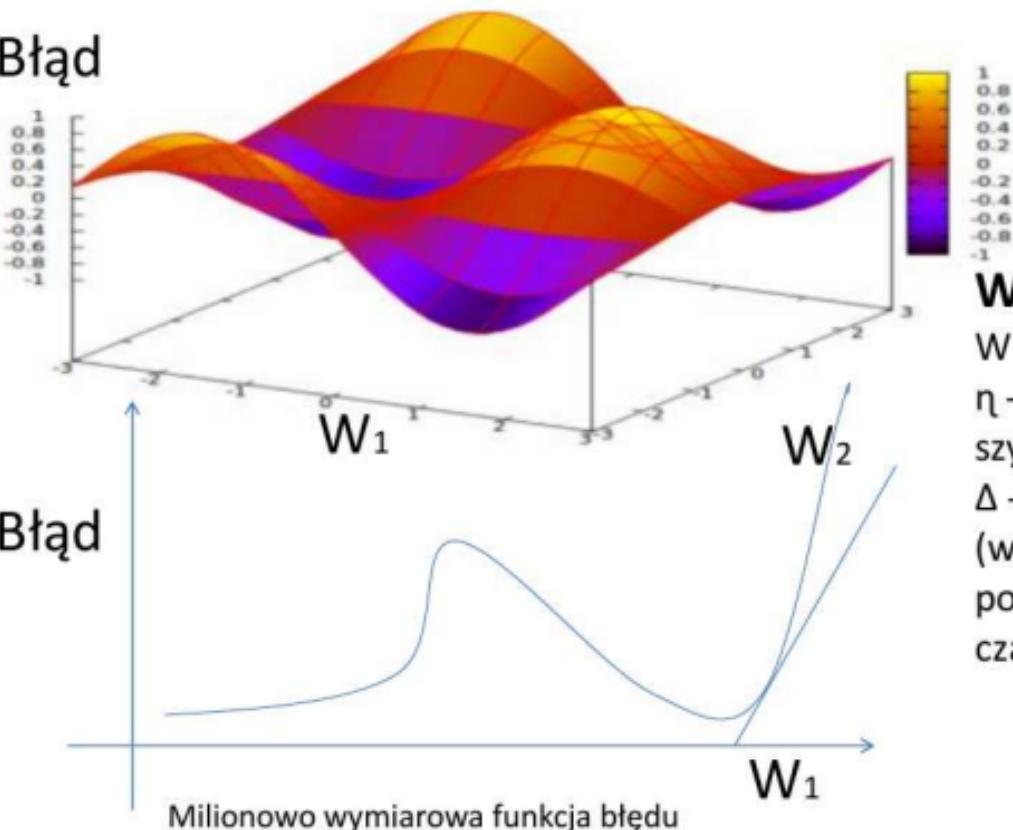
Algorytm propagacji wstecznej ang. back propagation algorithm

Algorytm oblicza wielkość korekty wag połączeń neuronów rozmieszczonych w sąsiednich warstwach sieci. Oparty jest on na minimalizacji funkcji błędu (np.: sumy kwadratów błędów) uczenia z wykorzystaniem optymalizacyjnej metody największego spadku.



Funkcja błędu

Błąd



$$W_n = W_p - \eta \Delta$$

W – wektor of
 η - współczynnik
szybkości uczenia
 Δ - gradient
(wektor
pochodnych
częstekowych)

Milionowo wymiarowa funkcja błędu

Pojęcia związane z uczeniem sieci neuronowych

- **Przeuczenia (ang. Overfitting)**: zbyt mała ilość danych w porównaniu do liczby parametrów (zanikanie), zwiększenie zbioru uczącego, zmniejszenie liczby parametrów, losowe wykluczanie neuronów
- **Zanikanie gradientu (ang. Gradients disappears)** zamieniając funkcję aktywacji lub wzmacniając gradient przez regularyzację
- **Zbieżność/konwergencja procesu uczenia (ang. Convergence of learning)** zbieżność procesu optymalizacji zmiana funkcji błędu
- **Ocena poprawności uczenia (ang. Correctness assessment)**: dokładność i precyza, Classical MSE / mutual entropy (CE -cross entropy; Averaged CE (ACE- averaged CE); F1
- **Generalizacja (ang. Generalization)** – cel uczenia

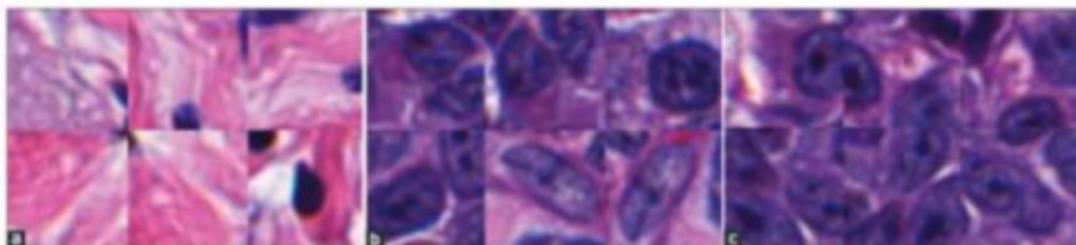
Augmentacja (ang. augmentation)

To proces powielenia danych uczących przez:

- **Transformacje obrazów** (ang. adversarial transformation); translacje, obroty, przekształcenia afiniczne?
- **Dołożenia pewnych zakłóceń** – rozjaśnienie, przyciemnienia, dodanie szumu zmiana tonacji, itp

J Pathol Inform 2016; 1:29

<http://www.jpathinformatics.org/content/7/1/29>



Podsumowanie o sieciach CNN

Sztuczna sieć neuronowa

- Poprawa efektywności rozpoznawania, klasyfikacji i segmentacji obrazów przez żmudny **proces optymalizacji sieci – uczenia się**, który przypomina proces Darwinowskiego doboru naturalnego
- Na etapie korzystanie z nauczonej sieci mamy na **wejściu dane, a na wyjściu prawdopodobieństwo** że obiekt należy do którejś ze znanych kategorii **i tylko tyle!**
- **Nie ma „rozumienia”**, bo nie ma kontekstu; nie można wyobrazić sobie konsekwencji, itp.





CNN wady i zalety

Możliwości CNN:

- Nadzorowane podejście do uczenia maszynowego polega na **szkoleniu modelu statystycznego** przy użyciu zestawu obrazów ze zbioru uczącego, dla których istnieją etykiety i oznaczenia przygotowane przez ekspertów w dziedzinie. Wyuczony model odwzorowuje automatycznie wybrane obiekty na **klasy/kategorie**, czyli uczy się z doświadczeń.
- **Połączenie różnych informacji w modelu** np.: danych klinicznych i obrazów diagnostycznych daje lepsze wyniki predykcji klasyfikacji.



CNN wady i zalety

CNN mają pewne ograniczenia, które są do pokonania:

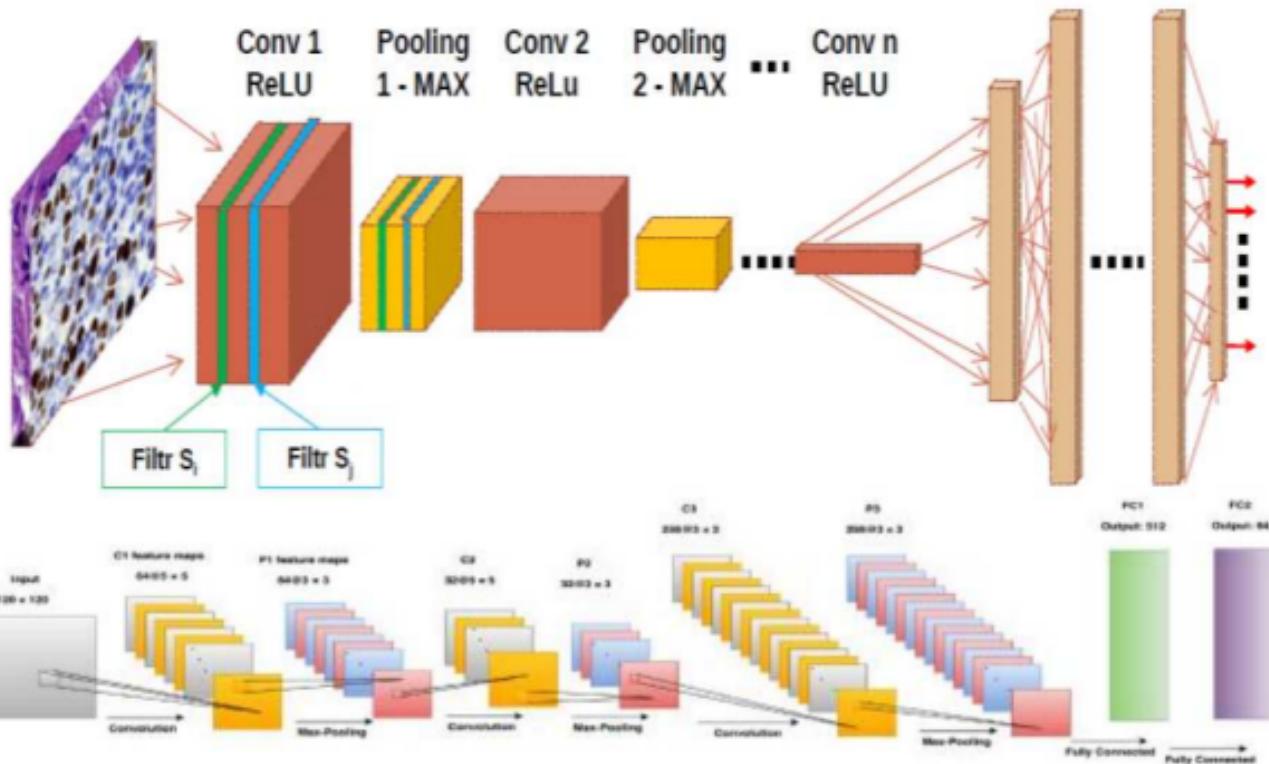
- Uczenie się sieci wymaga ogromnej ilości danych treningowych
- Uczenie się sieci wymaga dużej mocy obliczeniowej
- Proces uczenia się wymagają odpowiedniej architektury wydajnego systemu komputerowego mogą być złożone i często muszą być ścisłe dostosowane do konkretnej aplikacji (programowalne tablice bramek (FPGA), procesory graficzne (GPU) i specyficzne dla aplikacji układy scalone (ASIC)) są badane w celu wykorzystania równoległości struktury obliczeniowej sieci neuronowych, bardziej niż równoległe procesory)
- Powstałe **modele mogą nie być łatwo interpretowalne**
- Długi czas uczenia się (metoda prób i błędów),
- Występowanie nadmiernego dopasowania w przypadku niewystarczającej ilości/liczby danych.
- Trudności w zapewnieniu konwergencji procesu uczenia się.
- **Wymagają od użytkownika zrozumienia, jak należy interpretować wyniki oddawane przez wyuczony model zjawiska/wyuczoną sieć, aby wysunąć sprawdzalne hipotezy dotyczące badań**

Przykłady sieci

Sieci, ich autorzy i rok ich powstania

Year	CNN	Developed by	Place	Top-5 error rate	No. of parameters
1998	LeNet(8)	Yann LeCun et al			60 thousand
2012	AlexNet(7)	Alex Krizhevsky, Geoffrey Hinton, Ilya Sutskever	1st		60 million
2013	ZFNet()	Matthew Zeiler and Rob Fergus	1st		
2014	GoogLeNet(19)	Google	1st		4 million
2014	VGG Net(16)	Simonyan, Zisserman	2nd		138 million
2015	ResNet(152)	Kaiming He	1st		

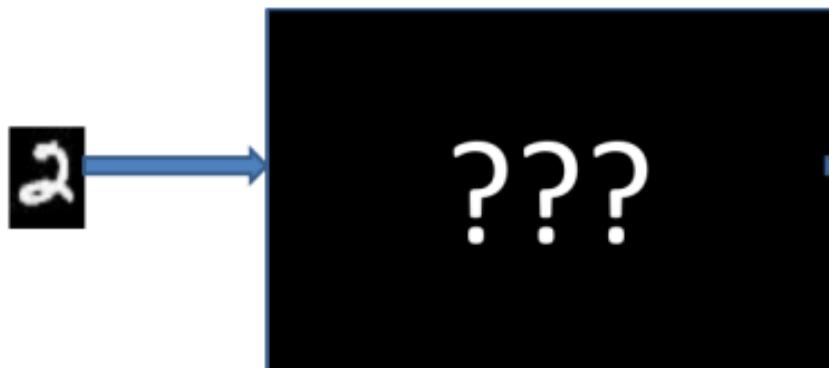
Podstawowa architektura sieci CNN



Rozpoznawanie cyfr zapisanych ręcznie

np.: na czekach - Case study

Jak z obrazu uzyskać informację o cyfrze?



- p(obiekt~0)
- p(obiekt~1)
- p(obiekt~2)
- p(obiekt~3)
- p(obiekt~4)
- p(obiekt~5)
- p(obiekt~6)
- p(obiekt~7)
- p(obiekt~8)
- p(obiekt~9)

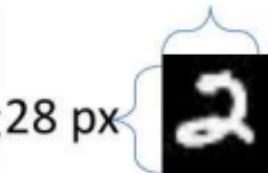
Baza obrazów MNIST

(Modified National Institute of Standards and Technology)



- 70 000 obrazów ręcznie pisanych cyfr
- 55 000 uczenie, 5 000 walidacja
- 10 000 testy
- 28x28 pikseli
- Wydany w 1998 roku
- 500 różnych osób
- Obraz + etykieta (jaka to cyfra)
- Skala kolorów znormalizowana do 0.0...1.0

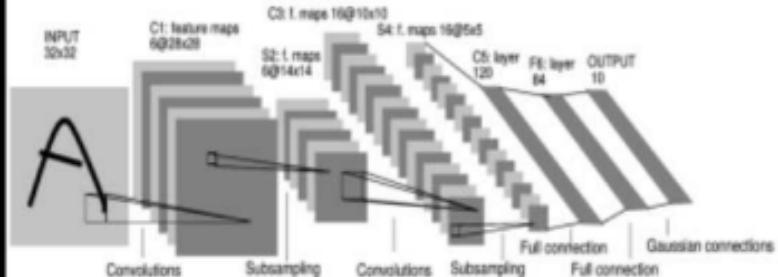
28 px $28 \times 28 = 784$ pikseli



28 px

Używamy do nauki:
THE MNIST DATABASE of
handwritten digits
<http://yann.lecun.com/exdb/mnist/>

Sieć LeNet



```
37 lenet5 = keras.models.Sequential()  
38  
39 lenet5.add(keras.layers.Conv2D(filters=6, kernel_size=(5, 5),  
40         activation='relu', input_shape=(32,32,1)))  
41 lenet5.add(keras.layers.AveragePooling2D())  
42 lenet5.add(keras.layers.Conv2D(filters=16, kernel_size=(5, 5),  
43         activation='relu'))  
44 lenet5.add(keras.layers.AveragePooling2D())  
45 lenet5.add(keras.layers.Flatten())  
46  
47 lenet5.add(keras.layers.Dense(units=120, activation='relu'))  
48 lenet5.add(keras.layers.Dense(units=84, activation='relu'))  
49 lenet5.add(keras.layers.Dense(units=10, activation = 'softmax')) # output
```

Model: "sequential_1"

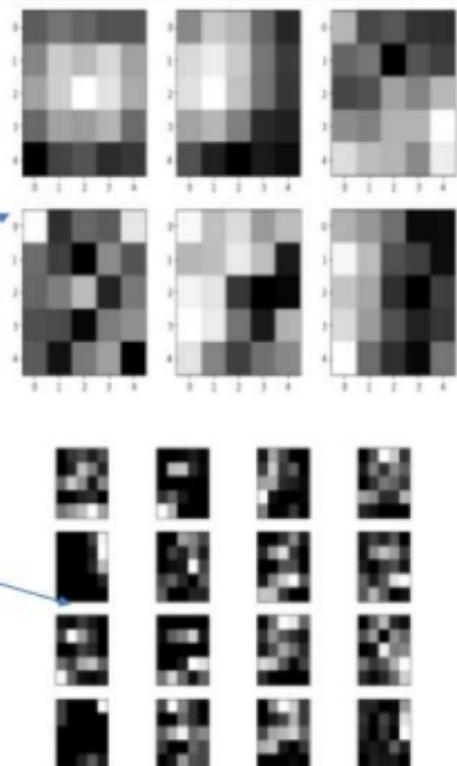
Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 28, 28, 6)	156
average_pooling2d_2 (AveragePooling2D)	(None, 14, 14, 6)	0
conv2d_3 (Conv2D)	(None, 10, 10, 16)	2416
average_pooling2d_3 (AveragePooling2D)	(None, 5, 5, 16)	0
flatten_1 (Flatten)	(None, 400)	0
dense_3 (Dense)	(None, 120)	48120
dense_4 (Dense)	(None, 84)	10164
dense_5 (Dense)	(None, 10)	850
Total params: 61,706		
Trainable params: 61,706		
Non-trainable params: 0		

Praca zaliczeniowa na APO
studenta mgr inż. Brylewa

Uczenie się sieci to optymalizacja ze względu na cel

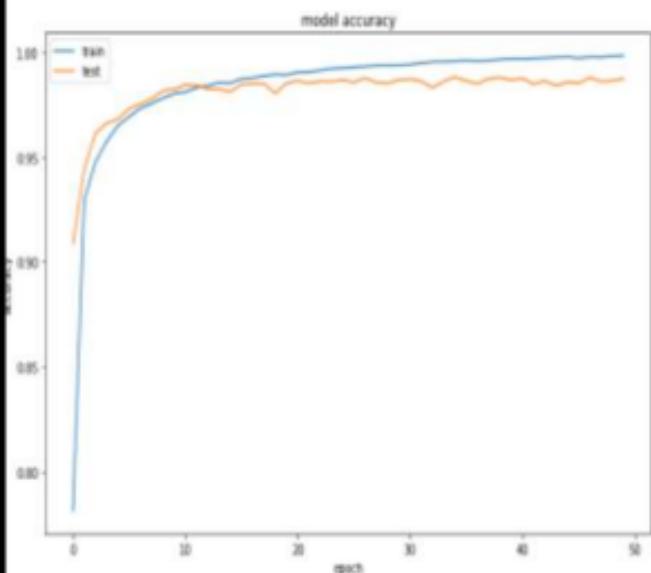
Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 28, 28, 6)	156
average_pooling2d_2 (Average)	(None, 14, 14, 6)	8
conv2d_3 (Conv2D)	(None, 10, 10, 16)	2416
average_pooling2d_3 (Average)	(None, 5, 5, 16)	8
flatten_1 (Flatten)	(None, 400)	0
dense_3 (Dense)	(None, 120)	48120
dense_4 (Dense)	(None, 84)	10164
dense_5 (Dense)	(None, 10)	850
<hr/>		
Total params: 61,706		
Trainable params: 61,706		
Non-trainable params: 0		

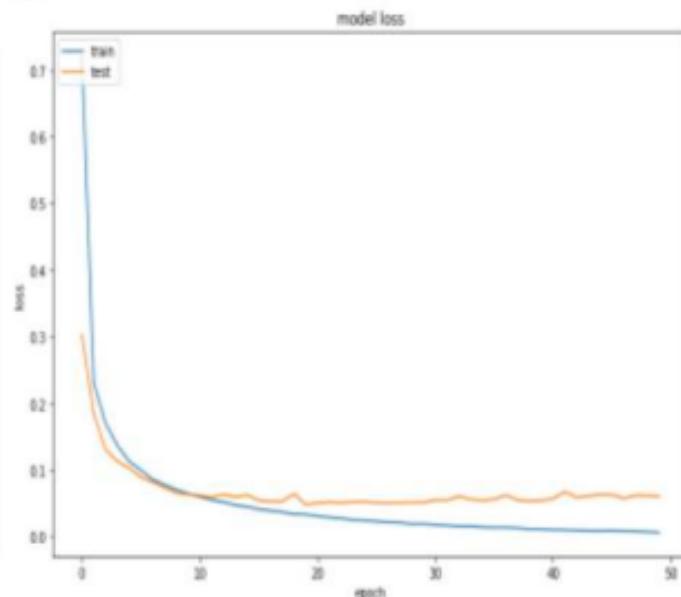


$[-0.0789532 \ -0.02418987 \ 0.10222547 \ 0.00066162 \ -0.04760809 \ 0.09279951 \ -0.01421791 \ -0.07095297 \ 0.05741353 \ -0.01717872]$

Parametry wyuczenia sieci

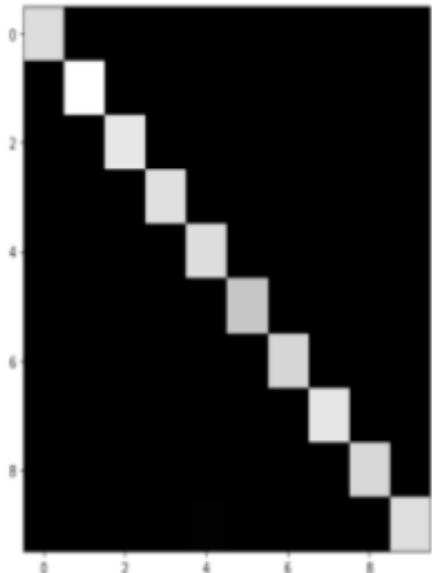


Najlepszego dopasowania



Minimalizacji błędu

Kontrola błędów



	0	1	2	3	4	5	6	7	8	9
0	974	0	0	0	0	0	2	1	3	0
1	0	1130	1	0	0	1	2	0	0	1
2	4	1	1020	0	3	0	0	2	2	0
3	1	0	2	993	0	6	0	4	2	2
4	0	0	0	0	976	0	2	2	0	2
5	2	0	0	7	0	875	2	2	3	1
6	2	2	0	0	1	5	947	0	1	0
7	0	3	4	1	1	0	0	1015	2	2
8	1	0	1	4	3	3	1	2	955	4
9	2	2	0	4	10	3	0	4	0	984

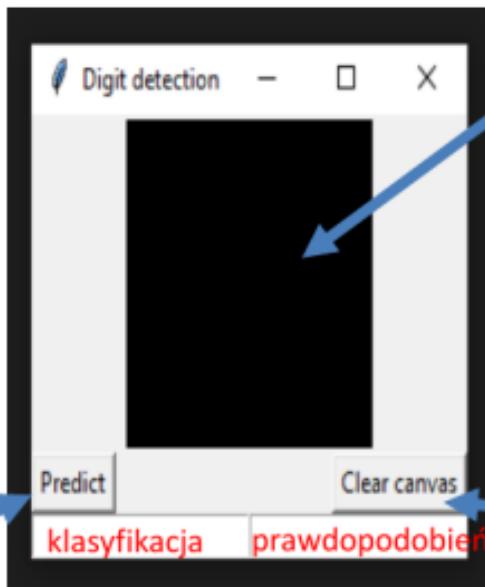
```
result = lenet5.evaluate(X_test, y_test)
dict(zip(lenet5.metrics_names, result))
```

```
313/313 [=====] - 0s 1ms/step - loss: 0.0490 - accuracy: 0.9869
```

```
{'loss': 0.04901135712862015, 'accuracy': 0.9868999719619751}
```

Wizualizacja macierzy pomyłek

Moduł wykonawczy do rozpoznawania znaków narysowanych przez użytkownika

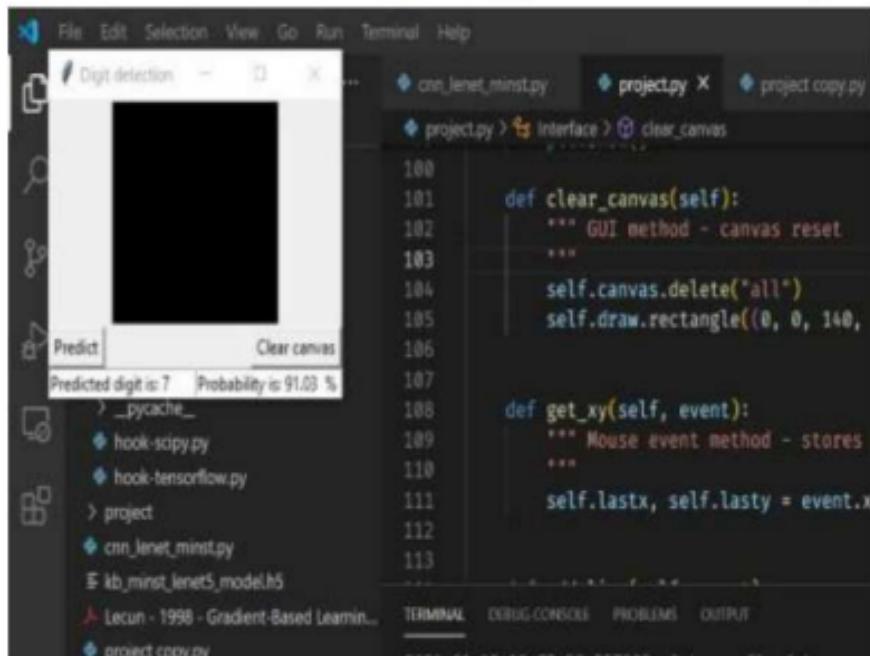


Zamiana obrazu na macierz 32x32 px, wysłanie do modelu, wyliczenia prawdopodobieństw przynależności do jednej z 10 klas i wybór prawdopodobieństwa maksymalnego

Powierzchnia na której można narysować cyfrę do rozpoznania

Czyszczenie powierzchni do rysowania

Działanie programu rozpoznającego



- Python 3.7
- Numpy
- Matplotlib
- Tensorflow 2.4 – wymaga AVX w zestawie instrukcji CPU!!!
- CUDA 10.1
- SciKit
- Tkinter - GUI

Moduł inferencji

U-net

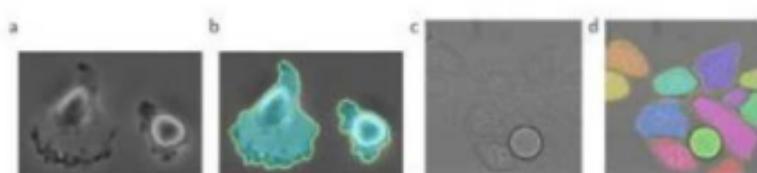
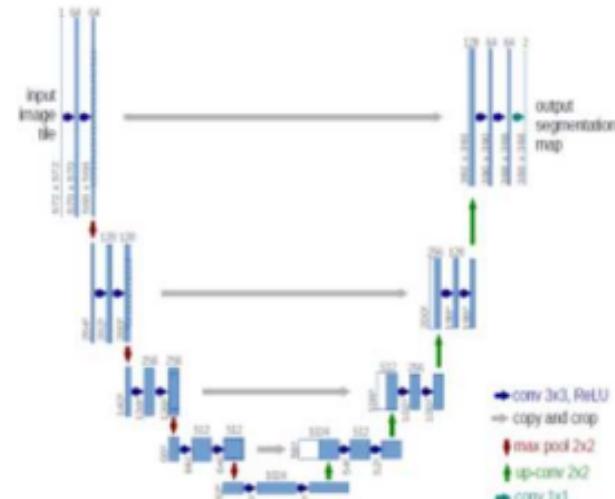
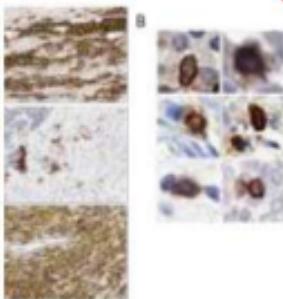
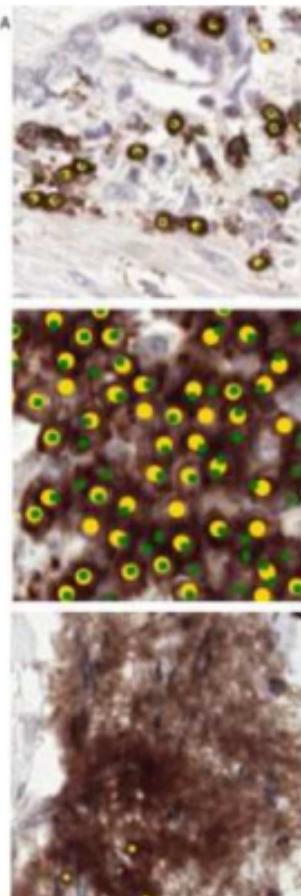
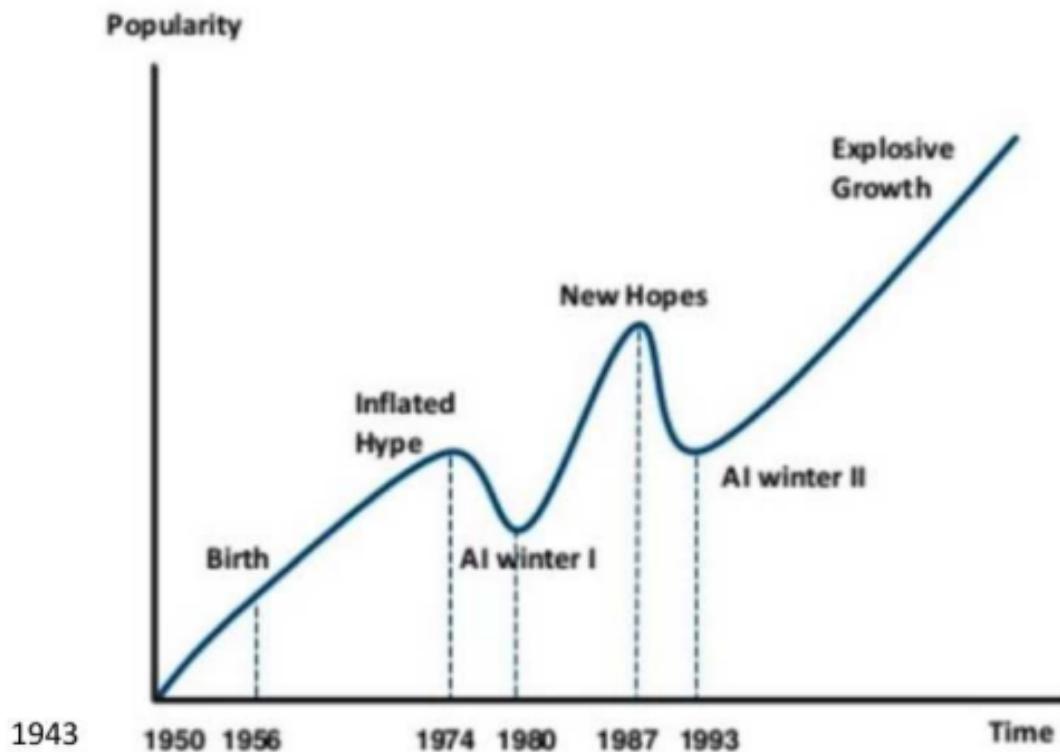


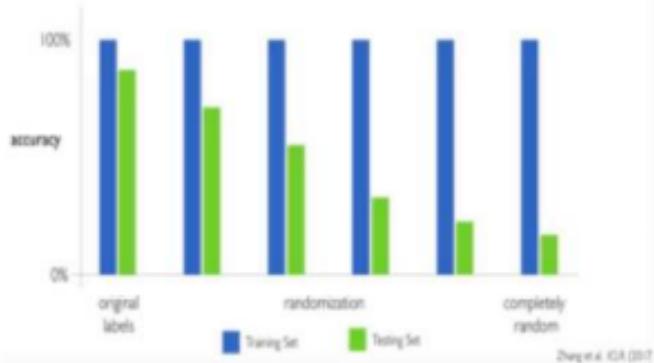
Fig. 4. Result on the ISBI cell tracking challenge. (a) part of an input image of the “PhC-U373” data set. (b) Segmentation result (cyan mask) with manual ground truth (yellow border) (c) input image of the “DIC-HeLa” data set. (d) Segmentation result (random colored masks) with manual ground truth (yellow border).

Historia i aktualne problemy NN



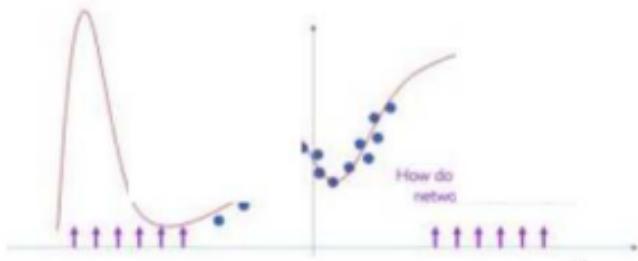
Ograniczenia DNN

- Można wytrenować sieć, która pasuje do kompletnie przypadkowych podziałów na klasy (na danych treningowych) da 100% dokładności
- Możemy wnioskować o funkcji aproksymującej jedynie w zakresie danych zbioru uczącego.



Zhang et al. ICLR 2017

Neural networks are excellent function approximators
...when they have training data

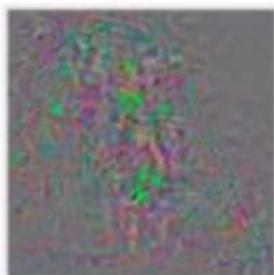


Ograniczenia DNN

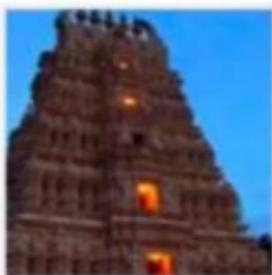
- Uwaga! na proces augmentacji bo czasem produkujemy „podobne przeciwnieństwa” (ang. *adversarial examples*)



Original image
Temple (97%)



Perturbations



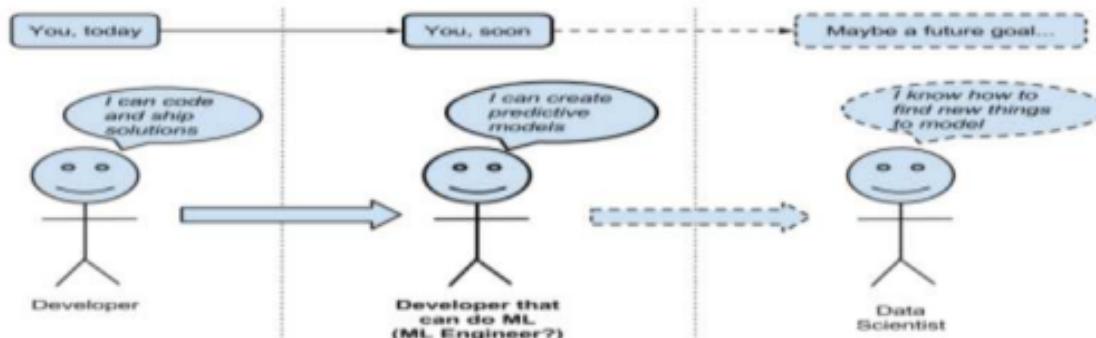
Adversarial example
Ostrich (98%)

Despois. "Adversarial examples and their implications" (2017).

Ograniczenia DNN

- Dużo danych bo dużo współczynników do optymalizacji
- Obliczeniowo i wdrożeniowo (GPU) kosztowny
- Mogą być oszukiwane przez „podobne przeciwnieństwa”
- Słabo można zbadać model, jego odporność i niepewność i inne parametry
- Model i oprogramowanie jest **czarną skrzynką** której trudno ufać – **szczególnie ekspertom/lekarzom**
- Nie każde podejście zawsze daje wynik w postaci dobrze optymalizowanej - potrzeba **prób i błędów**
- **Potrzebna wiedza eksperta** (wąskie gardło)

Praca dewelopera oprogramowania w dobie przyszłych mocno zaawansowanych sieci



Obawy społeczne związane z sztuczna inteligencją

- Przypominają obawy z XIX w sprawie **mechanizacji pracy** lub obawy o Puszkę Pandory
- **Czy technologia może coś chcieć**, być bezwzględnym megalomanem? **Nie ma woli i celu** – to ludzie mają chęć dominacji i gromadzenia zasobów i dokonują projekcji na AI
- Powody nieufności do technologii to głównie: **nieprzejrzystość** i brak dostępu do procesu ich uczenia się sieci (same wyszukują cechy ich zakresy i same się optymalizują wagi, a jak coś idzie nie tak to nadzorujący uczenie się maszyny właściwie nie wiedzą dlaczego, czy to można naprawić. Nadzorujący ponawia proces oparty na przypadkowym podawaniu przykładów ze zbioru uczącego).
- Zbiory uczące przygotowuje **człowiek** i to jego wiedza jest w nich zawarta i przez niego **wyselacjonowana** (postulat: **akredytacje dla deweloperów**).
- Technologia sztucznej inteligencji rozwijana się stopniowo, jest poprawiana pod względem bezpieczeństwa i skuteczności, projektowana, tak aby spełniać różne praktyczne kryteria
- Aktualne sukcesy AI wynikają z ich ogromnej siły obliczeniowej pozwalającej na przeanalizowanie ogromnej ilości danych uczących, czyli przypominanie sawantka czyli genialnego idiota – film „Rainmen” z Dusinem Hofmanem.
- Słowa Alisona Gopinka (University of Clifornia w Berkeley): **„Jak na razie znacznie więcej szkody narobić może naturalna głupota niż sztuczna inteligencja.”**

Przykłady sieci

STEGANOGRAFIA

STEGANOGRAFIA

Nauka o komunikacji w taki sposób, by obecność komunikatu nie mogła zostać wykryta, czyli **ukrywanie/hermetyzacji informacji** (ang. *information hiding*)

Słowo „steganografia” pochodzi z języka greckiego i oznacza ukryte pismo

Zaleta steganografii w stosunku do szyfrowania informacji jest ukrycie samego faktu porozumiewania się stron.

Ukrywanie informacji obrazowej (lub tekstowej) w obrazie
(image watermarking) w różnych celach:

Cele stosowania steganografii

1. Ochrona praw autorskich (ukryty obraz (lub tekst) pełni rolę znaku wodnego (*watermark*)), w tym przypadku konieczną własnością obrazu ukrytego jest (poza odpornością na usunięcie przez czynniki zewnętrzne) jego wystarczająca **odporność** na działanie standardowych operacji przetwarzania obrazów (filtracja, kompresja, zniekształcenia geometryczne itp.),
2. Ochrona autentyczności obrazu (ukryty obraz pełni rolę znaku wodnego (*watermark*)), w tym przypadku konieczną własnością obrazu ukrytego jest (poza odpornością na usunięcie przez czynniki zewnętrzne) jego wystarczająca **podatność** na działanie operacji przetwarzania obrazów (filtracja, kompresja, zniekształcenia geometryczne itp.).
3. Praktyczne przesyłania informacji

Historia steganografii

- Histajos w niewoli króla perskiego Dariusza postanowił przesłać informację do swego zięcia Arystagorasa z Miletu stosując następujący fortel: na wygolonej głowie swego orasa niewolnika **wytatuował informację**, gdy niewolnikowi odrosły włosy posłał go z oficjalnym, mało istotnym listem
- Egipcie i Chinach powszechnie stosowano **atrament sympathyczny**.
- Demaratus ostrzegł Grecję o ataku, pisząc go bezpośrednio na drewnianym podłożu woskowej tabletki przed nałożeniem powierzchni wosku pszczelego.
- **Mikrokropki:** Pierwsze próby miniaturyzowania przesyłanych informacji za pomocą mikrofotografii podjęto w 1871 r. podczas wojny francusko-pruskiej, kiedy to przesyłano raporty do oblężonego przez Niemców Paryża w postaci prostokątów o wymiarach 3 cm × 4 cm. Udoskonalenia tej techniki dokonał wywiad niemiecki Abwehra na początku lat 40. XX wieku.
- punkt o średnicy 1 mm wykonany przez specjalne urządzenie będące połączeniem aparatu fotograficznego i mikroskopu, zawierający zminiaturyzowane dane tekstowe lub fotografie i szczegółowe rysunki techniczne. Skala miniatyrizacji wynosi ok. 1:300, oznacza to możliwość pomniejszenia kartki formatu A4 zawierającej np. tekst do wielkości pojedynczej kropki znajdującej się w tekście pisanej czcionką normalnej wielkości listu.
- Współcześnie technika mikrokropiek (wykonywanych w tysiącach sztuk w technice grawerowania laserowego) jest wykorzystywana także komercyjnie np. do zabezpieczania żetonów w kasynach przed podrabianiem, znakowania samochodów lub cennych przedmiotów.
- **Zapisy kodem pod znaczkiem na liście**
- Podczas II Wojny Światowej Velvalee Dickinson nazywana **Doll Woman** japoński szpieg w Nowym Yorku z lalkami którymi handlowała przesyłała informacje szpiegowskie.
- **Jeremiah Denton** wielokrotnie mrugał oczami w alfabetie Morse'a podczas telewizyjnej konferencji prasowej w 1966 roku, po porwaniu przez porywaczy z Wietnamu Północnego, wypowiadając „T-O-R-T-U-R-E”.
- W 1968 r. członkowie **załogi statku wywiadowczego USS Pueblo**, przetrzymywani przez Koreę Północną jako więźniowie, komunikowali się w języku migowym na zdjęciach.
- Obraz Koranu miał zaszyfrowane informacje o zamachu na World Trade Center

Ukrywanie obrazu w obrazie - nazwy obrazów

1. obraz ukrywający [**p**]
2. obraz ukrywany [**h**]
3. obraz ukrywany po przekształceniu [**hmod**]
4. obraz ukrywający wraz z obrazem ukrytym [**ph**]

Obrazy ukrywające z gradacją poziomów szarości

Obraz ukrywający: na 1 piksel obrazu ukrywanego przypada 1 bajt (8 bitów)

Obraz ukrywany może i powinien być uproszczony: na 1 piksel obrazu może przypadać 1, 2, 3, 4, 5, 6, 7, 8 bitów, co odpowiada 2,4,8,16,32,64,128,256 poziomom szarości (**M**)

W praktyce, ze względu na potrzebę ograniczenia wpływu obrazu ukrywanego na wygląd obrazu ukrywającego, stosowane są wartości **M<<256**, a obraz ukrywany zapisywany jest na najmniej znaczących bitach obrazu ukrywającego.

Zajęcia badanie czy:

przy wzroście wartości **liczby bitów obrazu ukrywanego** następuje coraz większa zmiana wyglądu obrazu ukrywającego i jednocześnie coraz wyraźniejsze uwidocznienie obrazu ukrywanego w tym obrazie.

Obrazy ukrywające kolorowe np. w formacie RGB

Obraz ukrywający: na 1 piksel obrazu przypada 3 bajty (24 bity) odpowiadające 3 składowym R, G, B

Przykładowy zapis piksla obrazu ukrywanego: 2 najmłodsze bity składowej R oraz po jednym najmłodszym bicie składowych G i B (razem 4 bity, co odpowiada obrazowi o $M=16$ poziomach)

Zapis informacji tekstowej obrazie RGB: 1 znak – 8 bitów co oznacza że do zapisu 1 znaku można wykorzystać 2 piksle.

**Dodatkowe kodowanie obrazów ukrywanych => lepsze ich ukrycie
(np. przemieszanie poszczególnych piksli)**

Operacje najczęściej stosowane w procesie ukrywania i odtwarzania obrazu

- Jednopunktowe jednoargumentowe**

- progowania, redukcji poziomów szarości, rozciągania, uniwersalne operacje punktowe (UOP)

- Jednopunktowe dwuargumentowe**

- arytmetyczne (dodawanie, odejmowanie)
 - logiczne (suma (OR), iloczyn (AND))

Koniec wykładu



Egzamin będzie zdalny

Terminy egzaminu: 02.07.2021 i 09.07.2021

- Egzamin będzie na każdym etapie egzaminem testowym zdalnym.
- Będzie wykonywany przy pomocy oprogramowania **Inspera**.
- Kursy dla studentów na temat tego oprogramowania już się odbyły, ale można je jeszcze raz przejść rozwiązuając udostępnione testy.

↶ ↷ | **B** *I* U \times_r \times^2 | $\frac{\partial}{\partial x}$ $\frac{\partial^2}{\partial x^2}$ | \otimes \otimes^2 | \odot Σ | % | =

+ Wstaw element

Napisz, co to jest obraz cyfrowy.



- Punkty
- Główna ilustracja
- Panel PDF
- Etykiety

Napisz, co to jest obraz cyfrowy.



Outlook Web App x Tłumacz Google x Inspera Assessment x obraz cyfrowy - Szukaj w Google x +

← → C google.com/search?q=obraz+cyfrowy&rlz=1C1N&XL_plPL771PL771&oiq=obraz+cyfrowy&oq=chrome_6957j0l7.5430j08&sourceid=chrome&ie=UTF-8

Aplikacje Chrome Citibank Online - M... Credit Agricole Ban... PRACA & GRANTY... UBBK | Uczelniowy B... Home - Lumosity zdrofit Tłumacz Google

Google

obraz cyfrowy

X 🔍

Wszystko Grafika Zakupy Wiadomości Filmy Więcej Ustawienia Narzędzia

Około 47 800 000 wyników (0.56 s)

Obraz cyfrowy (ang. digital image) – dwuwymiarowy obraz zapisany w postaci binarnej, możliwy do przechowywania w pamięci cyfrowej. Pojęcie najczęściej dotyczy obrazu zapisanego w postaci rastrowej, w której jest on przedstawiony jako dwuwymiarowa macierz pikseli.

pt.wikipedia.org › wiki › Obraz_cyfrowy ↗

Obraz cyfrowy – Wikipedia, wolna encyklopedia

Informacje o fragmentach z odpowiedzią Przeslij opinię

www.szkolazpasia.pl > rastrowa > istota-obrazu-cyfrowego

ISTOTA OBRAZU CYFROWEGO - Informatyka Jacek Turlejski
ISTOTA OBRAZU CYFROWEGO 3.1. PIKSEL. Obrazy cyfrowe, odzwierciedlające oryginały, noszą także nazwę obrazów bitmapowych, map bitowych lub po ...

www.zsk.ict.pwr.wroc.pl > zsk > loc > in_obi_wyk_1 ↗ PDF

Obraz cyfrowy - Politechnika Wrocławska
Jak wyrazić cechy obrazu przy pomocy liczb? • Podstawy przetwarzania obrazów cyfrowych. Jacek Jarnicki Politechnika Wrocławska. 3. 1. Obraz cyfrowy i jego ...

https://allegro.pl/oferta/120x80-duzy-obraz-cyfrowy-natura-89737630117utm_feed=aa34192d-eee2-4419-9a9a-de66b9dfac28

Zobacz: obraz cyfro...

Hama Slimline Basic 8" Cyfrowa ramka do zdjęć 229,00 zł RTV Euro AGD Przez: PMshopping

Graffiti B... Obraz cy... 194,00 zł artspesaz...

T 7" 8" 10" mm/18,5/20,3/25,4 mm/45/50/63 mm/18/20/25 mm/45/50/63

120x80 Du... Cytomat N...

PŁENNY OPIS EN 0...dock PŁENNY OPIS EN 0...dock



+ Wstaw element

Co to jest histogram obrazu cyfrowego?

Histogram - to

wykres słupkowy, słupkowy wykres, wykres kolumnowy, kolumnowy wykres, graficzny sposób, opis statystyczny

przedstawiający liczbę wystąpień, częstość wystąpień, ilość pikseli, rozkład empiryczny, rozkład wszystkich poziomów szarości potencjalnie występujących w obrazie.

- Punkty
- Główna ilustracja
- Panel PDF
- Etykiety
- Zaawansowane

def histogramu

Co to jest histogram obrazu cyfrowego?

Histogram - to to wykres zawierający 256 punktów na poziomej osi, które określają jasność pikseli na zdjęciu.



(wykres słupkowy, słupkowy wykres) przedstawiający w zakresie [0,255]



(liczbę wystąpień, częstość wystąpień) wszystkich poziomów szarości potencjalnie występujących w obrazie .

Histogram

[edytuj]

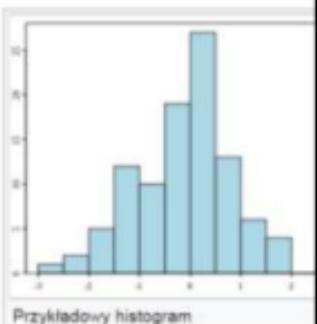
Histogram – jeden z graficznych sposobów przedstawiania rozkładu empirycznego cechy. Składa się z szeregu prostokątów umieszczonych na osi współrzędnych. Prostokąty te są z jednej strony wyznaczone przez przedziały klasowe (patrz: szereg rozdzielczy) wartości cechy, natomiast ich wysokość określona przez liczebności (lub częstotliwości, ewentualnie gęstość prawdopodobieństwa) elementów wpadających do określonego przedziału klasowego.^[1] Jeśli histogram pokazuje liczebności, a nie gęstość prawdopodobieństwa, wówczas szerokości przedziałów powinny być równe.

Interpretacja histogramu

[edytuj] [edytuj kod]

Przykładowa interpretacja histogramu z rysunku obok:

- luka w histogramie:
 - podejrzenie nieprawidłowego odczytu (brak danych),
 - podejrzenie błędu urządzenia pomiarowego.
- histogram z dwoma wierzchołkami:
 - tzw. rozkład dwumodalny, który powstaje często, gdy badana populacja jest połączeniem dwóch odseparowanych populacji, np. połączono wyroby z różnych procesów wytwórczych.



Zobacz też [edytuj] [edytuj kod]

sample.wig

PEŁNY OPS EN 0...doc



Pokaż w

← → | B I U ×, ×² | ← → | 0 Σ | ↻ | ← → |

Wstaw element

Na obrazie o M = 6 i N = 4

1	2	3	4
1	0	0	0
0	5	4	1
1	4	3	0
0	0	0	0

wykonaj operację progowania według progu th = 2. Zapisz operację w postaci tablicy LUT:

0	1	2	3	4	5
0,0	0,0	0,0	5,5	5,5	5,5

oraz obraz wynikowy

0,0	0,0	0,0	0,0
0,0	5,5	5,5	0,0
0,0	5,5	5,5	0,0
0,0	0,0	0,0	0,0

Progowanie

Jest to taka wersja operacji zmniejszenia ilości poziomów szarości do dwóch, dla której istnieje możliwość arbitralnego wyboru wartości progu (p_1) czyli szarości granicznej, od której przyporządkowujemy wyższy poziom szarości (najczęściej biały) i poniżej której przyporządkowujemy niższy próg szarości (najczęściej czarny).

$$q = \begin{cases} L_{\min} & \text{dla } p \leq p_1 \\ L_{\max} & \text{dla } p > p_1 \end{cases}$$

Punkty

Główna ilustracja

Panel PDF

Etykiety

Punkty

Główna ilustracja

Panel PDF

Etykiety

Na obrazie o M = 6 i N = 4

1	2	3	4
5	5	4	5
4	0	0	5
4	0	1	5
4	5	5	5

wykonaj operację progowania binarnego z progiem $th = 3$. Zapisz operację przy pomocy tablicy LUT:

0	1	2	3	4	5
0,0	0,0	0,0	0,0	1,1	1,1

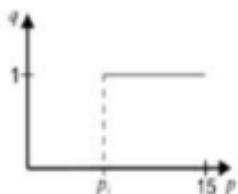
oraz obraz wynikowy:

5,5	5,5	5,5	5,5
5,5	0,0	0,0	5,5
5,5	0,0	0,0	5,5
5,5	5,5	5,5	5,5

Operacja progowania (binaryzacji)

$$p_1 = 5$$

$$q = \begin{cases} 0 & \text{dla } p \leq p_1 \\ 1 & \text{dla } p > p_1 \end{cases}$$



15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

1	1	0	0	0
1	1	1	0	0
0	0	1	1	1
0	0	0	0	0
1	1	1	1	1

[Punkty](#)
[Główna ilustracja](#)
[Panel PDF](#)
[Etykiety](#)

Na obrazie o $M = 6$ i $N = 4$

1	2	3	4
1	0	0	2
0	5	4	1
1	4	3	0
0	0	0	0

wykonaj operację progowania odwrotnego z progiem $th = 2$. Zapisz tę operację w postaci tablicy LUT:

0	1	2	3	4	5
5.5	5.5	5.5	0.0	0.0	0.0

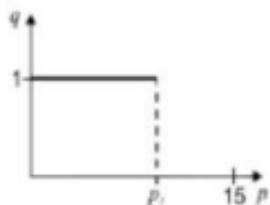
oraz obraz wynikowy

5.5	5.5	5.5	5.5
5.5	0.0	0.0	5.5
5.5	0.0	0.0	5.5
5.5	5.5	5.5	5.5

Operacja odwrotnego progowania (binaryzacji)

$$p_1 = 5$$

$$q = \begin{cases} 1 \text{ dla } p \leq p_1 \\ 0 \text{ dla } p > p_1 \end{cases}$$



15	15	0	0	2
13	13	15	0	0
0	0	7	14	14
0	1	2	3	4
15	14	13	12	11

0	0	1	1	1
0	0	0	1	1
1	1	0	0	0
1	1	1	1	1
0	0	0	0	0

Jak rozwiązywać zadania obliczeniowe?

- Zadania z liczeniem odległości

Metryka Euklidesowa

$$\rho_1(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{v=1}^n (x_v^\mu - x_v^\eta)^2}$$



Metryka uliczna (Manhattan, city block distance):

$$\rho_2(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n |x_v^\mu - x_v^\eta|$$

Metryka Czebyszewa (maksymalna)

$$\rho_3(\underline{x}^\mu, \underline{x}^\eta) = \max_{1 \leq v \leq n} |x_v^\mu - x_v^\eta|$$

Jak rozwiązywać zadania problemowe?



- Czym różni się akwizycja w przypadku tych trzech obrazów dloni?
- Czym różni się zapis informacji o obrazie dla tych trzech obrazów dloni?



Życzę powodzenia na egzaminie