



Monitorowanie i Diagnostyka w Systemach Sterowania

**Wydział Elektrotechniki i Automatyki
Katedra Elektrotechniki, Systemów Sterowania i Informatyki
Dr hab. inż. Michał Grochowski**



Monitorowanie i Diagnostyka w Systemach Sterowania

na studiach II stopnia specjalności: Systemy Sterowania i Podejmowania Decyzji

Rozpoznawanie wzorców i twarzy

w wykładzie wykorzystano materiały z następujących źródeł:

Blog: Data Science and Robots by Brandon Rohrer, <http://brohrer.github.io/blog.html>

Błaszowski P. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych.

Praca inżynierska, Politechnika Gdańska 2012. Promotor dr inż. Michał Grochowski

Błaszowski P. Wykrywanie, rozpoznawanie i śledzenie ruchomych obiektów poprzez niezależną platformę monitorującą, przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, Politechnika Gdańska 2013.

Promotor dr inż. Michał Grochowski

CS231n: Convolutional Neural Networks for Visual Recognition, 2016. Stanford University

Sikora M. Inteligentny system rozpoznawania twarzy w czasie rzeczywistym – implementacja sprzętowa.

Praca magisterska, Politechnika Gdańska 2012. Promotor dr inż. Michał Grochowski

Wenecki M. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, Politechnika Gdańska 2008. Promotor dr inż. Michał Grochowski

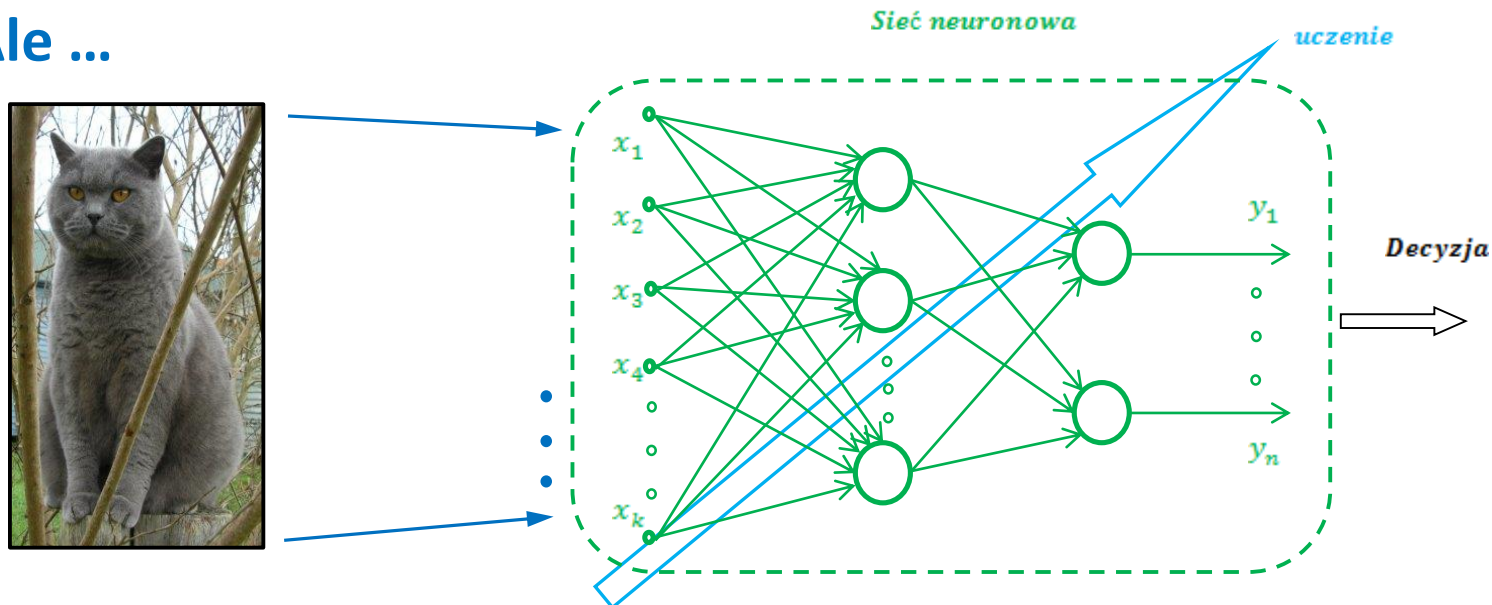
Opracował: dr inż. Michał Grochowski

Wprowadzenie

Rozpoznawanie/klasyfikacja wzorców

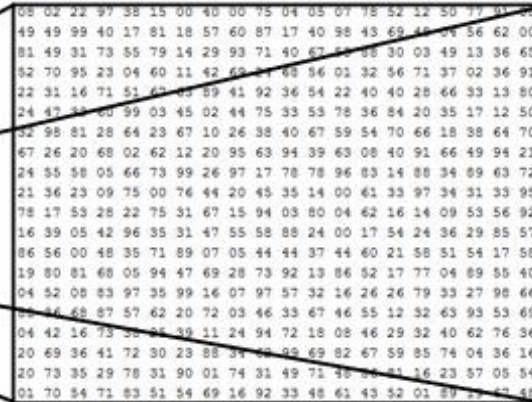
Zadanie sprowadza się do podania obrazów na wejście dedykowanego klasyfikatora np. sieci neuronowej czy SVM, odpowiednim jego nauczaniu (tryb off line) i wykorzystaniu do bieżącego rozpoznawania/identyfikacji.

Ale ...



Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców



What the computer sees

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców



05	02	22	97	38	15	00	40	00	75	04	05	07	78	52	12	50	77	90	09
49	49	99	40	17	81	18	57	60	87	17	40	98	43	69	40	04	56	62	00
81	49	31	73	55	79	14	29	93	71	40	67	53	88	30	03	49	13	36	65
52	70	95	23	04	60	11	42	65	41	69	56	01	32	56	71	37	02	36	91
22	31	16	71	51	65	33	59	41	92	36	94	22	40	40	28	66	33	13	80
24	47	33	60	99	03	45	02	44	75	33	53	78	36	84	20	35	17	12	50
32	98	81	28	64	23	67	10	26	38	40	67	59	54	70	66	18	38	64	70
67	26	20	68	02	62	12	20	95	63	94	39	63	08	40	91	66	49	94	21
24	55	58	05	66	73	99	26	97	17	78	78	96	83	14	88	34	89	63	72
21	36	23	09	75	00	76	44	20	45	35	14	00	61	33	97	34	31	33	95
78	17	53	28	22	75	31	67	15	94	03	80	04	62	16	14	09	53	56	92
16	39	05	42	96	35	31	47	55	58	88	24	00	17	54	24	36	29	85	57
86	56	00	48	35	71	89	07	05	44	44	37	44	60	21	58	51	54	17	58
19	80	81	68	05	94	47	69	28	73	92	13	86	52	17	77	04	89	55	40
04	52	08	83	97	35	99	16	07	97	57	32	16	26	26	79	33	27	98	66
17	46	68	87	57	62	20	72	03	46	33	67	46	55	12	32	63	93	53	69
04	42	16	73	35	85	39	11	24	94	72	18	08	46	29	32	40	62	76	36
20	69	36	41	72	30	23	88	34	68	89	69	82	67	59	85	74	04	36	16
20	73	35	29	78	31	90	01	74	31	49	71	48	84	11	16	23	57	05	54
01	70	84	71	83	51	54	69	16	92	33	48	61	43	52	01	89	17	67	48

What the computer sees

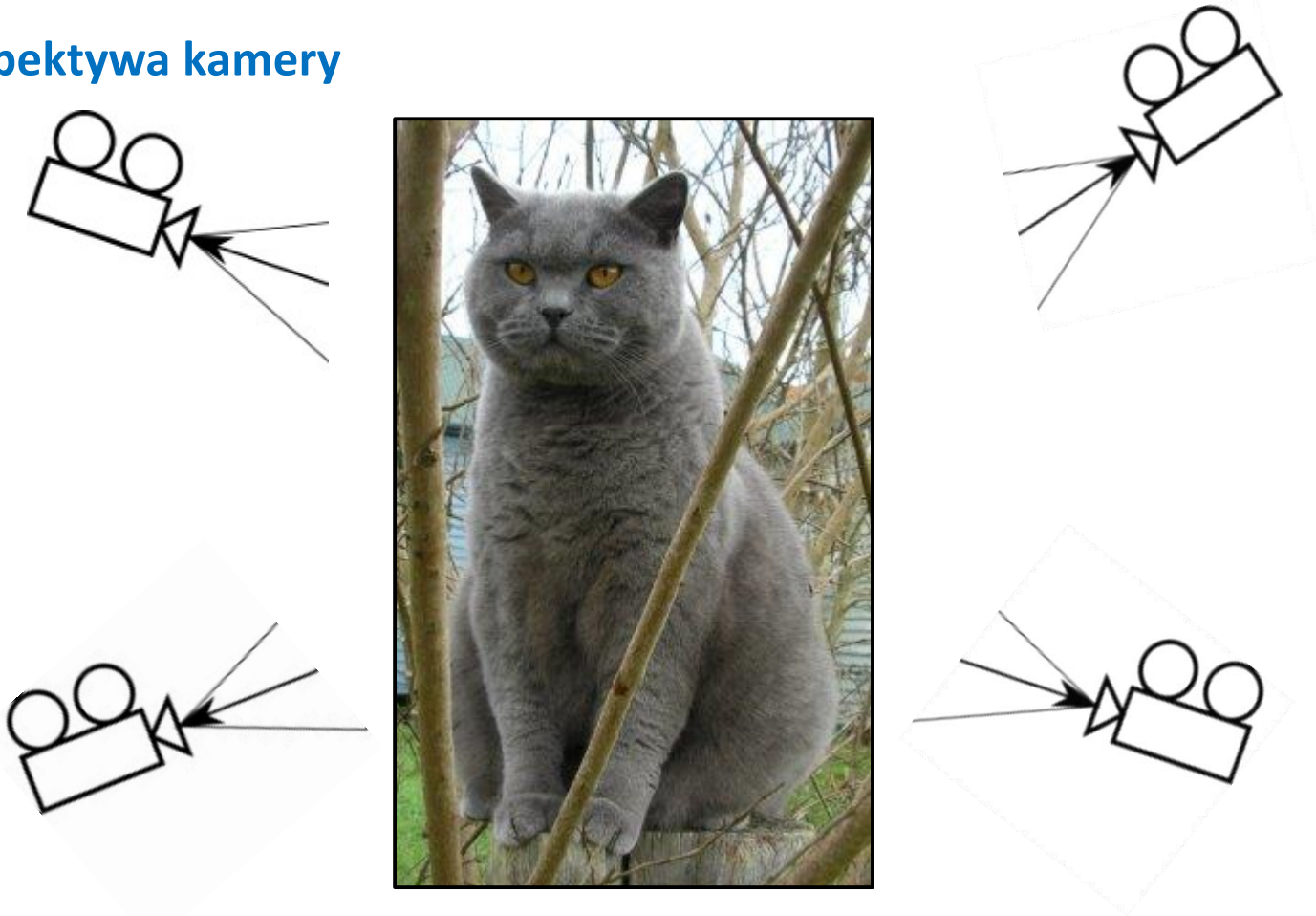
- zdjęcia np. w standardzie RGB, więc:
- 3 macierze, każda o rozmiarze np.: 1280x1024 pikseli co daje: $3 \times 1\,280 \times 1\,024 = 3\,932\,160$!!!;
- w każdej z macierzy zakodowane kolory w postaci wartości z zakresu 0-255;
- kamera przesyła próbki np. 60 klatek na sekundę.

Powstaje zadanie praktycznie **numerycznie nierozwiązywalne**, o nikłej przydatności dla użytkownika

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- Perspektywa kamery



źródło: CS231n: Convolutional Neural Networks for Visual Recognition, 2016. Stanford University

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- Oświetlenie



źródło: CS231n: Convolutional Neural Networks for Visual Recognition, 2016. Stanford University

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- „Zniekształcenia”



Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- „Przeszkody”



Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- Tło



źródło: CS231n: Convolutional Neural Networks for Visual Recognition, 2016. Stanford University

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

- **Różnorodność**



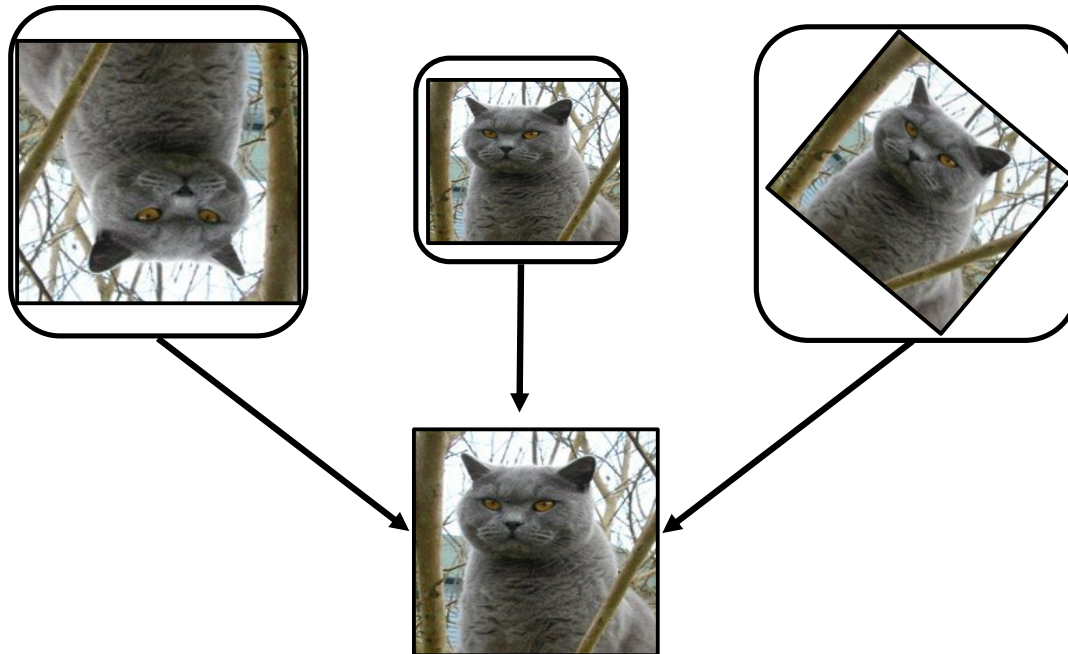
źródło: CS231n: Convolutional Neural Networks for Visual Recognition, 2016. Stanford University

Wprowadzenie

Wyzwania (problemy) związane z rozpoznawaniem/klasyfikacją wzorców

Ponadto ...

- obrazy mogą być różnych rozmiarów (różna odległość obiektów od kamery);
- obrazy mogą być w różnej orientacji.



Konieczne staje się odpowiednie przetwarzanie wstępne !

Przetwarzanie wstępne obrazów

Przestrzeń barw

- Model RGB;
- Model HSV;
- Model CMY (CMYK);
- Model skali szarości;
- Model binarny;
- I inne ...

Przetwarzanie wstępne obrazów

Przestrzeń barw

- Model RGB
(Red – czerwony, Green – zielony, Blue – niebieski)

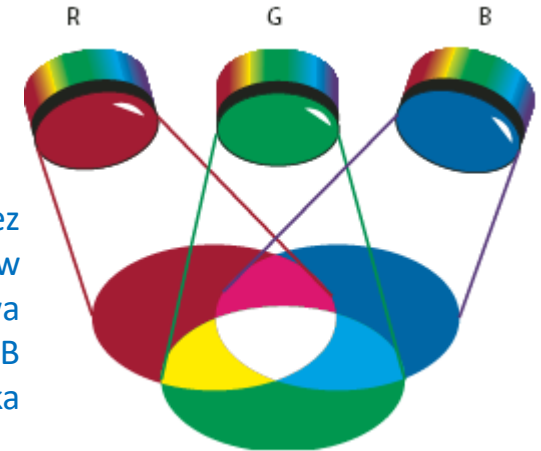
Dużą część kolorów z widzialnego widma światła można otrzymać poprzez łączenie ze sobą w różnych proporcjach i nasyceniach trzech kolorów podstawowych: czerwonego, zielonego i niebieskiego. Kolory RGB nazywa się *addytywnymi*, ponieważ po połączeniu wszystkich składowych RGB powstaje biel. Oznacza to, że całe światło odbijane jest z powrotem do oka obserwatora.

Tryb kolorów RGB (oparty na modelu RGB) umożliwia korzystanie z wartości kolorów. Każdej ze składowych RGB danego koloru są przypisywane określone wartości natężenia z zakresu od 0 (czerni) do 255 (biel).

Na przykład, przypisując składowej R liczbę 246, składowej G liczbę 20 i składowej B liczbę 50, uzyskuje się kolor jasnoczerwony.

Kiedy wartości trzech składowych są sobie równe powstaje odcień szarości.

Gdy wszystkie składowe mają wartość 255, powstaje czysta biel, a składowe równe 0 dają czystą czerni.



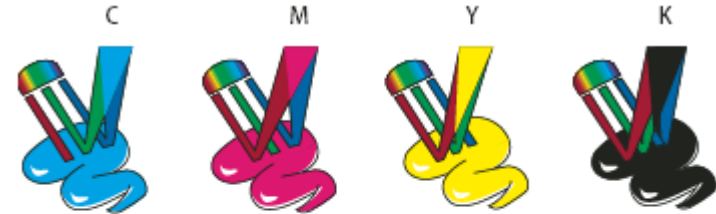
źródło:

http://help.adobe.com/pl_PL/illustrator/cs/using/WS714a382cdf7d304e7e07d0100196cbc5f-6295a.html

Przetwarzanie wstępne obrazów

Przestrzeń barw

- Model CMY (CMYK)
(Cyan – cyjan, Magenta – magenta,
Yellow – żółty, blackK – czarny)



Model RGB jest oparty na prawach rządzących emisją światła, model CMYK opiera się na prawach dotyczących absorpcji fal świetlnych. W przypadku drukarek, na przykład, chodzi o absorpcję światła przez farby naniesione na papier. Gdy białe światło pada na półprzezroczyste farby, część jego widma jest pochłaniana. Te barwy, które nie zostały zaabsorbowane, odbijane są z powrotem do oka.

Zgodnie z zasadami optyki połączenie czystych pigmentów koloru niebieskozielonego (C, od ang. “cyan”), karmazynowego (M, od ang. “magenta”) i żółtego (Y, od ang. “yellow”) powinno skutkować absorpcją całego światła, a więc dawać czystą czerń. Z tego względu wymienione kolory nazywa się *kolorami subtraktywnymi*. Aby uzyskać prawdziwą czerń, trzeba użyć czwartej farby – czarnej (K). Litera *K* służy do oznaczania czerni dlatego, że litera *B* (ang. black – czerń) jest już używana jako symbol barwy niebieskiej (ang. blue).

źródło:

http://help.adobe.com/pl_PL/illustrator/cs/using/WS714a382cdf7d304e7e07d0100196cbc5f-6295a.html

Przetwarzanie wstępne obrazów

Przestrzeń barw

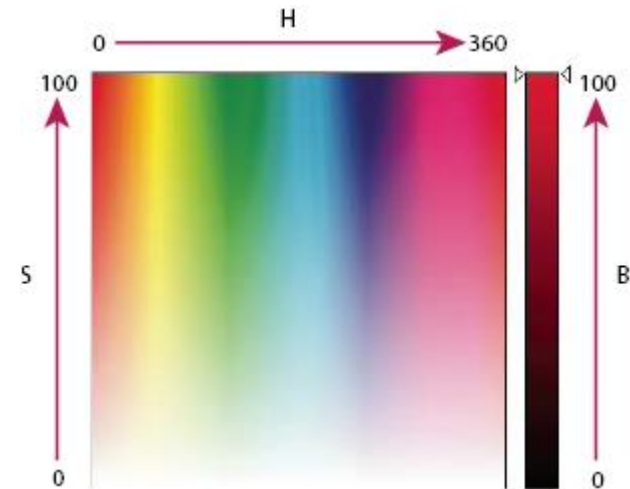
- Model HSV zwany także HSB (Hue – barwa/odcień), Saturation – nasycenie, Value – wartość, Brightness – jasność)

Model HSB powstał w oparciu o sposób postrzegania barw przez człowieka i opisuje trzy główne cechy koloru:

Barwa to właściwy kolor, powstający wskutek odbicia lub przejścia światła przez obiekt. Miarą barwy jest jej położenie na standardowym kole kolorów, wyrażone w stopniach od 0° do 360°. W zwykłym języku barwa jest określana po prostu nazwą odpowiedniego koloru, np. czerwony, pomarańczowy czy zielony.

Nasycenie to stopień czystości barwy (nazywany czasami *chrominancją*). Nasycenie reprezentuje intensywność szarości w stosunku do samej barwy i jest mierzone jako wartość procentowa, gdzie 0% oznacza barwę szarą, a 100% barwę całkowicie nasyconą. Na standardowym kole kolorów nasycenie rośnie od środka ku krawędziom.

Jasność jest to względna jaskrawość koloru, mierzona jako wartość procentowa, gdzie 0% odpowiada czerni, a 100% bieli.



źródło:

http://help.adobe.com/pl_PL/illustrator/cs/using/WS714a382cdf7d304e7e07d0100196cbc5f-6295a.html

Przetwarzanie wstępne obrazów

Przestrzeń barw

- Model skali szarości



źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Przetwarzanie wstępne obrazów

Przestrzeń barw

- Model binarny



źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia punktowe i arytmetyczne;
- Przekształcenia geometryczne;
- Przekształcenia morfologiczne;
- Filtracja.

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia punktowe i arytmetyczne (sumowanie obrazów)



obrazy sumowane

suma obrazów

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia punktowe i arytmetyczne (odejmowanie obrazów)



obrazy odejmowane

różnica obrazów

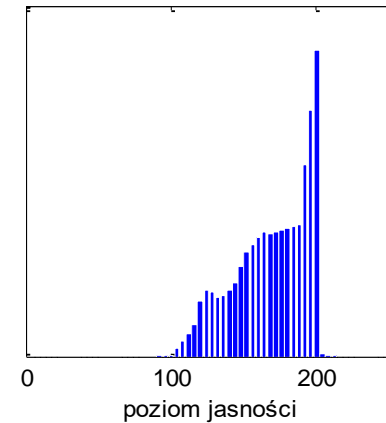
Przetwarzanie wstępne obrazów

Przekształcenia obrazów

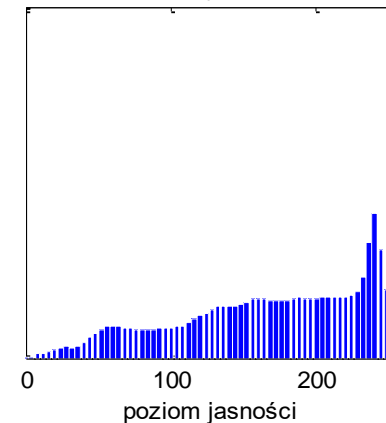
- Przekształcenia punktowe i arytmetyczne (liniowe)



obraz oka



rozjaśniony obraz oka



źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia punktowe i arytmetyczne (nieliniowe)



logarytmiczna operacja
punktowa



ekspotencjalna
operacja punktowa

źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

Przekształcenia geometryczne

- Obrót o kąt
- Skalowanie
- Normalizacja



Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia morfologiczne (dylatacja, erozja, otwarcie i zamknięcie)

Erozja

W dużym uproszczeniu erozję jest metodą zmniejszenia pola figur znajdujących się na obrazie poprzez odcięcie od każdej jego krawędzi, pasa o pewnej szerokości. Erozja usuwa z obrazu obszary odchylone od stosowanego szablonu. Efektem może być wygładzenie obiektów



obraz przed erozją



obraz po erozji

źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia morfologiczne (dylatacja, erozja, otwarcie i zamknięcie)

Dylatacja

Działanie dylatacji jest odwrotne do działania erozji. Upraszczając, można stwierdzić że dylatacja dopisuje pas o pewnej szerokości wokół obszarów wykrytych na obrazie.



obraz przed dylatacją



obraz po dylatacji

źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Przetwarzanie wstępne obrazów

Przekształcenia obrazów

- Przekształcenia morfologiczne (dylatacja, erozja, otwarcie i zamknięcie)
Erozja i dylatacja powodują zmiany pola obiektu, co nie jest korzystne w procesach analizy. Obie operacje wykorzystuje się często razem. Wykonanie na obrazie erozji, a następnie dylatacji powinno zachować podobne rozmiary obiektów, dając szereg zalet takich jak wygładzenie, usunięcie odchyłek, czy rozdzielenie nałożonych na siebie fragmentów. Działanie takie nazywa się **otwarcie**. Operacja odwrotna (dylatacja + erozja) nazywa się **zamknięciem**)



obraz przed otwarciem



obraz po otwarciu

źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Przetwarzanie wstępne obrazów

Filtracja obrazów

Filtracja obrazów jest operacją na pikselach w rezultacie której uzyskujemy przekształcony obraz. Filtracja jest przekształceniem kontekstowym gdyż do jej wykonania, dla każdego piksela obrazu przetworzonego potrzebna jest informacja o wielu pikselach obrazu źródłowego.

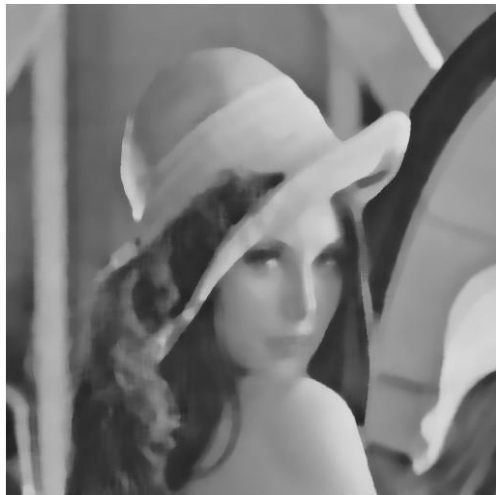
Celem filtracji jest:

- eliminacja (zmniejszenie) szumów w obrazie;
- wzmocnienie pewnych cech obrazu;
- wykrywanie pewnych cech w obrazie (np. krawędzi, kolorów);
- generalna poprawa „jakości” obrazu i usunięcie jego „wad”.

Podstawowe rodzaje filtrów:

- filtry dolnoprzepustowe i górnoprzepustowe;
- maksymalny i minimalny
- medianowy;
- krawędziowy
- gradientowe...

Przetwarzanie wstępne obrazów



Przykład wykorzystania filtru medianowego:

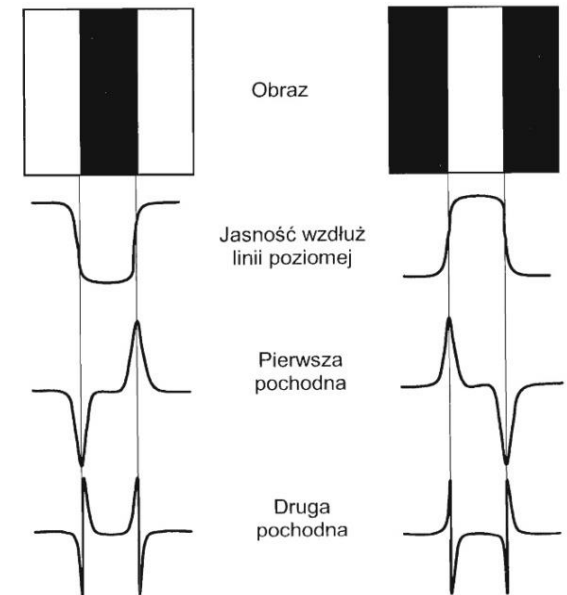
- a) zdjęcie oryginalne,
- b) wynik filtracji medianowej z wykorzystaniem maski 3x3,
- c) wynik filtracji medianowej z wykorzystaniem maski 7x7,
- d) wynik filtracji medianowej z wykorzystaniem maski 13x13

Przetwarzanie wstępne obrazów

Filtr Canney'go

Filtr Cannego służy do wykrywania krawędzi na obrazie. Wykorzystuje do tego pierwszą i drugą pochodną funkcji zmiany jasności pikseli na obrazie. Algorytm działa punkt po punkcie.

Istotnymi elementami filtru jest odsumowanie, filtry np. medianowe (usuwanie pojedynczych pikseli), progowanie (usunięcie nieistotnych krawędzi).

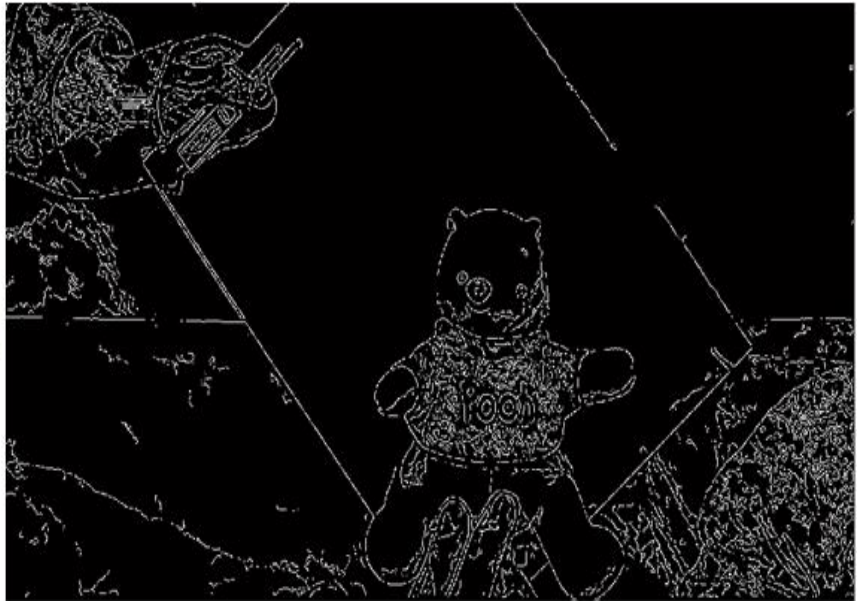


Wykrywanie krawędzi na podstawie pierwszej i drugiej pochodnej

Przetwarzanie wstępne obrazów

Filtr Canney'go

Przykład zastosowanie filtru Canny'ego do detekcji krawędzi



źródło:

Skorupa, M., Lewicki, D., Lewicki, T. Wykorzystanie metod przetwarzania obrazów i inteligencji obliczeniowej do identyfikacji biometrycznej osób na podstawie twarzy. Praca inżynierska. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki, Katedra Inżynierii Systemów Sterowania.

Przetwarzanie wstępne obrazów

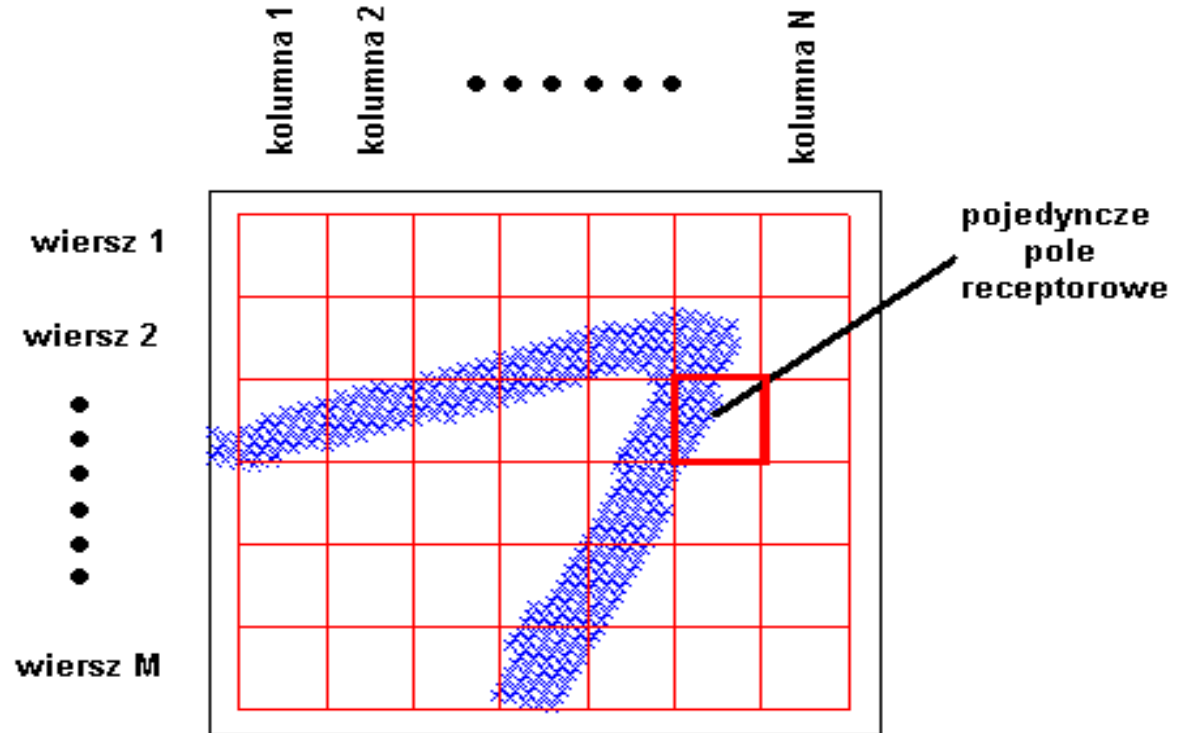
Redukcja wymiaru

- skala szarości, binaryzacja, progowanie, histogramy;
- siatki redukcyjne;
- PCA;
- k – means clustering;
- DCT (ang. Discrete Cosine Transform) – wykorzystywana w kompresji jpg;
- przekształcenie log-polar;
- transformacje Fouriera, Gabora, falkowa...

Przetwarzanie wstępne obrazów

Redukcja wymiaru

Siatka prostokątna



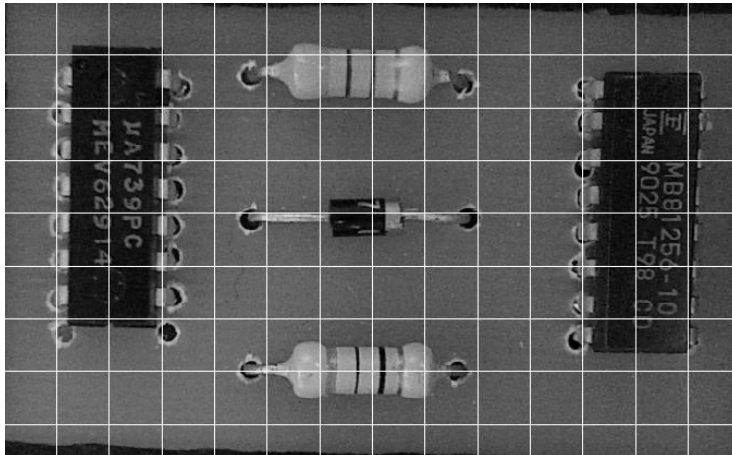
źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

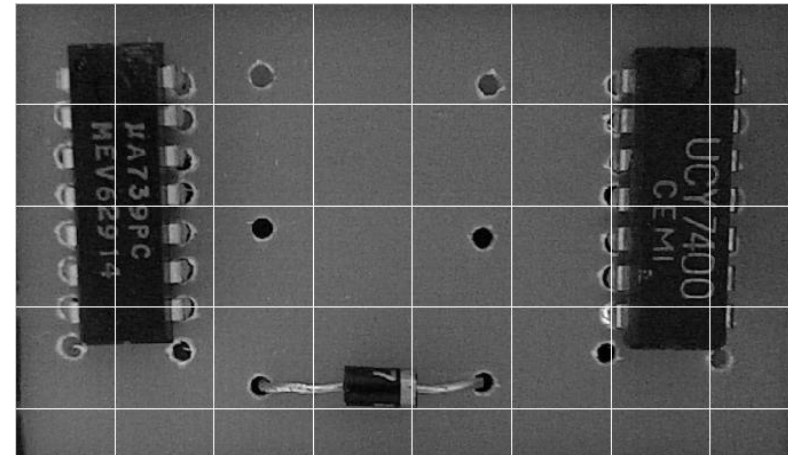
Redukcja wymiaru

Siatka prostokątna – dobór parametrów

Siatka o oczku 50 x 50



Siatka o oczku 90 x 90



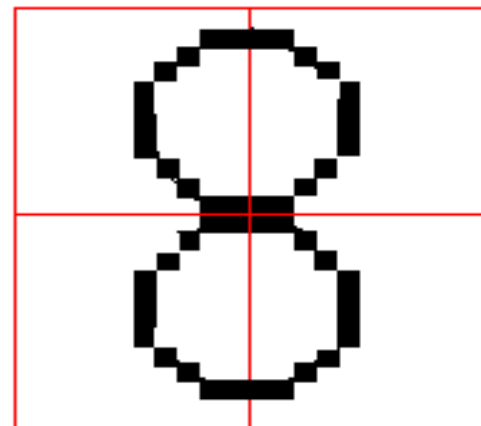
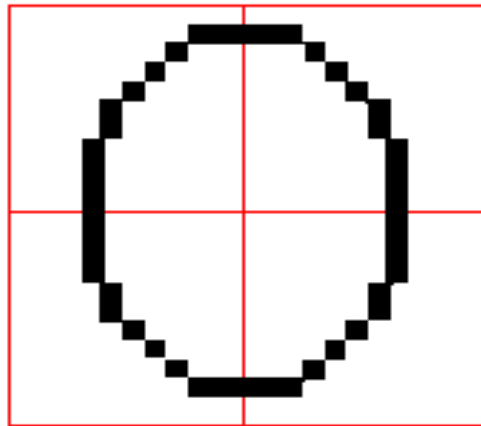
źródło: Mariusz Sikora. Wykorzystanie sztucznych sieci neuronowych do diagnostyki optycznej poprawności produkcji płytek drukowanych. Praca magisterska, PG 2010

Przetwarzanie wstępne obrazów

Redukcja wymiaru

Siatka prostokątna – dobór parametrów

Nałożenie siatki o źle dobranych parametrach $M=2$ i $N=2$ na cyfry 0 i 8.
W obu przypadkach wyjściem dla tej transformacji będą identyczne wartości



źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

Redukcja wymiaru

Siatka prostokątna z adoptowalnym rozmiarem pól receptorowych

$$X_{siatki} = 2 \cdot k \cdot \sigma_x$$

$$Y_{siatki} = 2 \cdot k \cdot \sigma_y$$

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

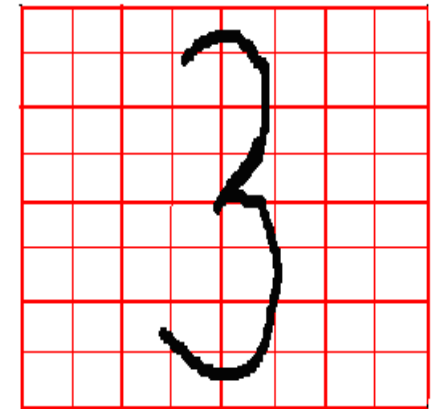
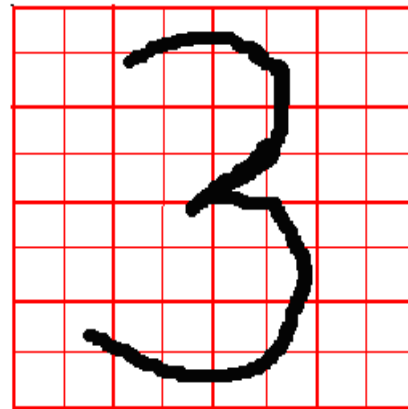
$$\sigma_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$$

gdzie:

σ_x, σ_y - odchylenie standardowe

\bar{x}, \bar{y} - wartości średnie współrzędnych x i y punktów należących do obiektu

x_i, y_i - współrzędne i -tego punktu należącego do obiektu



źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

Redukcja wymiaru

Siatka eliptyczna

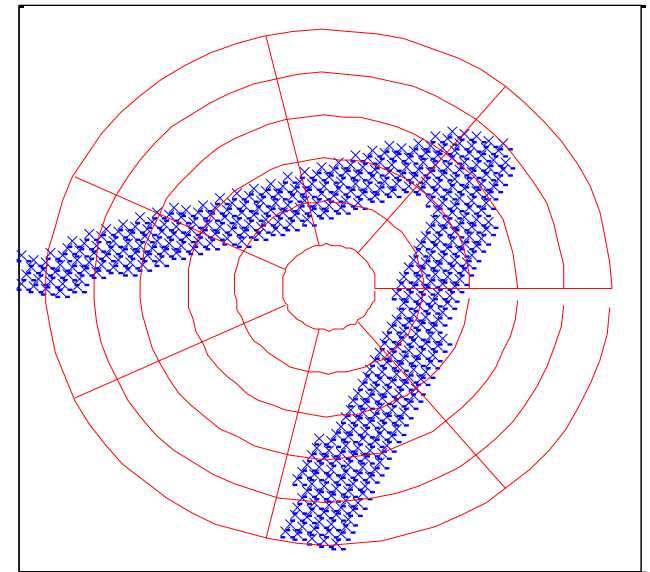
Liczba pól receptorowych:

$$L_{PR} = W \cdot (P - 1) + 1$$

gdzie:

P - liczba pierścieni siatki

W - liczba wycinków siatki
(równomiernie kwantowanych kątów)



źródło: Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, PG 2008

Przetwarzanie wstępne obrazów

Przykład: Rozpoznawanie cyfr

Sekcja oparta o:

Maciej Wenecki. Systemy wizyjne do przetwarzania i rozpoznawania obrazów przy wykorzystaniu sztucznych sieci neuronowych.

Praca magisterska, PG 2008

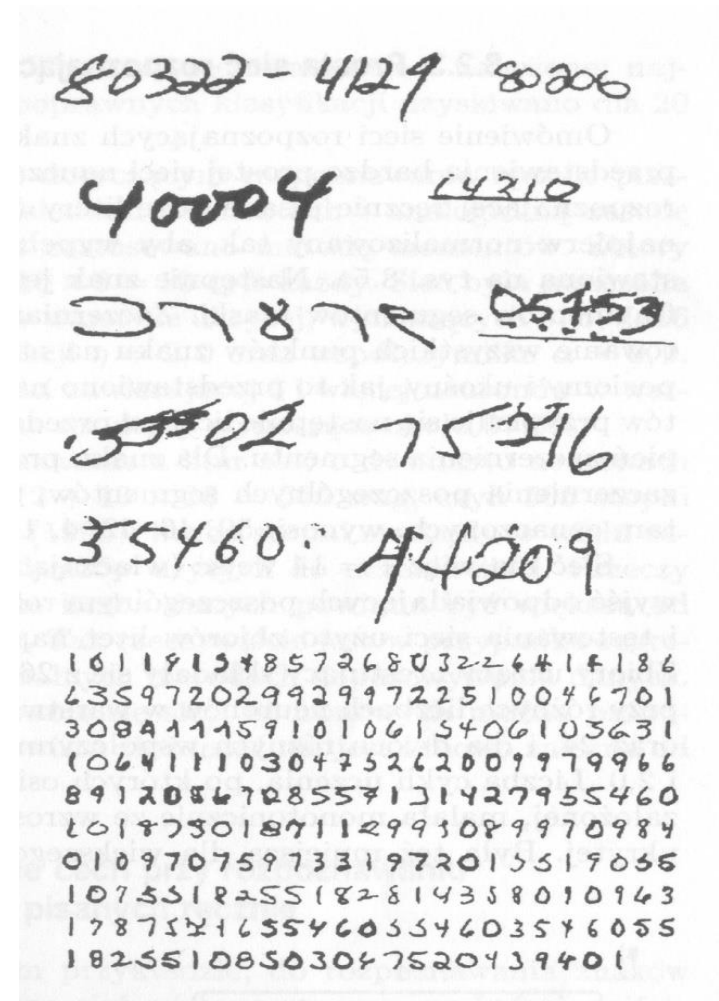
Promotor dr inż. Michał Grochowski

Rozpoznawanie wstępne cyfr

Wprowadzenie

Przykłady kodów miejscowości pisanych na listach oraz znormalizowanych pisanych cyfr

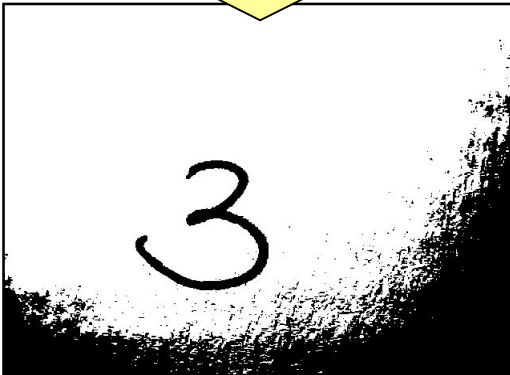
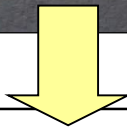
1 2 3 4 5 6 7 8 9 0
1 2 3 4 5 6 7 8 9 0
1 2 3 4 5 6 7 8 9 0
1 2 3 4 5 6 7 8 9 0



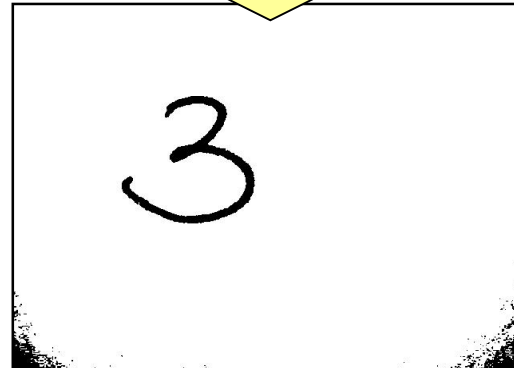
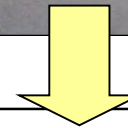
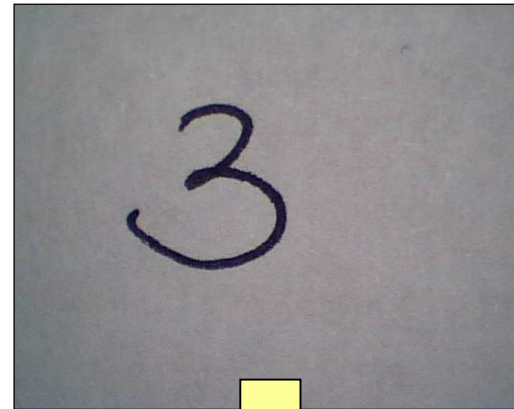
źródło: Żurada J., Barski M., Jędruch W. (1996)
Sztuczne sieci neuronowe – Wydawnictwo Naukowe PWN

Rozpoznawanie cyfr

Wpływ oświetlenia na progowanie



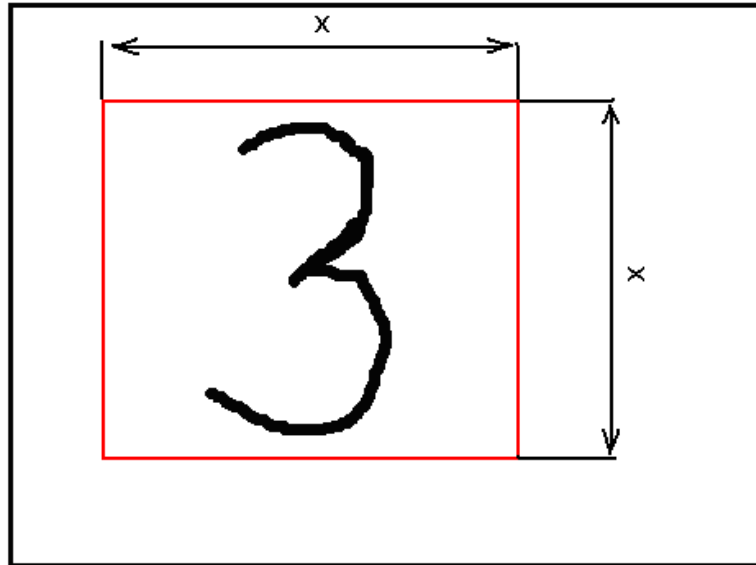
Błędnie dobrane oświetlenie



Poprawnie dobrane oświetlenie

Rozpoznawanie cyfr

Przycinanie do obiektu



Czarna ramka określa pierwotny wymiar obrazu,
czerwona ramka oznacza obszar obrazu po przycięciu

Rozpoznawanie cyfr

Zmiana reprezentacji obrazu

Współrzędne punktów (w układzie kartezyjskim) wyznaczone są z zależności:

$$x_i = k_i$$

$$y_i = W - w_i$$

gdzie:

x_i - współrzędna określająca położenie na osi X;

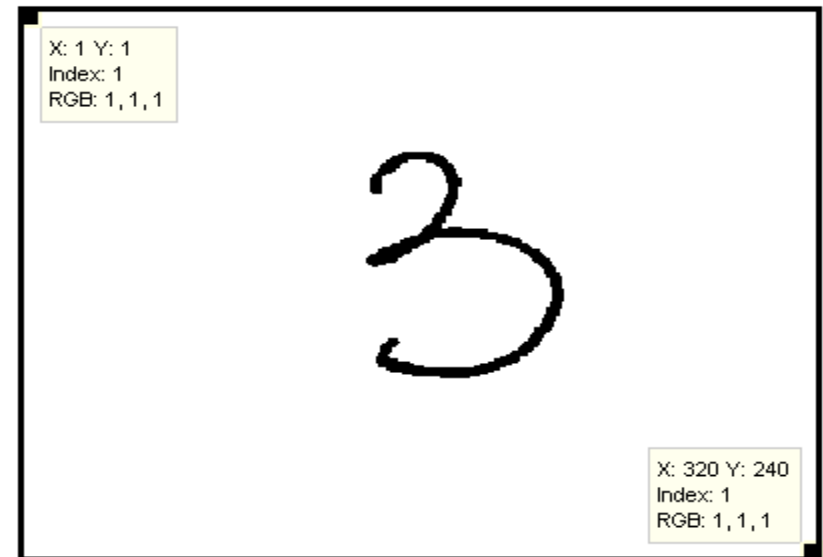
y_i - współrzędna określająca położenie na osi Y

k_i - numer kolumny i-tego punktu należącego do obiektu

w_i - numer wiersza i-tego punktu należącego do obiektu

W - liczba wierszy z których składa się macierz reprezentująca obraz

Reprezentacja obrazu w postaci macierzowej



Rozpoznawanie cyfr

Zmiana reprezentacji obrazu

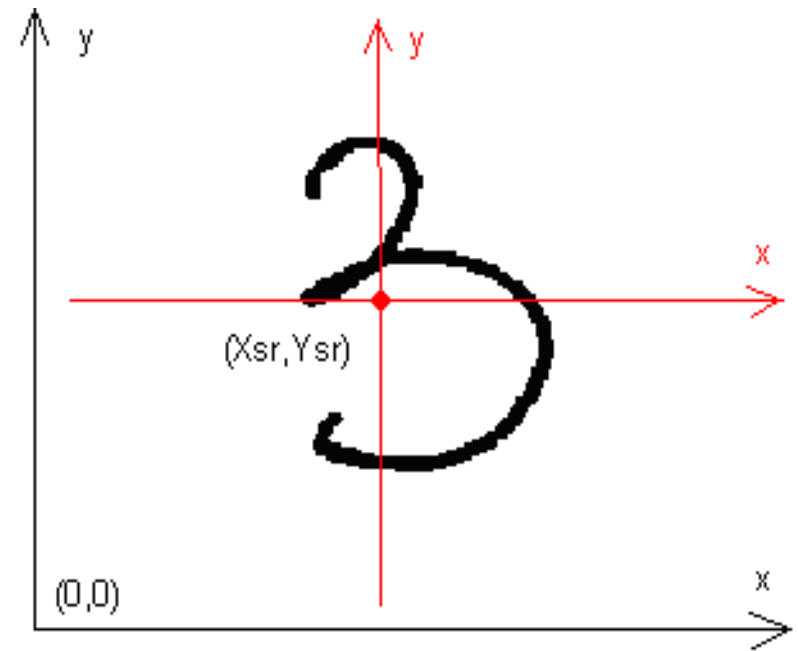
Wyznaczenie środka ciężkości figury:

$$x_{sr} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$y_{sr} = \frac{1}{N} \sum_{i=1}^N y_i$$

Następnie do współrzędnych każdego punktu należy dodać wektor:

$$v = [-x_{sr}, -y_{sr}]$$



Przeniesienie początku układu współrzędnych do środka ciężkości obiektu

Rozpoznawanie cyfr

Zmiana reprezentacji obrazu

Wyznaczanie głównej osi obiektu

Każda cyfra (i inny znak) posiada własną charakterystyczną oś, od której średnie odchylenie standardowe dla wszystkich punktów będzie najmniejsze. Dokonując rotacji obrazu, wokół środka ciężkości danej cyfry, tak aby ta charakterystyczna oś pokrywała się, np. z osią OY w układzie kartezjańskim, otrzymamy obraz znormalizowany.

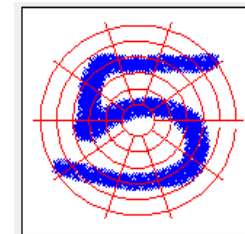
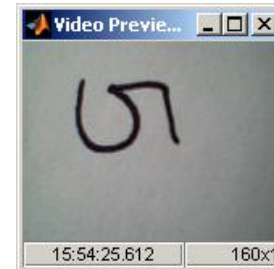
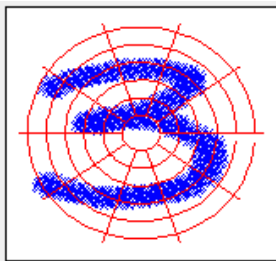
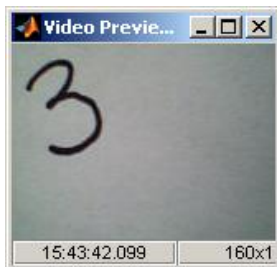
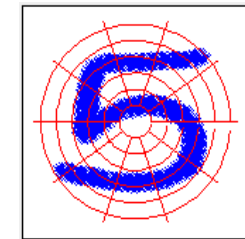
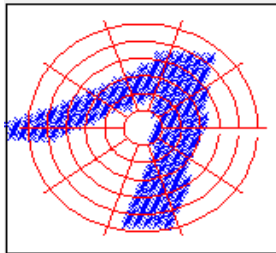
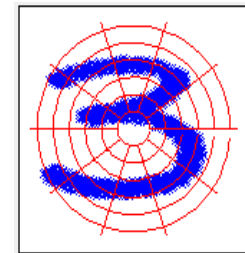
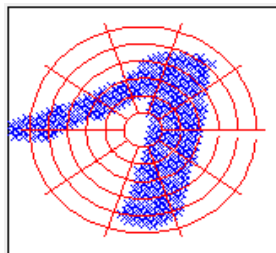
Wyznaczenie tej osi jest możliwe przy wykorzystaniu metody PCA.

Za dane treningowe należy przyjąć wszystkie punkty reprezentujące daną cyfrę, zapisane w układzie kartezjańskim, a otrzymany wektor własny, któremu odpowiada największa liczbowo wartość własna, będzie określał kierunek osi.

Rozpoznawanie cyfr

Zmiana reprezentacji obrazu

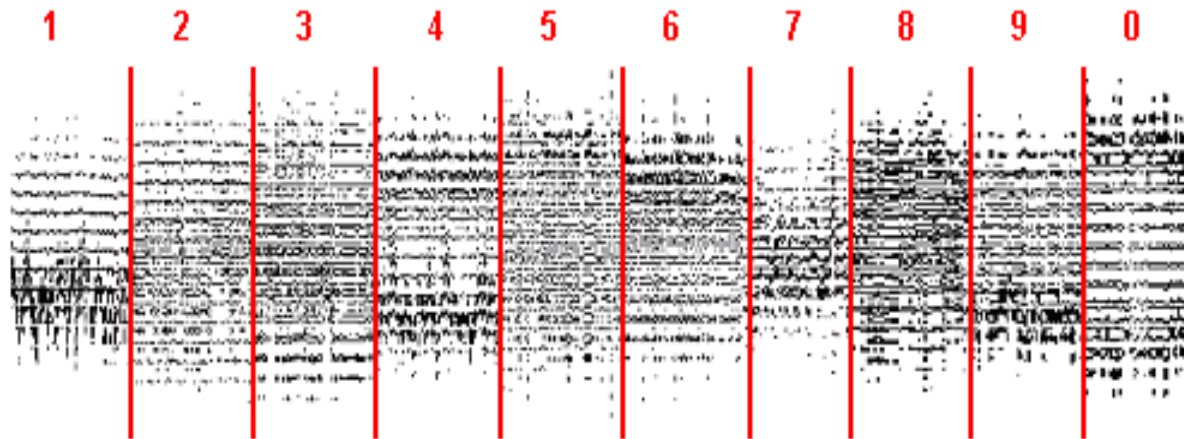
„Wyrównywanie” obiektu



Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek

Baza wzorców



Baza wzorców składa się z 1476 cyfr, z czego na każdą cyfrę przypada w przybliżeniu podobna liczba wzorców. Każda cyfra przedstawiona jest na jednakowej macierzy (macierz o wymiarze 20 x 20 pikseli), przy czym nie występują istotne różnice w rozmiarze tych cyfr oraz każda z nich pisana jest przez jedną osobę, jednym charakterem pisma i pod tym samym kątem.

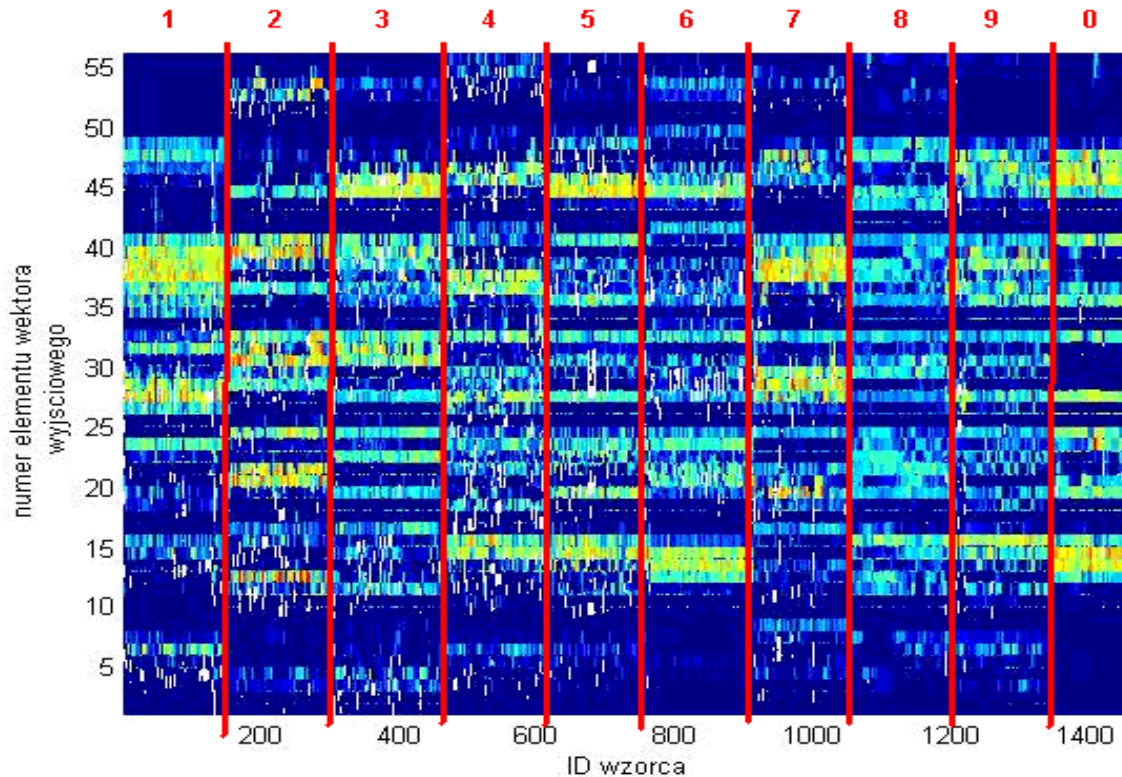
Ze względu na fakt, iż wejściem sieci neuronowej jest wektor, każdą cyfrę przedstawiono w postaci wektora składającego się z 400 elementów

Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek

Nakładamy siatkę prostokątną o ilości wierszy $M = 8$ i ilości kolumn $N = 7$

Uzyskujemy 56 jednakowy pól receptorowych (co stanowi ilość wejść do sieci neuronowej)



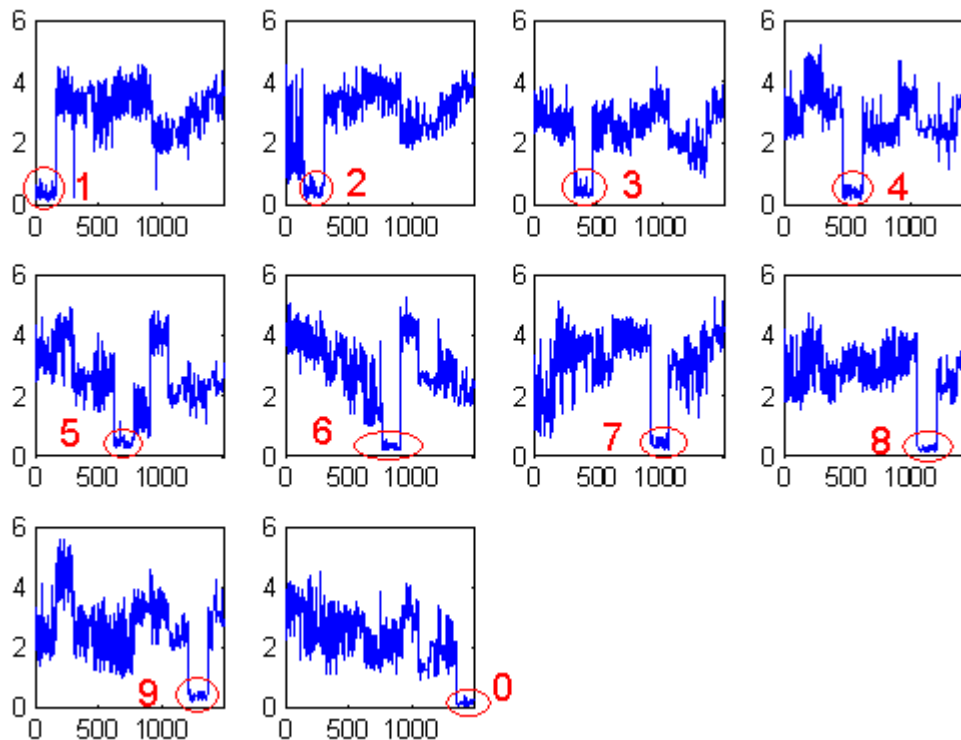
Podobieństwa
wykazują pary cyfr:

- 5 i 6;
- 8 i 9;
- 1 i 7.

Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

Tak przetworzone dane poddajemy obróbce PCA: redukcja wymiarów z 56 do 20 składników głównych (co stanowi ilość wejść do sieci neuronowej) oraz analiza miary SPE.



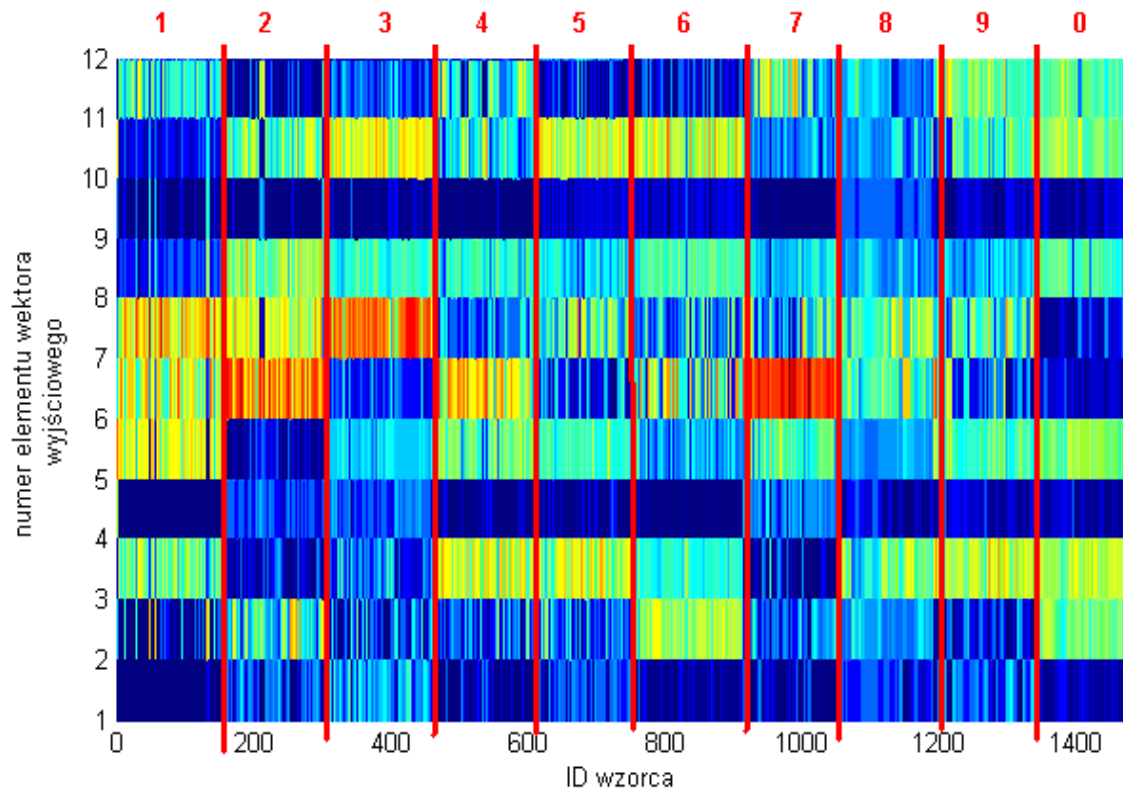
Porównanie wskaźnika SPE dla transformacji PCA wszystkich wzorców cyfr dla każdego modelu osobno

Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek

Nakładamy siatkę prostokątną o ilości wierszy $M = 4$ i ilości kolumn $N = 3$

Uzyskujemy 12 jednakowy pól receptorowych (co stanowi ilość wejść do sieci neuronowej)

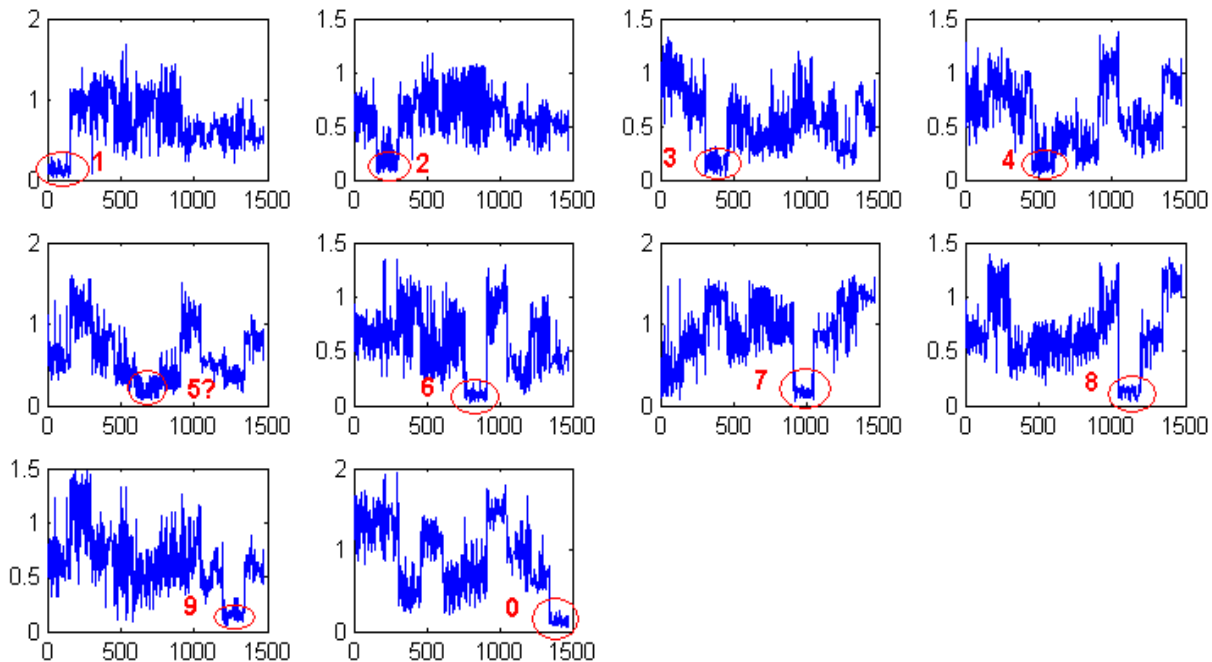


Brak wyraźne
zarysowanych różnic
pomiędzy znakami

Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

Redukcja wymiarów z 12 do 5 składników głównych (co stanowi ilość wejść do sieci neuronowej) oraz analiza miary SPE.



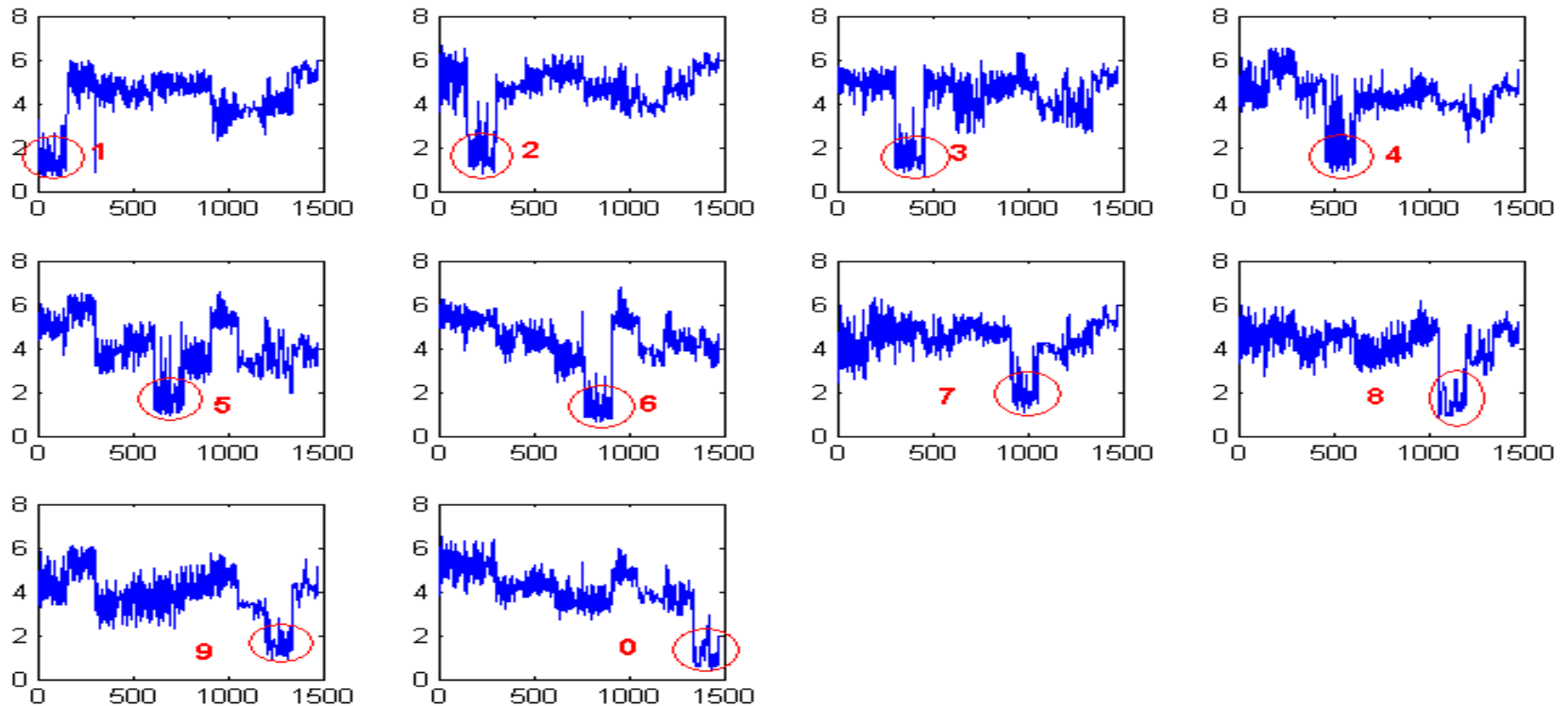
Porównanie wskaźnika SPE
dla transformacji PCA
wszystkich wzorców cyfr dla
każdego modelu osobno

Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

Porównanie miar SPE dla różnych ilości składników głównych R

R = 3

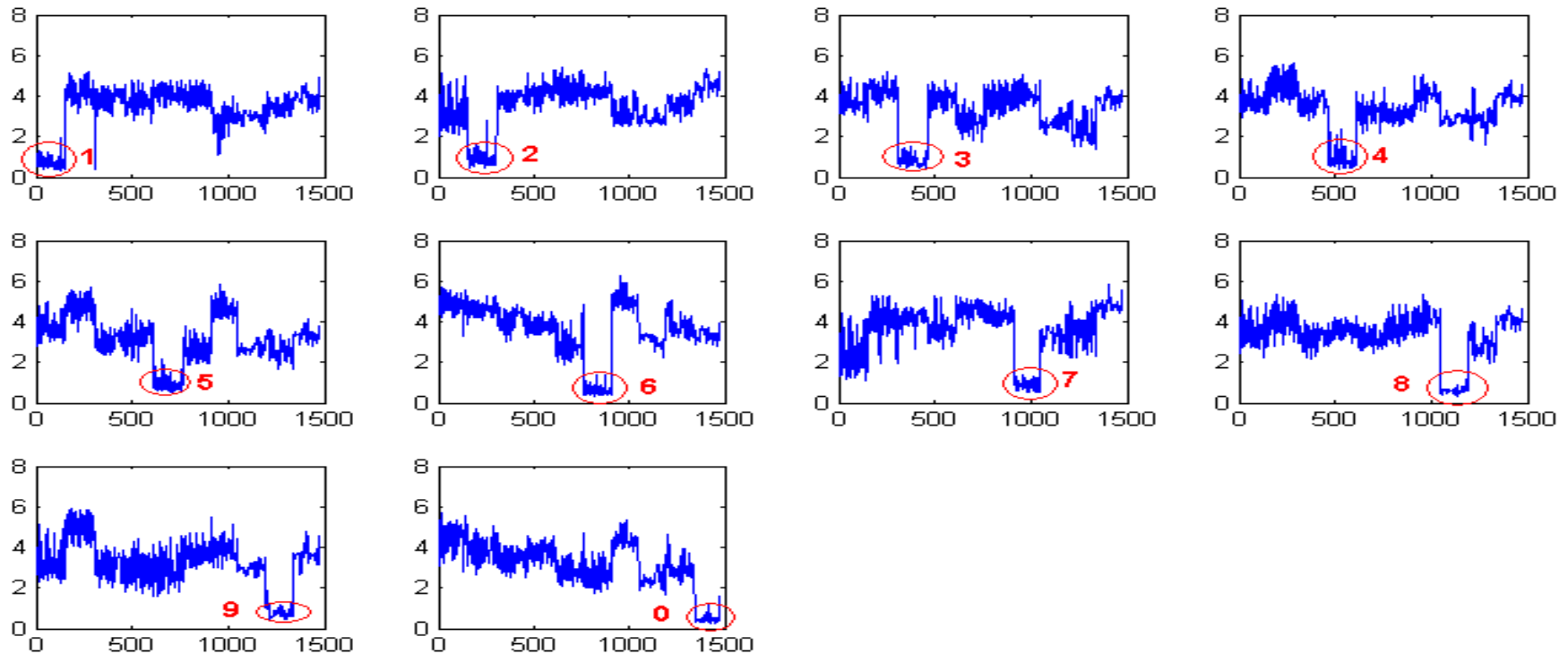


Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

Porównanie miar SPE dla różnych ilości składników głównych R

R = 10

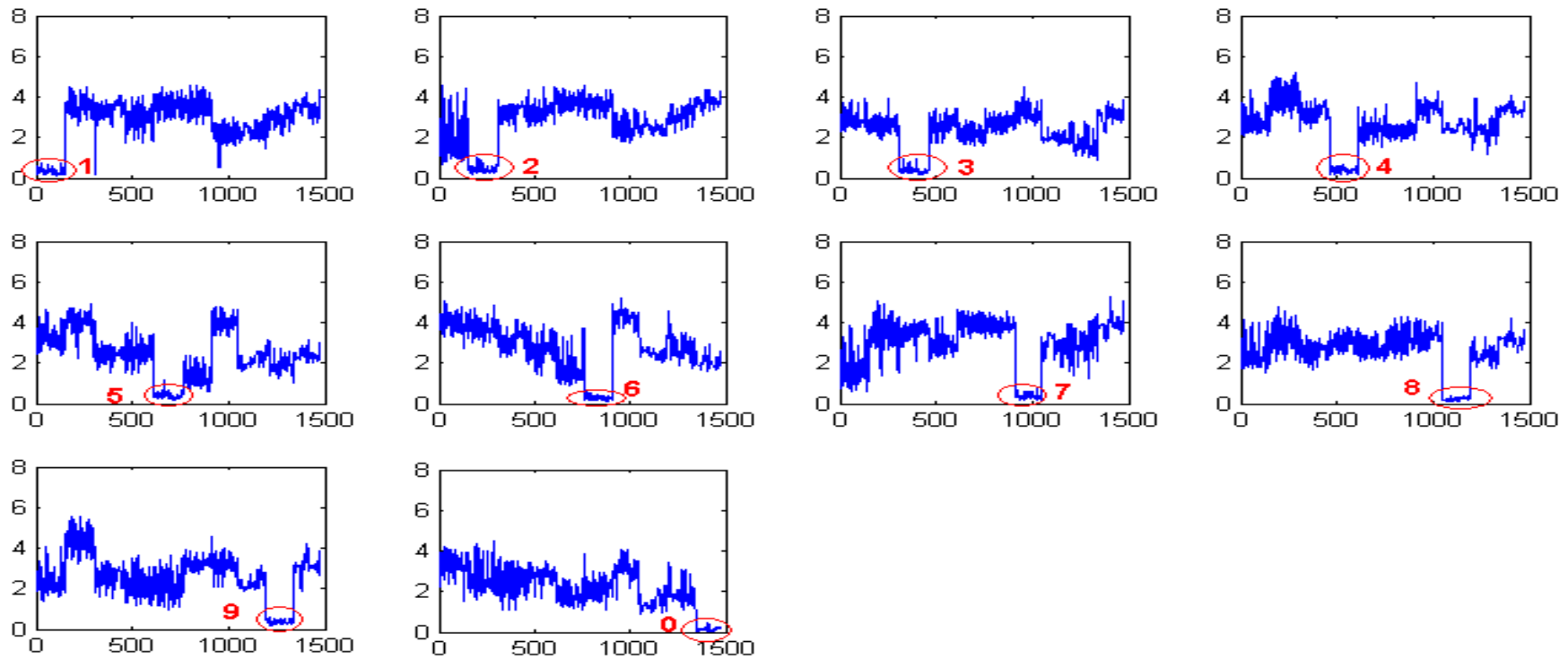


Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

Porównanie miar SPE dla różnych ilości składników głównych R

R = 20



Rozpoznawanie cyfr

Redukcja wymiaru – nakładanie siatek + analiza PCA

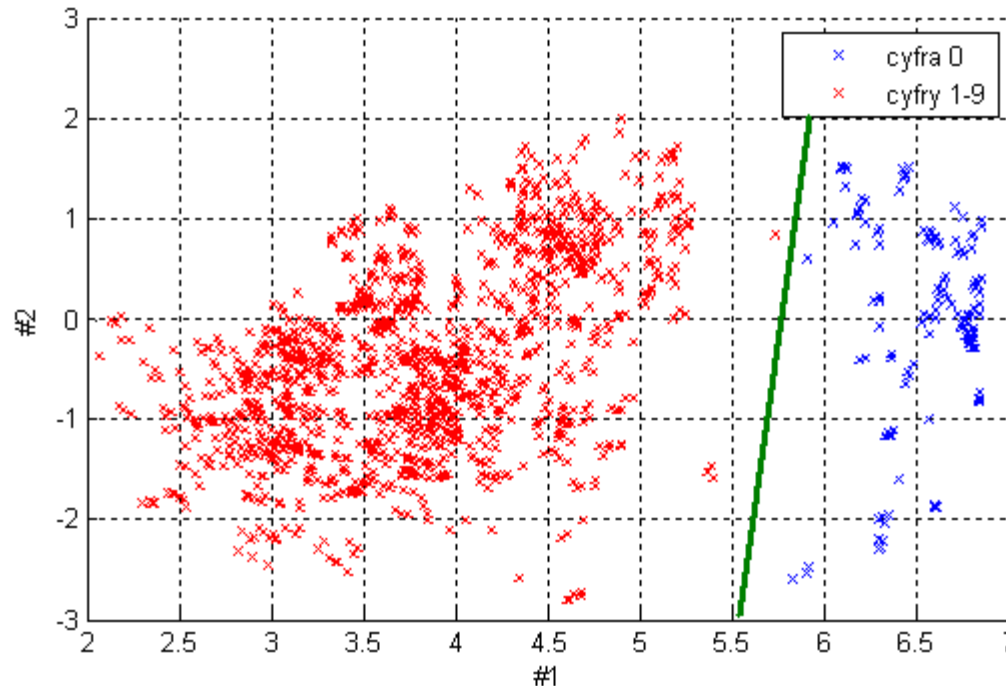
Porównanie miar SPE dla różnych ilości składników głównych R

		Siatka prostokątna		Siatka eliptyczna	
Liczba pól receptorowych		56	12	51	13
Liczba wymiarów wyjściowych po transformacji PCA		20	5	20	5
Rozpoznawalność [%]	Obrazy wzorcowe	99.86	95.28	99.80	95.26
	Obrazy testowe	95.16	85,82	96.27	79.85

Rozpoznawanie cyfr

Analiza doboru struktury sieci neuronowej

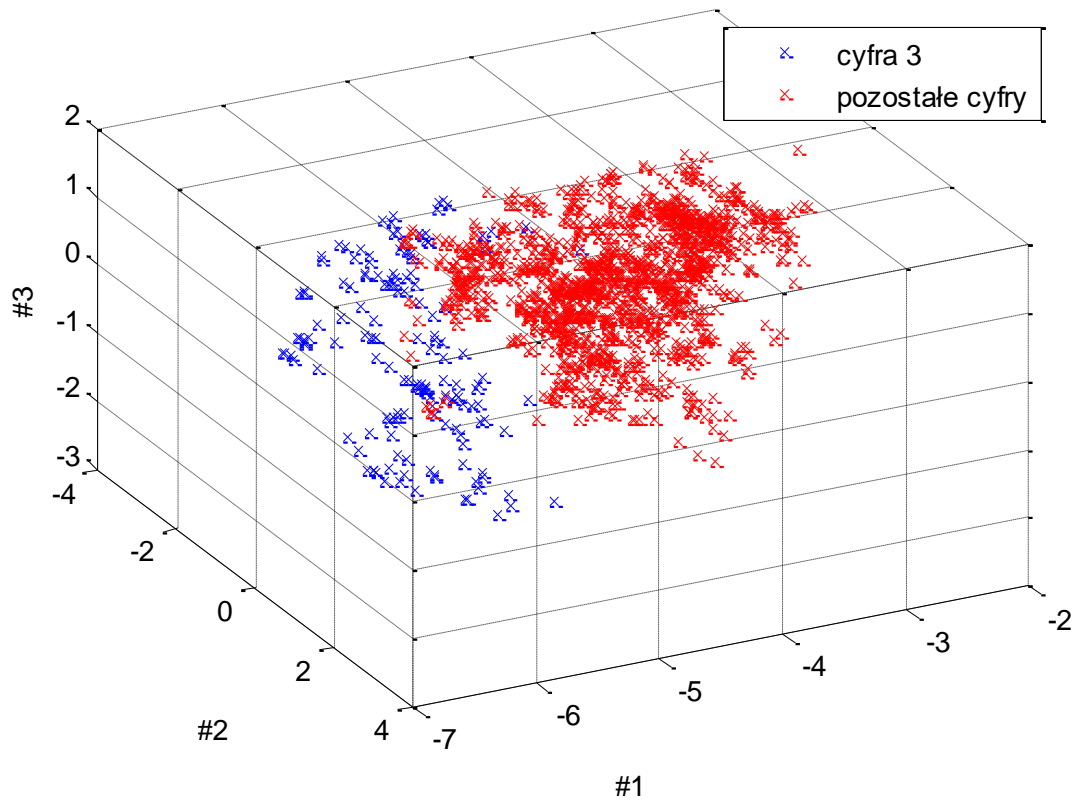
Graficzna prezentacja wzorców cyfr po transformacji PCA (dwa pierwsze istotne PCs).
Prosta oznaczona kolorem zielony wyznacza granicę między tymi klasami



Rozpoznawanie cyfr

Analiza doboru struktury sieci neuronowej

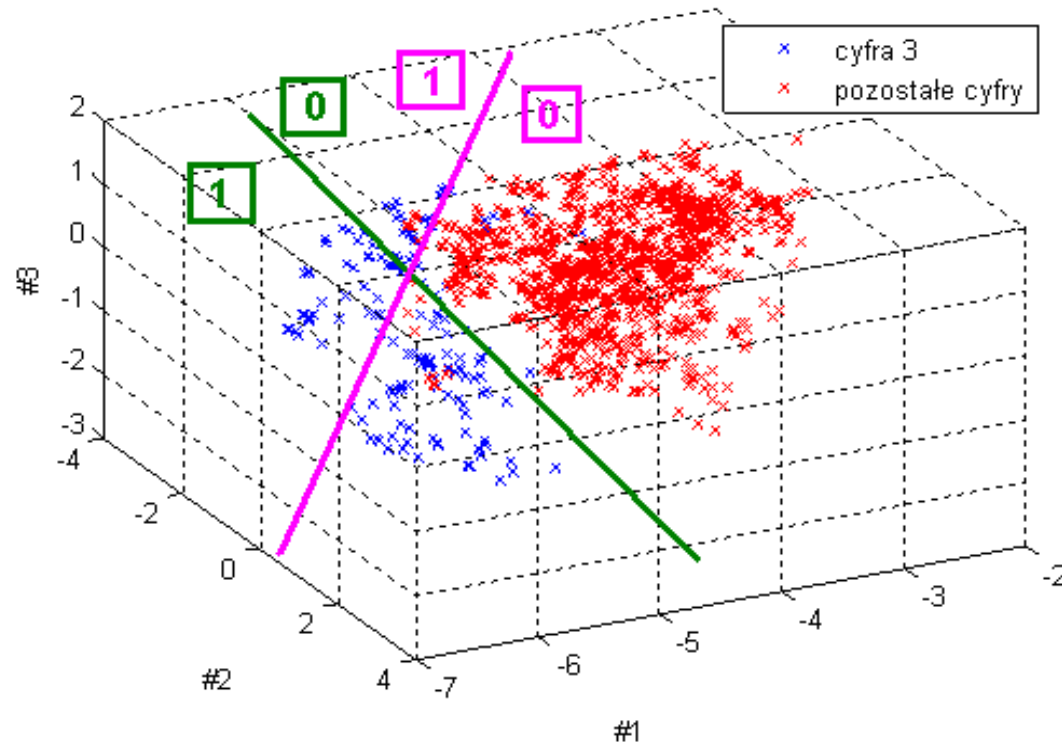
Graficzna prezentacja wzorców cyfr po transformacji PCA (trzy pierwsze istotne PCs).



Rozpoznawanie cyfr

Analiza doboru struktury sieci neuronowej

Graficzna prezentacja wzorców cyfr po transformacji PCA (trzy pierwsze istotne PCs).
Prosta oznaczona kolorem zielony wyznacza granicę między tymi klasami



Przykładowa struktura sieci:

- pierwsza warstwa:
2 neurony najlepiej tangens sigmoidalny;
- druga warstwa:
jeden neuron, może być liniowy.

Przetwarzanie wstępne obrazów

Przykład: Analiza znamion skórnych

Sekcja oparta o:

Agnieszka Mikołajczak: Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

Praca inżynierska, PG 2015

Promotor dr inż. Michał Grochowski

Julia Sas-Swistelnicza

Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

Praca inżynierska, PG 2016

Promotor dr inż. Michał Grochowski



**POLITECHNIKA
GDAŃSKA**

Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

Autor: inż. Agnieszka Mikołajczyk

Promotor: Dr inż. Michał Grochowski

Jak rozpoznać czy znamię jest nowotworem?



Metody oceny dermatoskopowej

- **ABCD(E);**
- **Menzies method;**
- **7-point rule;**
- 3-point rule;
- CASH acronym;
- CHAOS and clues;
- BLINCK algorithm.

Metody te są bardzo skuteczne ale stosować je oraz odpowiednio interpretować ich wyniki potrafią jedynie lekarze specjaliści.

Metoda ABCD(E)

- **A (Assymetry)** – Czy znamię jest symetryczne?
- **B (Border)** – Czy krawędź znamienia jest równa i gładka?
- **C (Color)** – Czy znamię jest jednolitego koloru?
- **D (Diffrental Structures)** – Czy można na nim zauważyć kropki, plamy lub inne nietypowe struktury?
- **E (Evolution)**- Czy znamię zmieniło się w ciągu ostatnich 3 miesięcy?

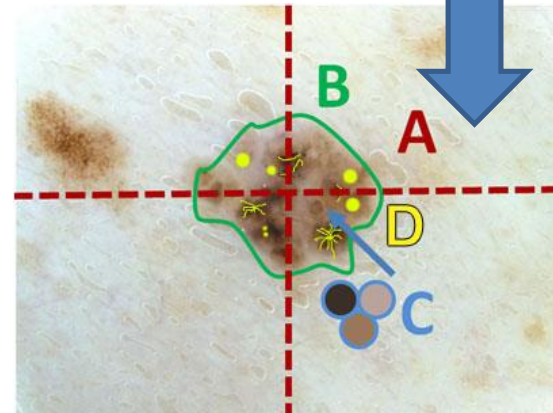
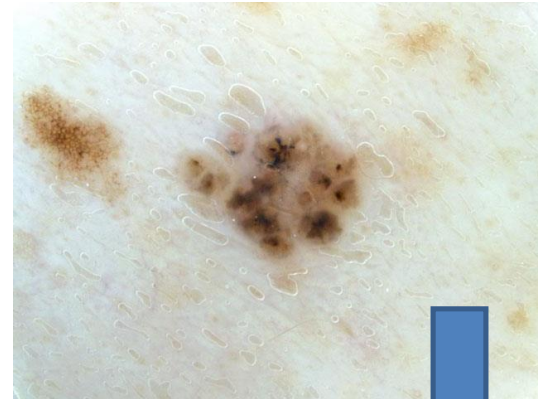
NIE!

NIE!

NIE!

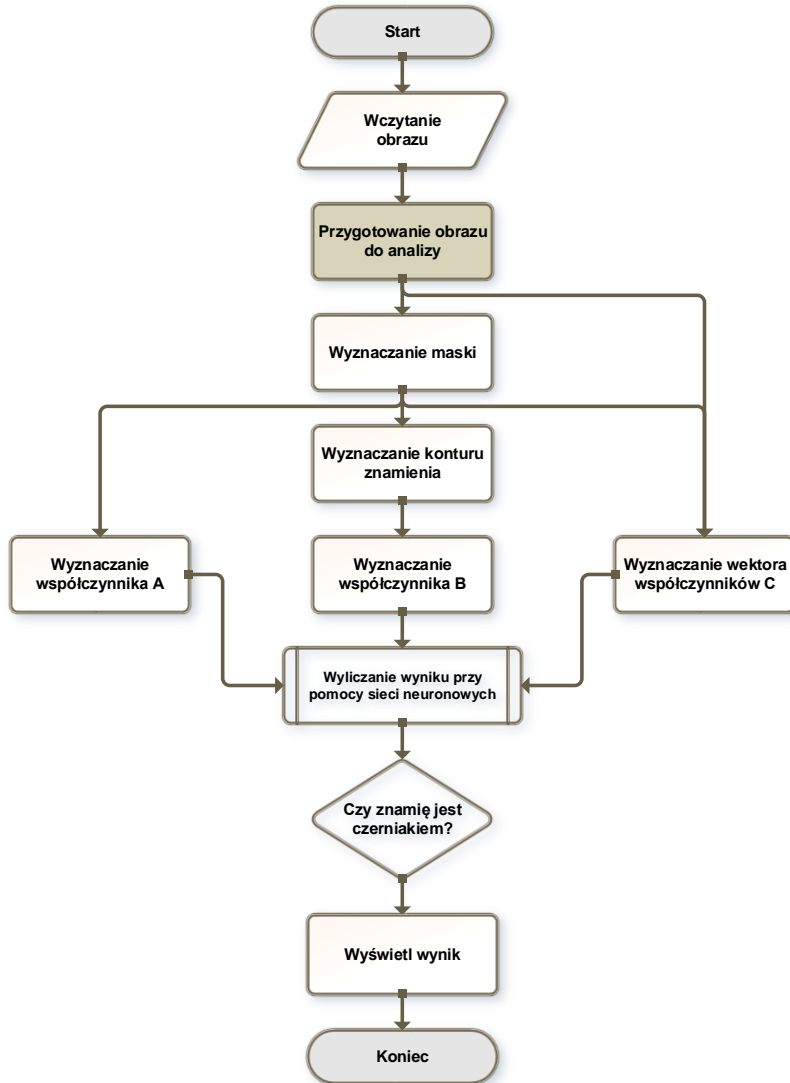
TAK!

TAK!

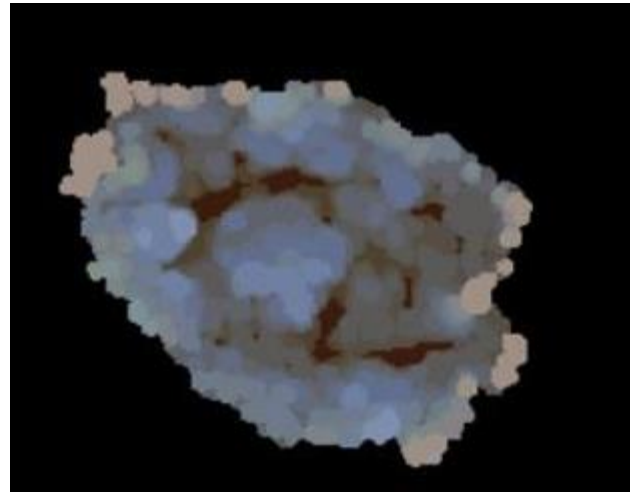


WYNIK: NOWOTWÓR!

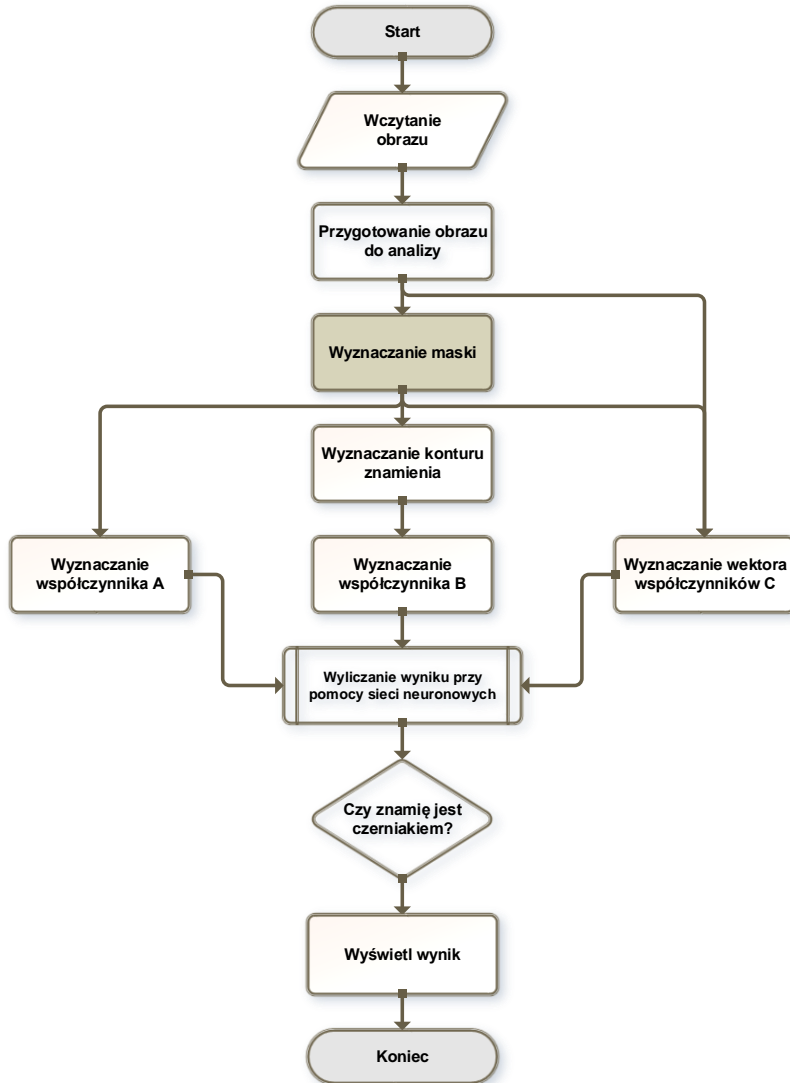
Metody przetwarzania obrazu



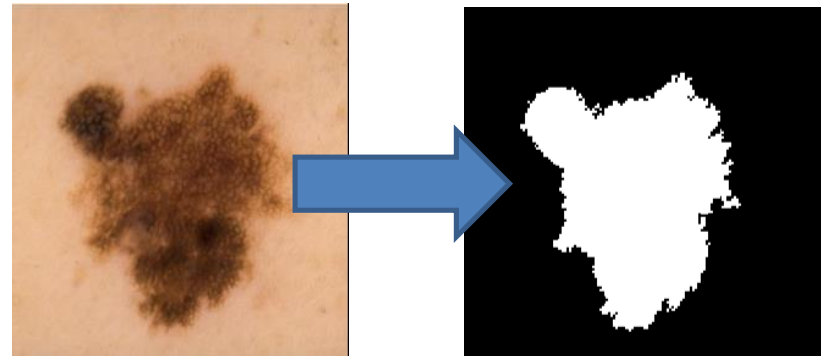
- Rozmycie, ograniczenie ilości kolorów, zwiększenie kontrastu;
- Konwersja do czerni i bieli;
- Zmiany morfologiczne;



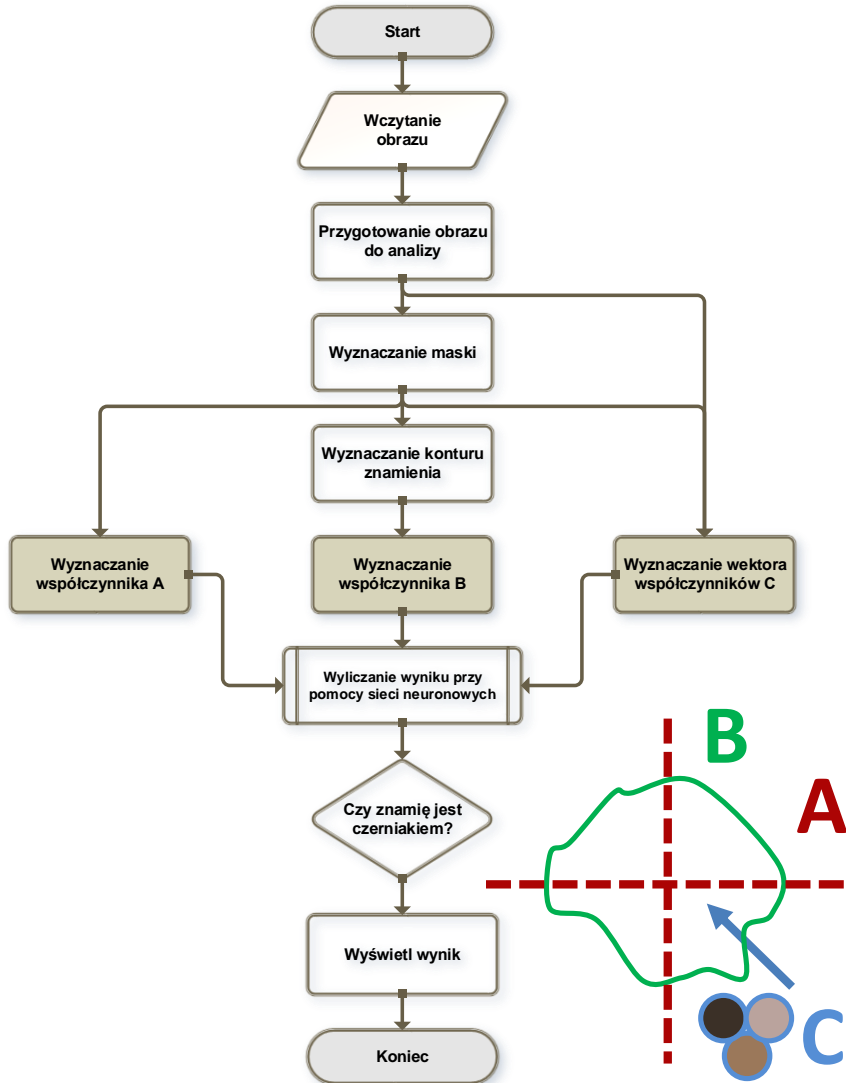
Segmentacja obrazu



- Nowa, innowacyjna metoda bazująca na wyznaczeniu maksimumów lokalnych obrazu;
- Praca na obrazie w skali szarości;
- Wysoka skuteczność działania.



Wyliczanie współczynników decydujących



- Współczynnik **A**- Symetria;

- Współczynnik **B**- krawędź znamienia;

- Współczynnik **C₁**

$$C_1 = \frac{R_{max} - R_{min} + G_{max} - G_{min} + B_{max} - B_{min}}{3}$$

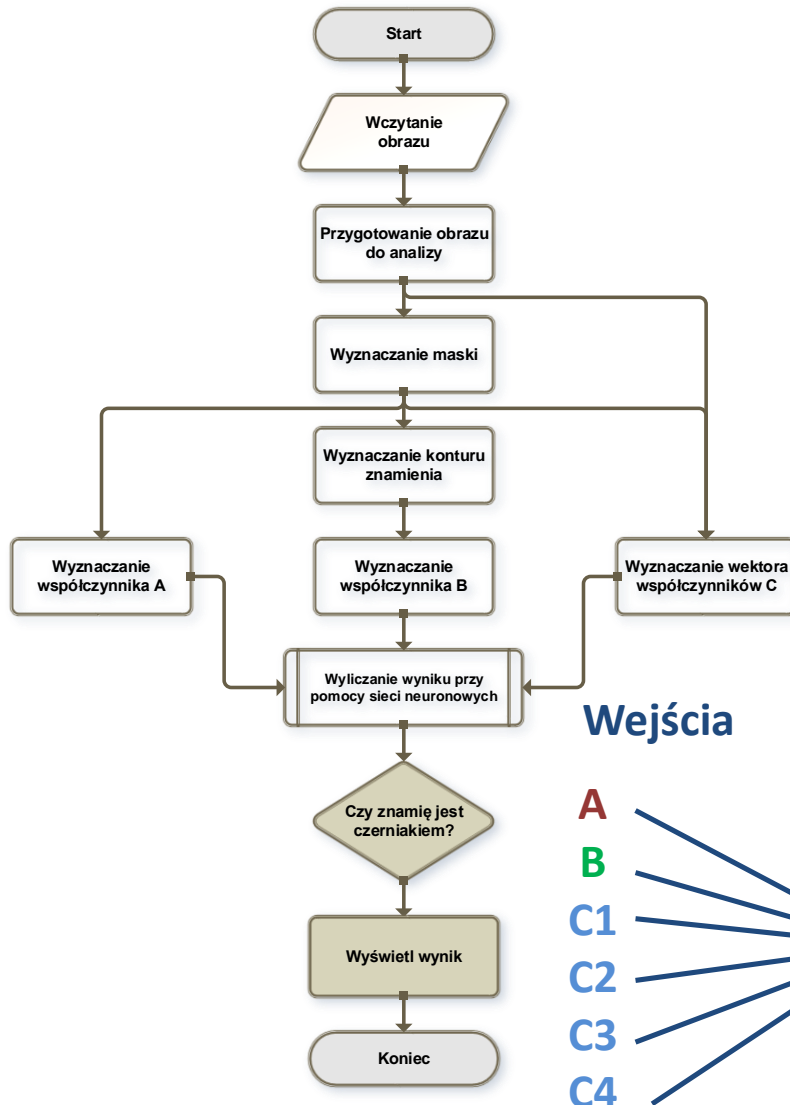
- Współczynnik **C₂** – W jakiej ilości pikseli w znamieniu dominuje kolor niebieski?;

- Współczynnik **C₃**- W jakiej ilości pikseli w znamieniu dominuje kolor zielony?;

- Współczynnik **C₄**

$$C_4 = \frac{C_{max} - C_{min}}{2}$$

Wnioskowanie



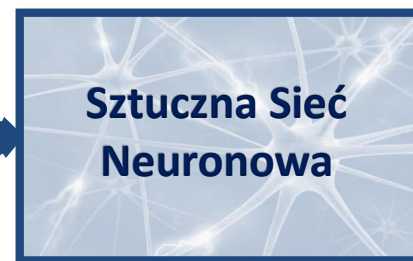
- Wnioskowanie przy pomocy sieci neuronowych;
- Wykorzystano jednokierunkową sieć trójwarstwową, z sigmoidalnymi funkcjami aktywacji;

Wejścia

Proces uczenia

Wyjścia

A
B
C1
C2
C3
C4



- Zdrowe znamię
- Znamię nowotworowe
- Znamię podejrzane

Wynik działania

- **Czułość równa 98%** (100 przypadków);
- **Swoistość równa 73.07%** (26 przypadków);
- Szybkie obliczanie wyniku.



POLITECHNIKA
GDAŃSKA

Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

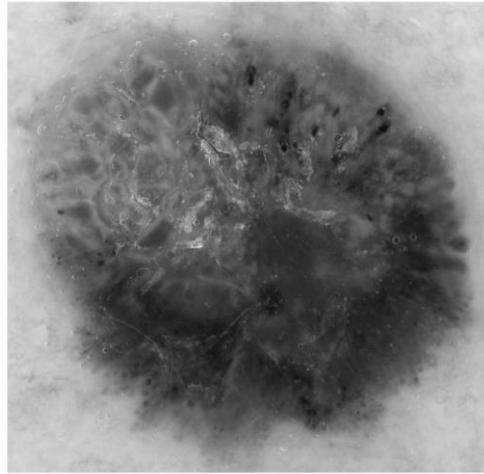
Autor: Julia Sas-Swistelnicza

Promotor: Dr inż. Michał Grochowski

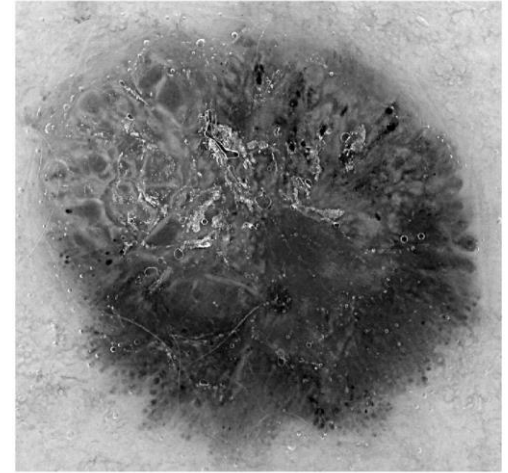
Cropped



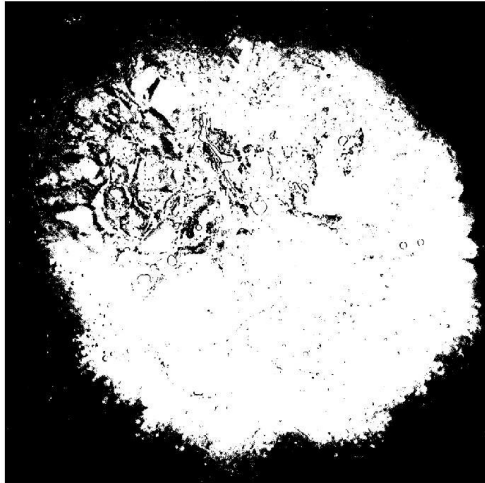
Obraz BW



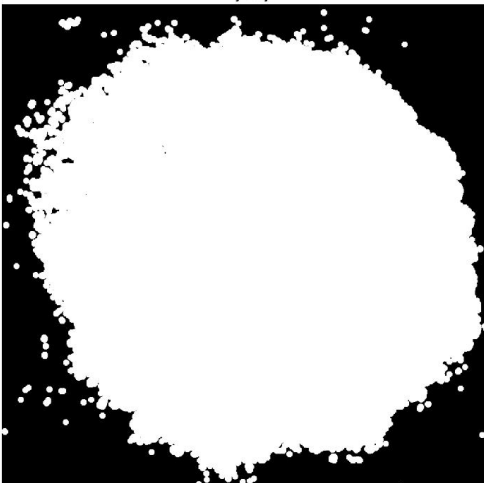
Zwiększony kontrast



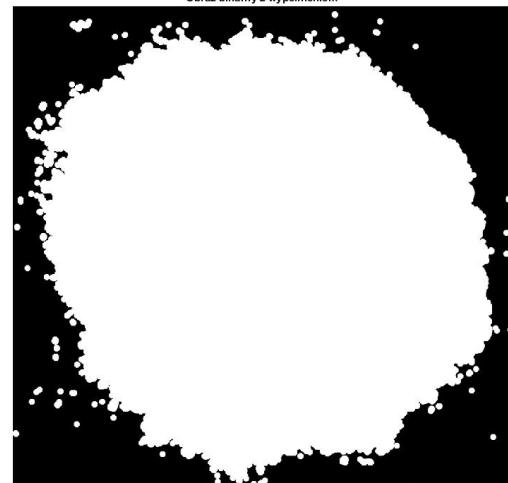
Progowanie metodą Otsu



Dylatacja

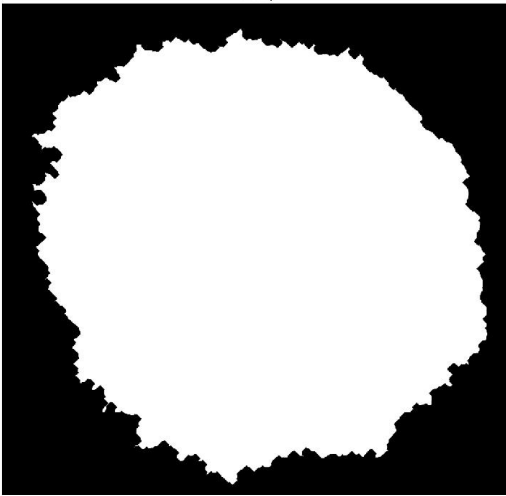


Obraz binarny z wypełnieniem

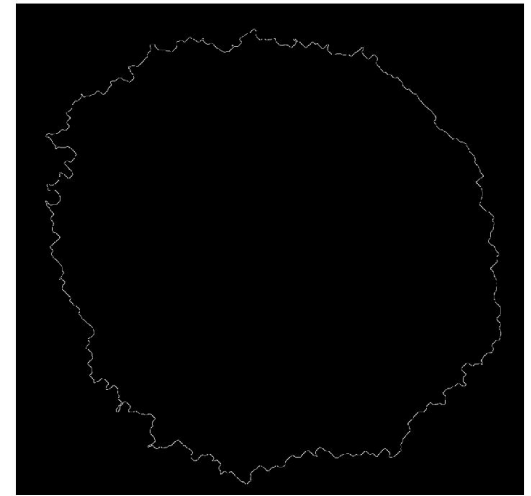
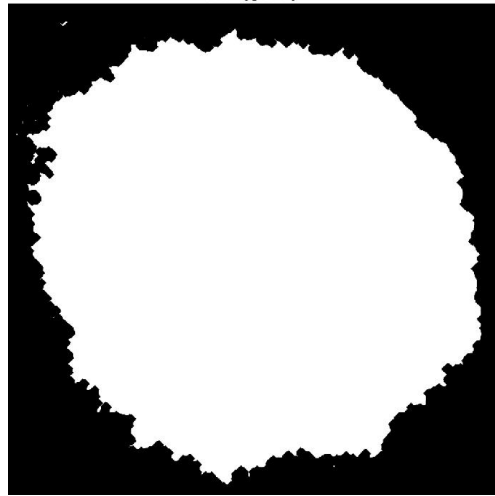


Julia Sas-Swistelnicka: Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

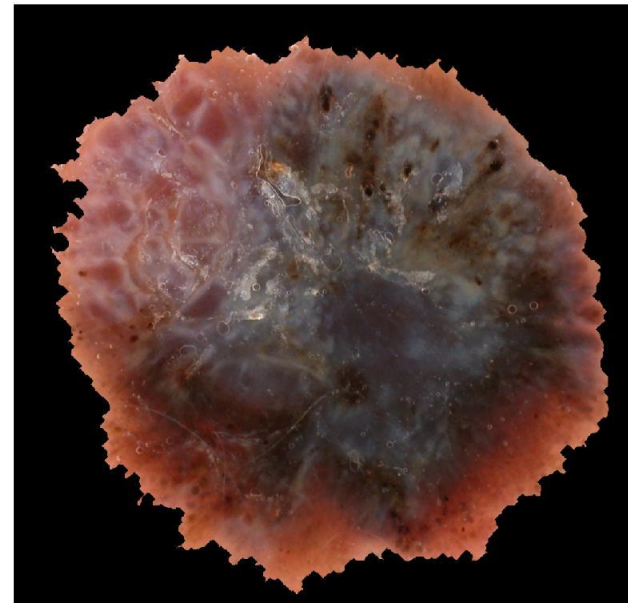
BWareaopen



Obraz wygładzony

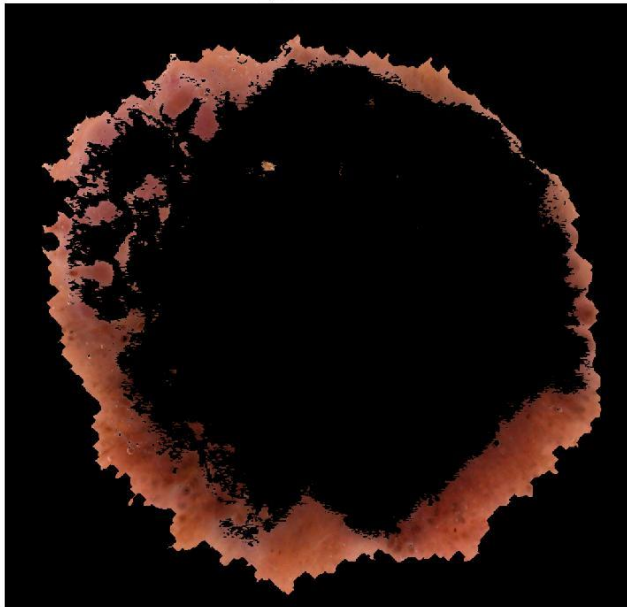


Kontur



Julia Sas-Swistelnicka: Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

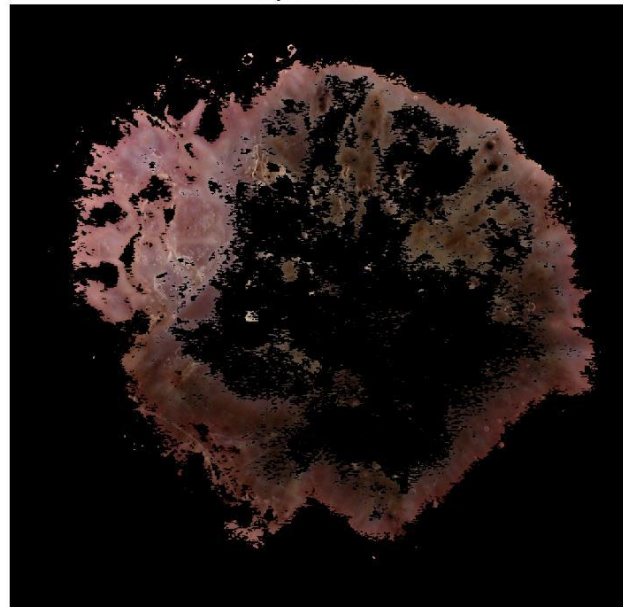
objects in cluster 1



objects in cluster 2



objects in cluster 3

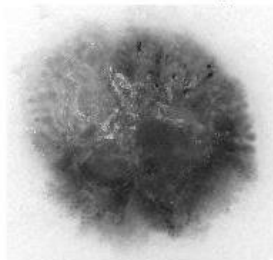


Julia Sas-Swistelnicza: Analiza znamion skórnych przy pomocy metod przetwarzania obrazu i algorytmów inteligencji obliczeniowej

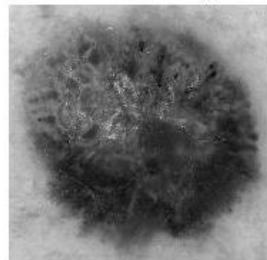
Obraz znamienia RGB



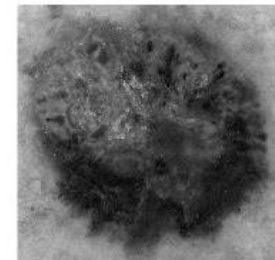
Kanał czerwony R



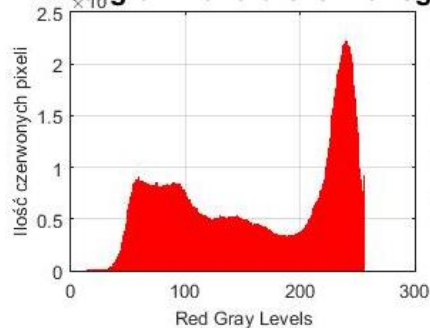
Kanał zielony G



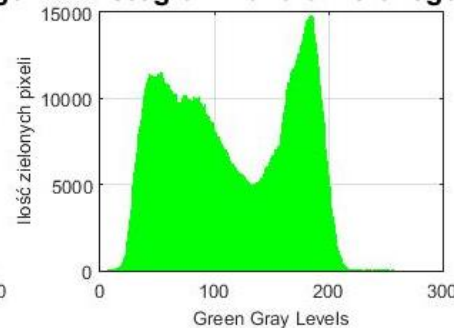
Kanał niebieski B



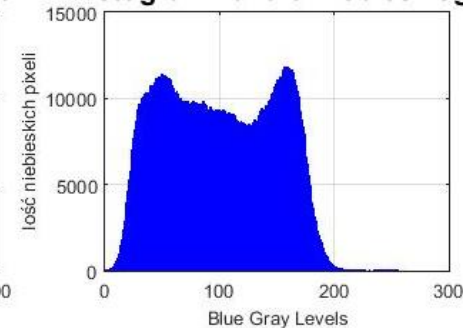
Histogram kanału czerwonego



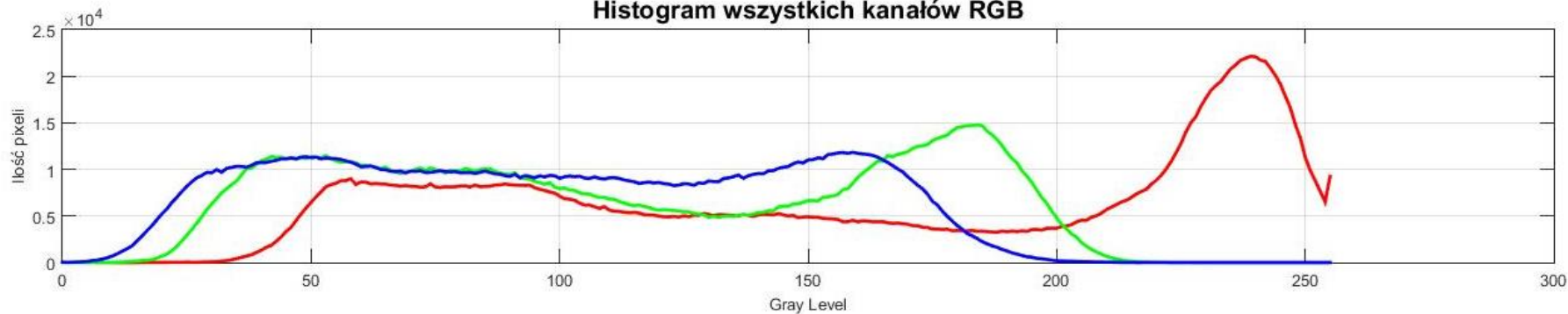
Histogram kanału zielonego



Histogram kanału niebieskiego



Histogram wszystkich kanałów RGB



Rozpoznawanie twarzy

Detekcja ruchu

Różnica kolejnych klitek



źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Detekcja ruchu

Różnica kolejnych klatek



źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Detekcja ruchu

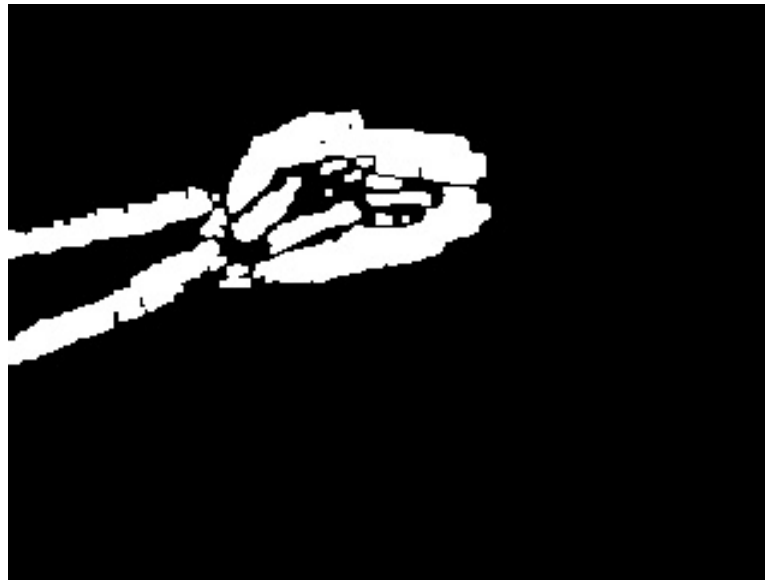
Różnica kolejnych klamek



źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Detekcja ruchu

Różnica kolejnych klatek



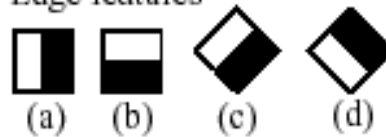
źródło: Przemysław Błaszowski. Rozpoznawanie obiektów dynamicznych przy wykorzystaniu sztucznych sieci neuronowych. Praca inżynierska, PG 2012

Detekcja obiektów – twarzy, elementów twarzy

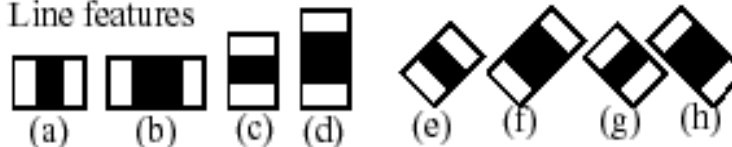
Klasyfikator Haar'a

Istotą działania klasyfikatora Haar'a (algorytm Viola-Jones) są tzw. cechy Haar'a. Cechy te zawierają informacje o zmianie wartości kontrastu pomiędzy prostokątnymi grupami pikseli. Zmiany te są wykorzystywane do określenia relatywnie jasnych i ciemnych obszarów. Dwie lub trzy sąsiadujące grupy o podobnej wariancji kontrastu tworzą cechę Haar'a. Cechy te mogą być łatwo skalowane, tak aby rozpoznawać obiekty o różnej wielkości.

1. Edge features



2. Line features



3. Center-surround features



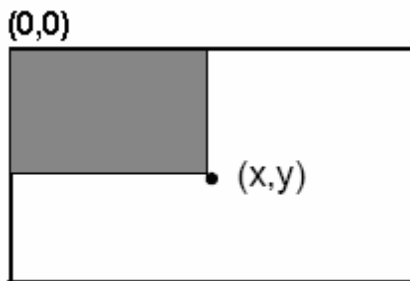
źródło: Wilson P.; Fernandez J., , "Facial feature detection using Haar classifiers" J. Comput. Sci. Coll. 21, p. 127-133, 2006.

Detekcja obiektów – twarzy, elementów twarzy

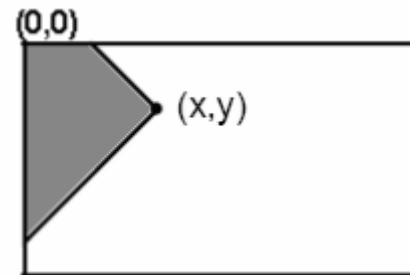
Klasyfikator Haar'a

Prostokątne cechy są obliczane za pomocą zmienionej reprezentacji zdjęcia, zwanej zdjęciem integralnym. Jest to macierz stworzona z pikseli, których wartości są sumami intensywności wszystkich pikseli położonych po lewej i ponad pikselem odpowiadającym zdjęciu pierwotnemu.

Cechy obrócone o 45° , takie jak cechy liniowe wymagają innej zmienionej reprezentacji zdjęcia. Obrócone zdjęcie integralne jest obliczane jako suma intensywności pikseli położonych o 45° po lewej i ponad pikselem odpowiadającym dla wartości x oraz poniżej dla wartości y .



Sumowany obszar obrazu integralnego



Sumowany obszar obrazu integralnego obróconego

Detekcja obiektów – twarzy, elementów twarzy

Klasyfikator Haar'a

Zanim algorytm przystąpi do detekcji obiektów, należy zbudować odpowiednie wzorce cech, które powinien wyszukiwać. Budowa klasyfikatora opiera się o metodę Gentle AdaBoost. Obraz o typowej rozdzielczości 320x240 pikseli zawiera miliony cech Haar'a.

Do uczenia kaskady potrzebne są tysiące zdjęć wysokiej rozdzielczości, przedstawiające zarówno twarze jak i inne obiekty. Na obrazach zawierających twarze, należy wskazać ich położenie tak, aby algorytm miał informacje które rejony są celem detekcji.

Algorytm w działaniu jest kaskadowym algorytmem działającym na zasadzie przesuwanego okna.

Przykład

Śledzenie obiektów – twarzy, elementów twarzy

Histogram Based Tracker - algorytm CAMshift

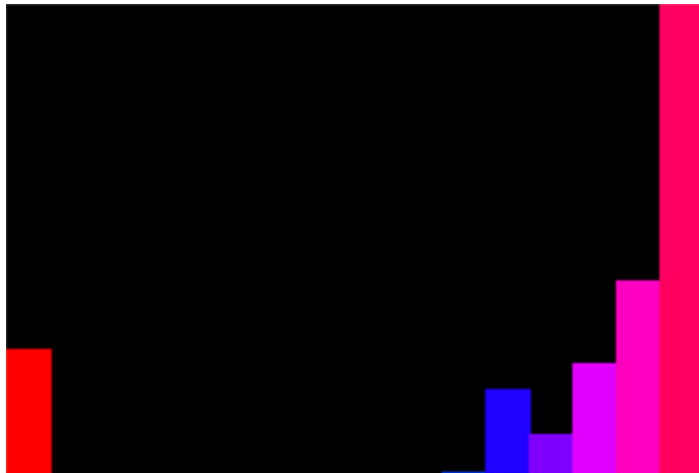
Histogram Based Tracker bazuje na algorytmie CAMshift, stworzonym przez Gary Bradski'ego z firmy Intel. Algorytm wykorzystuje informację o kolorze w video-sekwencji w celu lokowania i śledzenia obiektów. Program dla każdego piksela obrazu wyznacza prawdopodobieństwo przynależności do zbioru pikseli reprezentujących obiekt. Działa on w przestrzeni barw HSV, dokładniej na składowej H (odcień światła białego)

Działanie algorytmu CamShift można podzielić na cztery podstawowe kroki:
(Hewitt R.; , "How OpenCV's Face Tracker Works", SERVO Magazine, 2007)

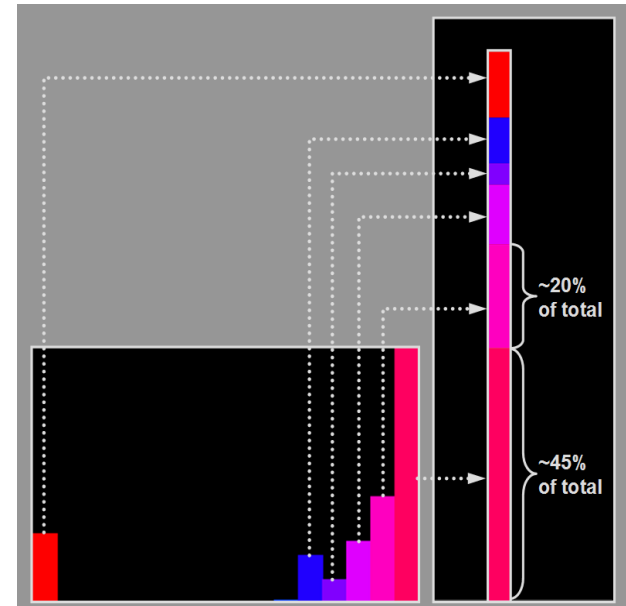
- Utworzenie histogramu kolorów reprezentującego twarz
- Obliczenie prawdopodobieństwa dla każdego piksela przychodzącej klatki
- Przesunięcie położenia prostokąta obrysującego twarz
- Obliczenie rozmiaru i kąta.

Śledzenie obiektów – twarzy, elementów twarzy

Histogram based tracker - algorytm CAMshift



Histogram twarzy w algorytmie CamShift



Przedstawienie histogramu jako udział pikseli w całości

źródło: Hewitt R. "How OpenCV's Face Tracker Works", SERVO Magazine, 2007

Śledzenie obiektów – twarzy, elementów twarzy

Histogram based tracker - algorytm CAMshift



Klatka z kamery



Obraz w reprezentacji
prawdopodobieństwa pikseli

źródło: Błaszowski P. Wykrywanie, rozpoznawanie i śledzenie ruchomych obiektów poprzez niezależną platformę monitorującą, przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, Politechnika Gdańska 2013

Przykład

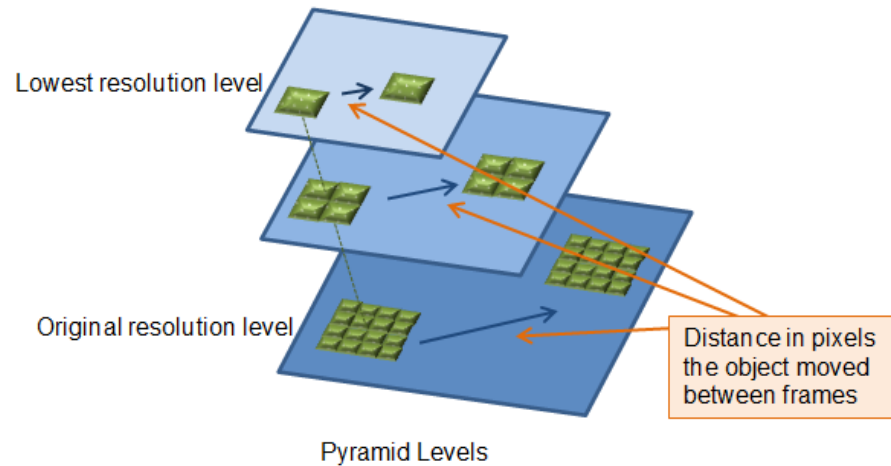
Śledzenie obiektów – twarzy, elementów twarzy

Point tracker – algorytm Kanade-Lucas-Tomasi

Point Tracker śledzi zbiór punktów wykorzystując algorytm Kanade-Lucas-Tomasi. Algorytm ten operuje na intensywnościach pikseli. Każdy śledzony punkt posiada pewne sąsiedztwo (domyślnie 31x31 pikseli), w którym algorytm iteracyjnie poszukuje jego nowej lokalizacji na kolejnych klatkach. Tracker generuje piramidę zdjęć, w której każdy poziom charakteryzuje się rozdzielczością zmniejszoną o połowę względem poprzedniego. Pozwala to na śledzenie punktu na kilku poziomach rozdzielczości, co umożliwia obsługiwanie dużych przemieszczeń punktów (większych niż ich sąsiedztwo).

Śledzenie obiektów – twarzy, elementów twarzy

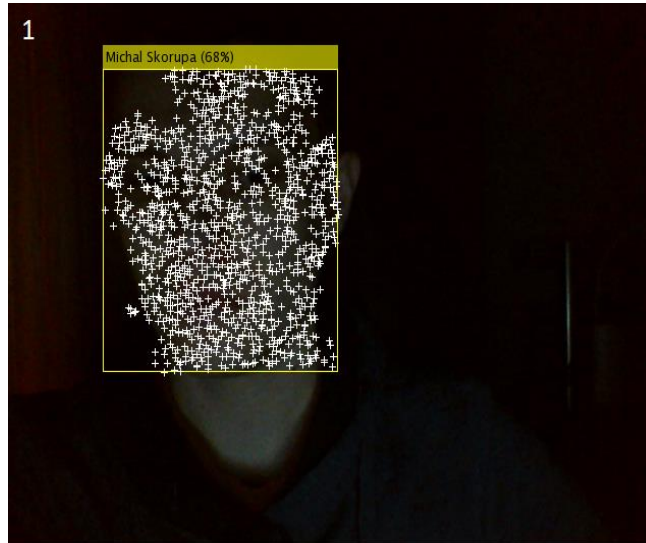
Point tracker – algorytm Kanade-Lucas-Tomasi



źródło: <http://www.mathworks.com/help/vision/ref/vision.pointtrackerclass.html>

Śledzenie obiektów – twarzy, elementów twarzy

Point tracker – algorytm Kanade-Lucas-Tomasi



źródło:

Skorupa, M., Lewicki, D., Lewicki, T. Wykorzystanie metod przetwarzania obrazów i inteligencji obliczeniowej do identyfikacji biometrycznej osób na podstawie twarzy. Praca inżynierska. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki, Katedra Inżynierii Systemów Sterowania.

Rozpoznawanie twarzy

Najważniejsze metody rozpoznawania twarzy:

- Eigenfaces;
- LBBH;
- FisherFaces;
- Gradient;
- Length;
- Resize;
- Histogram;
- Neural Networks;
- Convolutional Neural Networks (ConvNets, CNNs), Deep Neural Networks.

Rozpoznawanie twarzy

Metoda Resize

Najprostsza z metod. Polega na kolejnej redukcji wymiaru danych opisujących zdjęcie (skala odcieni szarości, wyrównywanie histogramem, binaryzacja, siatka redukcyjna)

Proces skalowania obrazów:

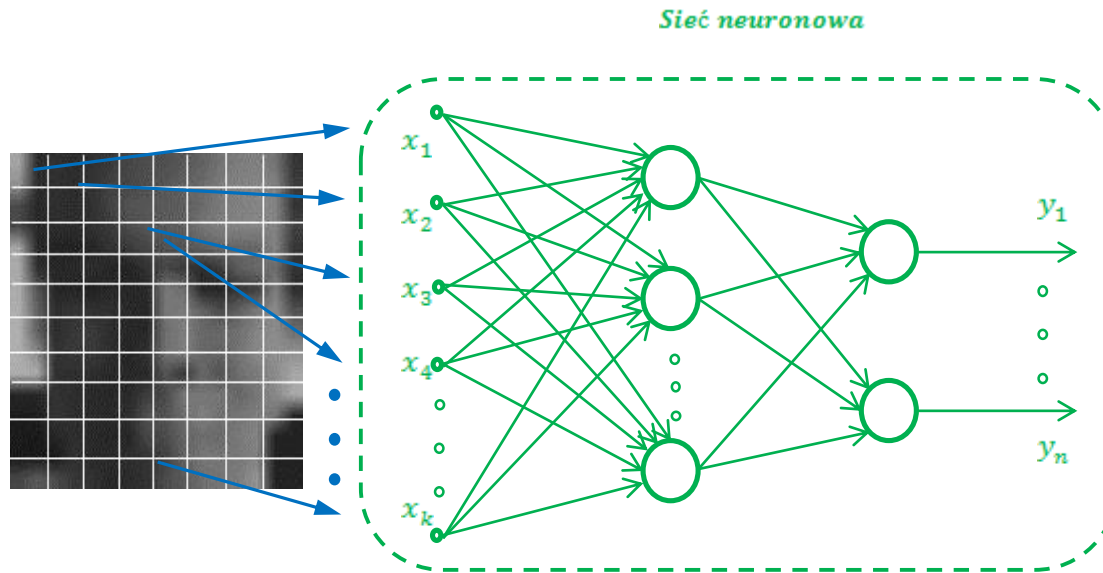
- oryginalnego (a)
- do obrazu w skali szarości (b)
- o wymiarach 32x32 pikseli (c)
- o wymiarach 16x16 pikseli (d)



źródło: Sikora M. *Inteligentny system rozpoznawania twarzy w czasie rzeczywistym – implementacja sprzętowa*. Praca magisterska, Politechnika Gdańska 2012.

Rozpoznawanie twarzy

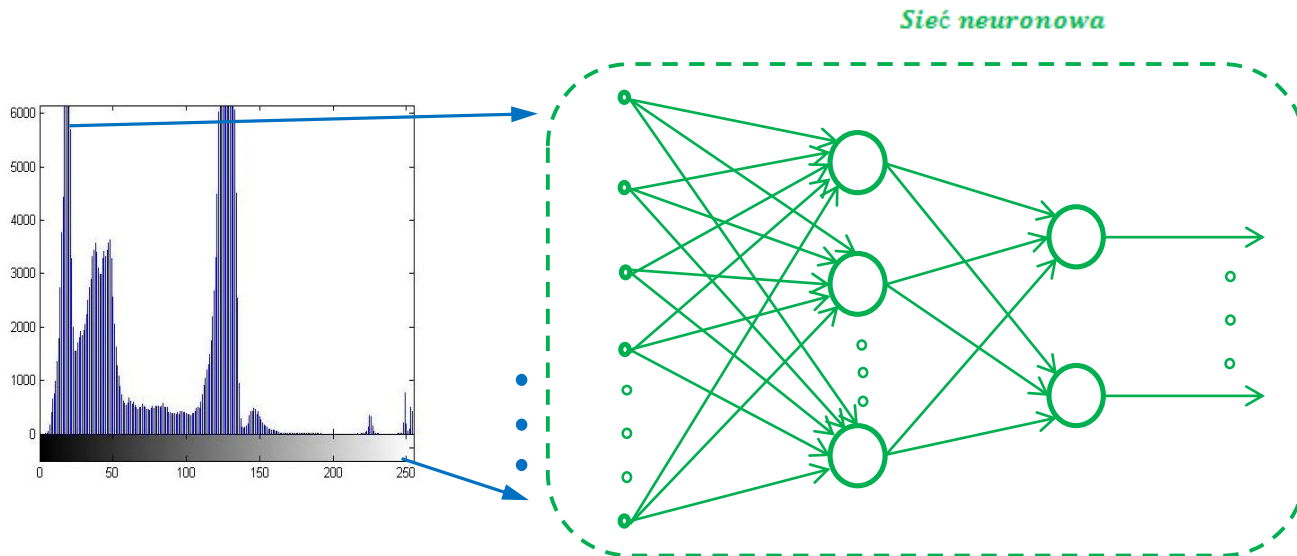
Metoda Resize z klasyfikatorem w postaci sieci neuronowych



Rozpoznawanie twarzy

Metoda oparta o histogramy z klasyfikatorem w postaci sieci neuronowych

Polega na przedstawieniu zdjęć w postaci histogramów i przekazywaniu ich jako wejścia do sieci neuronowej, która służy za klasyfikator.



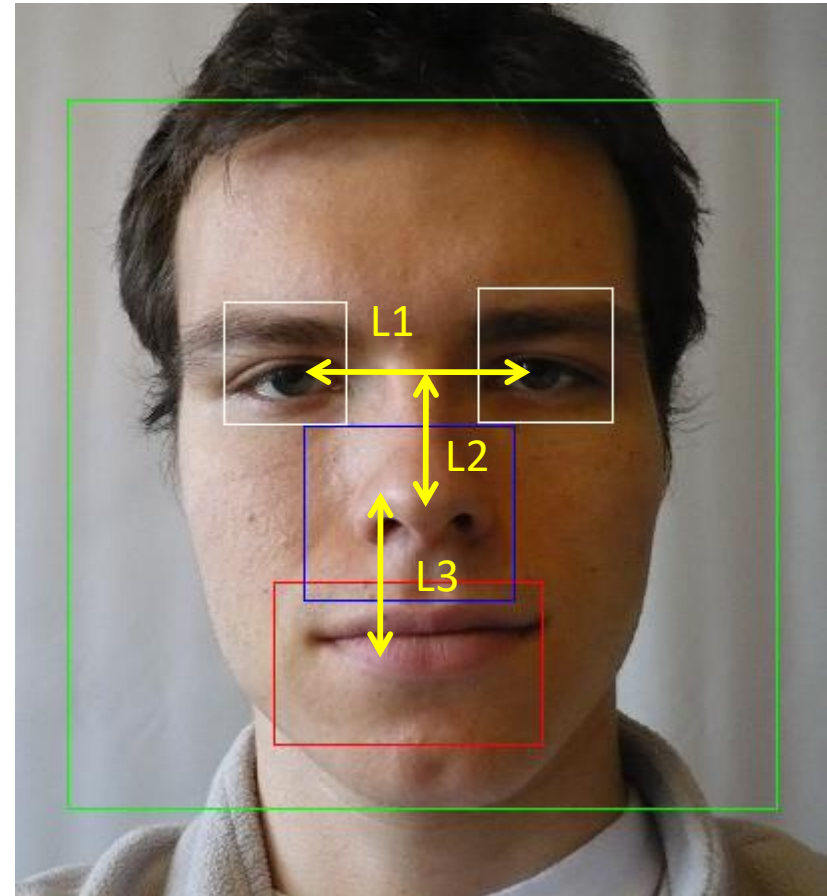
Rozpoznawanie twarzy

Metoda Length

Metoda bazuje na obliczaniu odległości pomiędzy charakterystycznymi punktami twarzy człowieka.

Do poprawnego zadziałania tej metody musimy z wynikiem pozytywnym odnaleźć na twarzy danej osoby takie elementy jak: obydwie oczy, nos oraz usta.

Przy wyszukiwaniu ww. elementów najlepiej skorzystać z estymatora Haar'a.



źródło: Sikora M. *Inteligentny system rozpoznawania twarzy w czasie rzeczywistym – implementacja sprzętowa*. Praca magisterska, Politechnika Gdańska 2012.

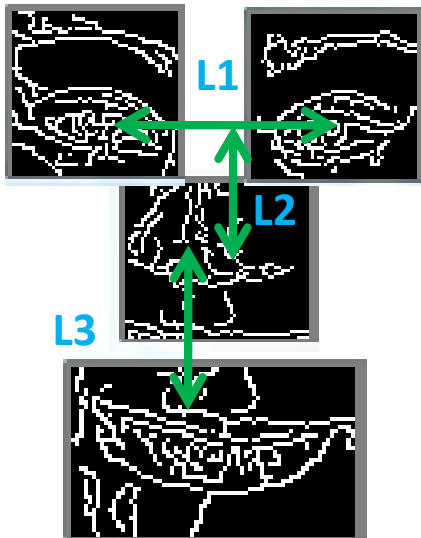
Rozpoznawanie twarzy

Metoda Length

Posiadając odnalezione te elementy przystępujemy do obliczenia środka każdego z nich.

Następnie obliczana jest odległość pomiędzy środkiem każdego z tych elementów.

W celu uniezależnienia wyników od odległości od kamery, a jedynie od kształtu ludzkiej twarzy należy obliczyć stosunki tych odległości.



Wykorzystanie trzech zależności:

- $L1/L2$;
- $L1/L3$;
- $L2/L3$.

Rozpoznawanie twarzy

Metoda LBPH

Na podstawie: Pietikäinen M.; Hadid A.; Zhao G.; Ahonen T.; "Computer Vision Using Local Binary Patterns", Machine Vision Group, University of Oulu, Finland, 2011

Local Pattern Binary Histogram jest metodą pozwalającą na zmianę zdjęcia w macierz wartości całkowitych opisujących mało-wymiarowe cechy obrazu. Histogram tych wartości jest następnie wykorzystywany do dalszej analizy.

LBPH nie bazuje na statystyce, więc nie jest potrzebna duża ilość zdjęć uczących.

Metoda LPB może równie dobrze działać przy tylko jednej próbce dla każdej z klas.

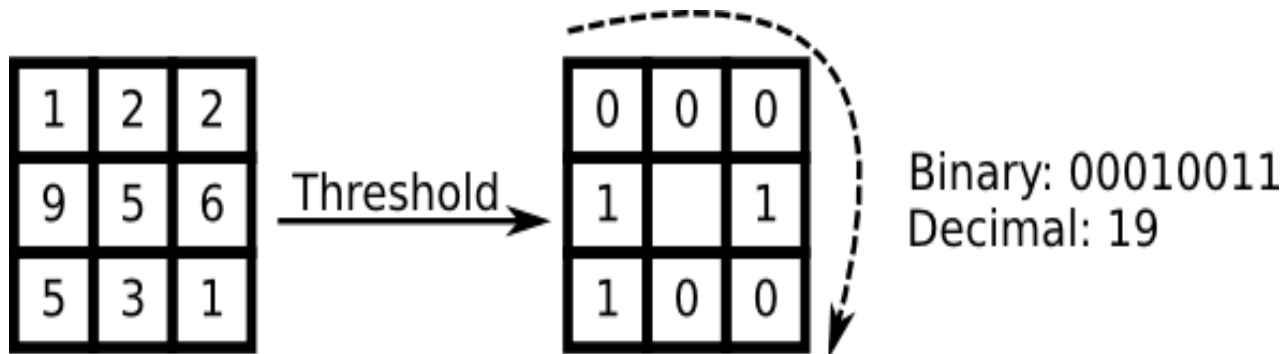
LBPH nie analizuje obrazu jako wielowymiarowego wektora, a opisuje go jedynie w postaci lokalnych cech obiektów.

Rozpoznawanie twarzy

Metoda LBPH

Główną ideą algorytmu jest sumowanie lokalnej struktury zdjęcia poprzez porównywanie każdego piksela z jego sąsiedztwem. Wybierany jest środkowy piksel kwadratowego obszaru (w wersji podstawowej 3x3 pikseli), a jego sąsiedztwo jest poddawane progowaniu. W zależności od tego, czy dany sąsiadujący piksel jest większy od progu, czy nie, przyjmuje wartość 1 lub 0. Następnie odczytuje się liczbę binarną zapisaną dookoła środkowego piksela. Dla ośmiu pikseli sąsiadujących istnieje 256 kombinacji, zwanych *Local Binary Patterns*.

Sposób analizy ramek 3x3 w metodzie LBPH

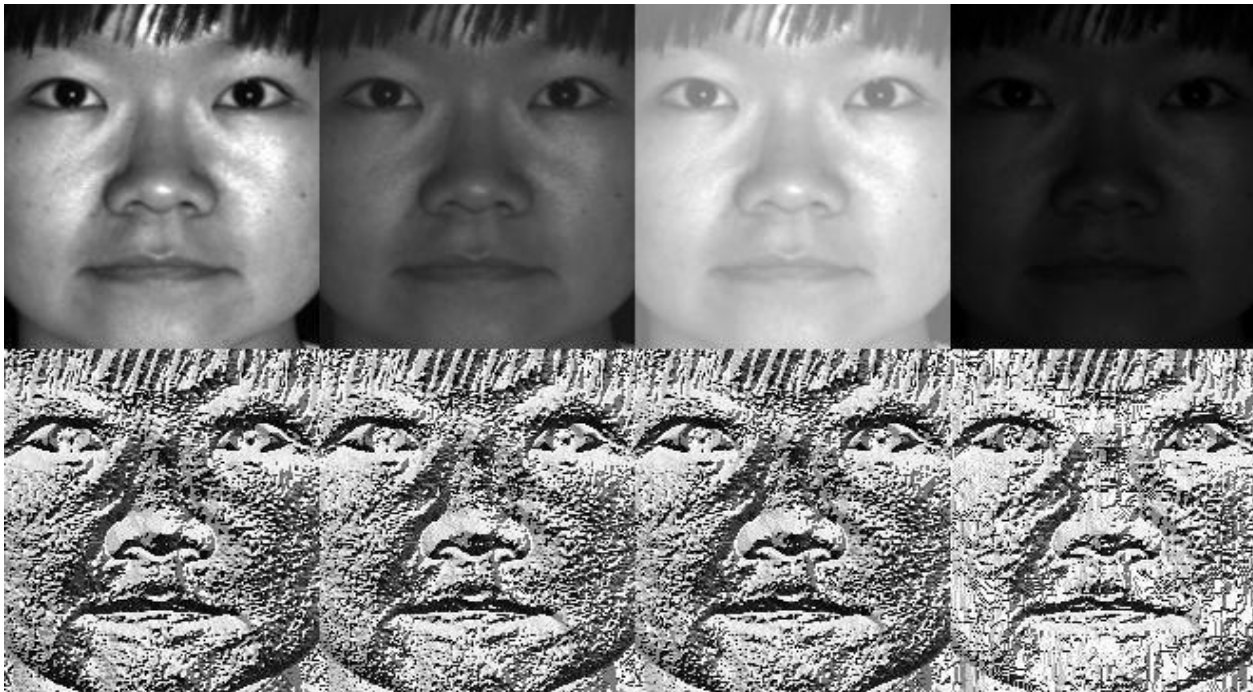


źródło: Hewitt R. "How OpenCV's Face Tracker Works", SERVO Magazine, 2007

Rozpoznawanie twarzy

Metoda LBPH

Przykład transformacji twarzy metodą Local Pattern Binary Histogram
w różnych warunkach oświetleniowych



źródło: Dokumentacja OpenCV
www.opencv.org

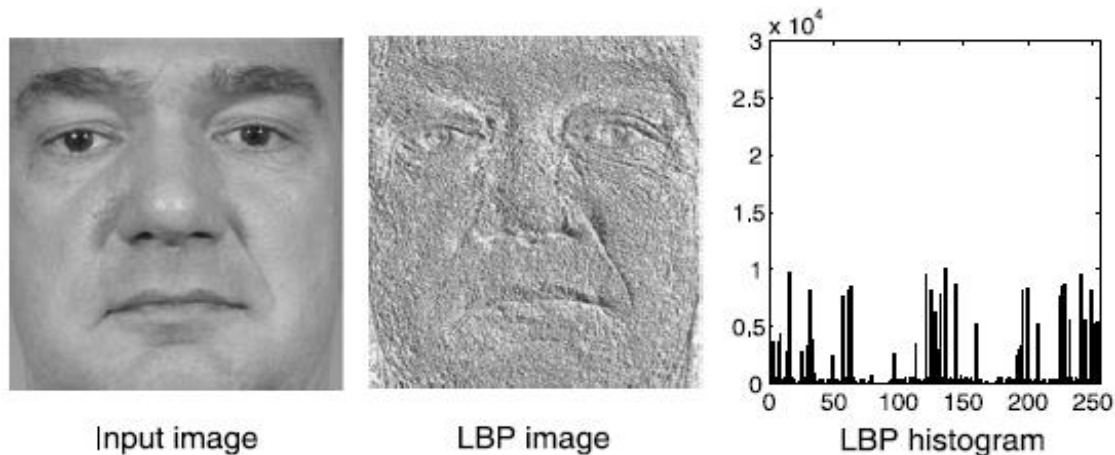
Rozpoznawanie twarzy

Metoda LBPH

Na podstawie: Pietikäinen M.; Hadid A.; Zhao G.; Ahonen T.; "Computer Vision Using Local Binary Patterns", Machine Vision Group, University of Oulu, Finland, 2011

Jako klasyfikator do analizy zdjęć w dziedzinie LBPH, dzieli się obraz na pewną liczbę obszarów i wykorzystuje się porównywanie histogramów.

Histogramy zapisane w postaci wektora, o kolejnych elementach będących liczbami pikseli w danym zakresie intensywności, mogą być porównywane w prosty sposób, za pomocą metody Najbliższego Sąsiada. Niekiedy stosuje się także bardziej zaawansowane techniki, jak PCA czy LDA w celu przeniesienia histogramu do innej przestrzeni.



źródło: Dokumentacja OpenCV
www.opencv.org

Rozpoznawanie twarzy

Metoda Eigenfaces

Na podstawie:

- *Błaszowski P. Wykrywanie, rozpoznawanie i śledzenie ruchomych obiektów poprzez niezależną platformę monitorującą, przy wykorzystaniu sztucznych sieci neuronowych. Praca magisterska, Politechnika Gdańska 2013;*
- *Turk M.; Pentland A.; , "Eigenfaces for Recognition", Journal of Cognitive Neuro-science, 1991.*

Metoda Eigenfaces wykorzystuje PCA w odniesieniu do przetwarzania obrazów, jednak przetwarzanie obrazów wymaga znacznie więcej obliczeń niż to występuje w przy przetwarzaniu „przemysłowych” danych pomiarowych.

Przykładowo zdjęcie o wymiarach 200x200 pikseli, stosując klasyczne PCA, wymagałoby obliczenia macierzy kowariancji o wymiarach 40 000 x 40 000 oraz wektorów własnych o takich samych długościach.

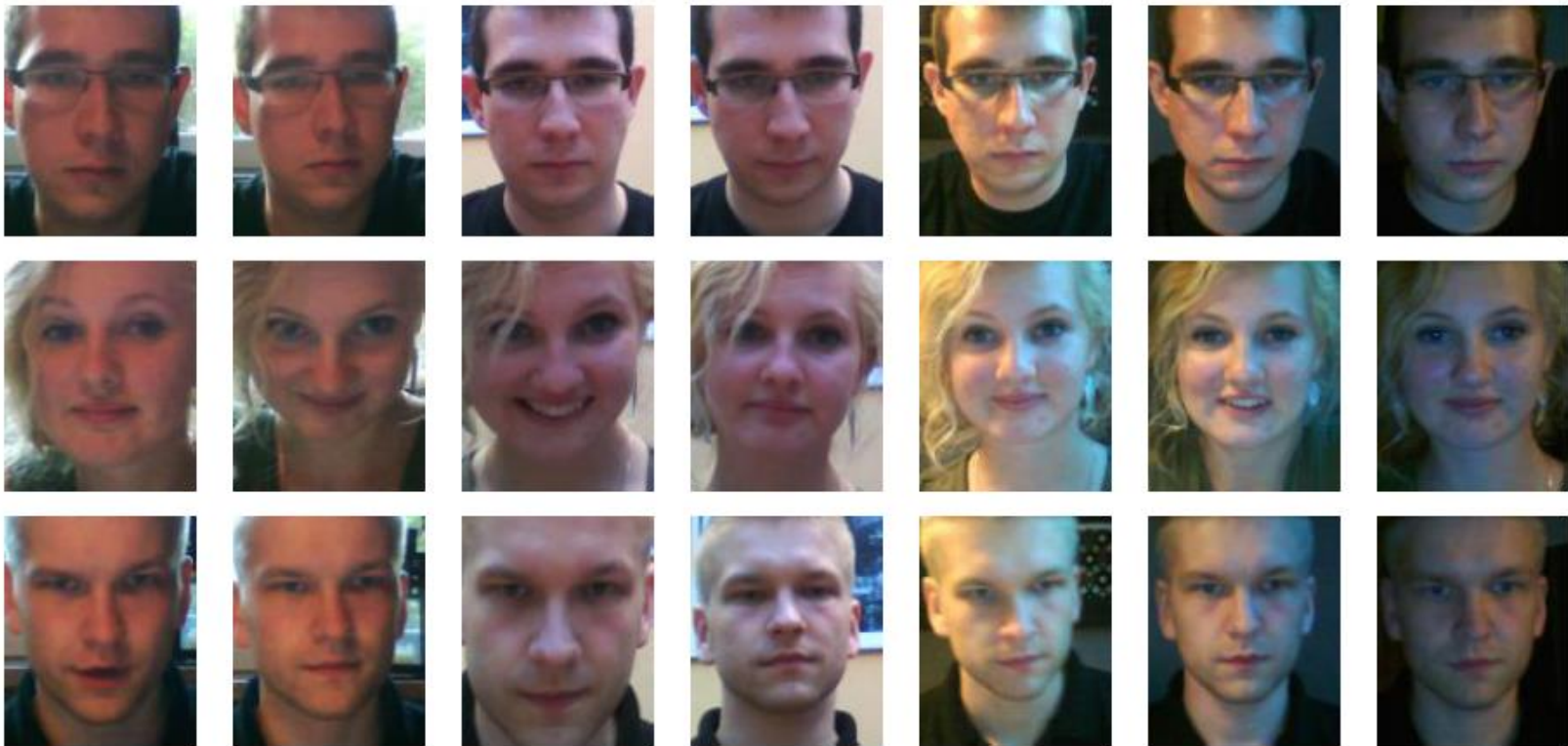
W związku z tym metoda Eigenfaces zawiera szereg modyfikacji przystosowującej algorytm PCA do efektywnej obróbki tak dużych zbiorów danych.

Rozpoznawanie twarzy

Metoda Eigenfaces

Pierwszą fazą algorytmu jest przygotowanie zbioru zdjęć uczących.

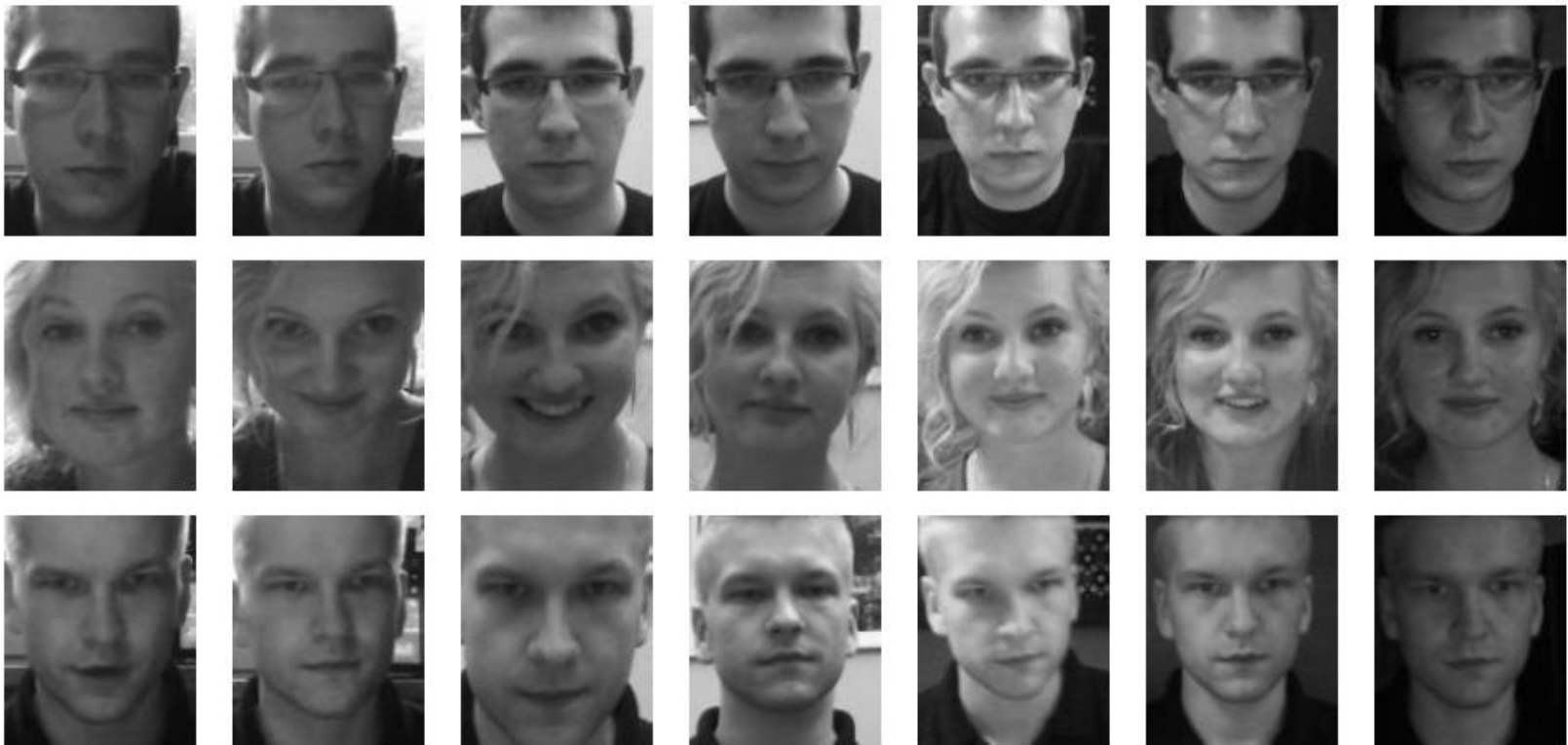
Przykładowy zbiór składający się z $M=21$ zdjęć, o rozmiarach: 100x120 pikseli



Rozpoznawanie twarzy

Metoda Eigenfaces

Kolejną fazą algorytmu jest przekonwertowanie zbioru zdjęć uczących do skali odcieni szarości.



Rozpoznawanie twarzy

Metoda Eigenfaces

Następnie należy obliczyć zdjęcie, będące średnią arytmetyczną wszystkich zdjęć uczących.



Ostatnim etapem przygotowywania próbek jest odjęcie od wszystkich zdjęć zdjęcia średniego.

Rozpoznawanie twarzy

Metoda Eigenfaces – Algorytm

Kolejne zdjęcia uczące (w skali szarości) zapisywane są w postaci wektorów: $\Gamma_1 \dots \Gamma_M$ o N^2 elementach, średnie zdjęcie twarzy Ψ można przedstawić jako:

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$$

Różnica pomiędzy poszczególnymi zdjęciami a wartością średnią Φ_i wynosi:

$$\Phi_i = \Gamma_i - \Psi$$

Kolejnym etapem algorytmu jest stworzenie macierzy A , składającej się z kolejnych znormalizowanych wektorów zdjęć:

$$A_{N^2M} = [\Phi_1 \ \Phi_2 \ \dots \ \Phi_M]$$

Rozpoznawanie twarzy

Metoda Eigenfaces – Algorytm

W klasycznym PCA macierz kowariancji (C) oblicza się, mnożąc macierz A przez jej transpozycję:

$$C = A_{N^2M} * A_{MN^2}^T$$

Następnie z tej macierzy oblicza się wektory i wartości własne. W przypadku przetwarzania obrazów daje to ogromną macierz o wymiarach $N^2 \times N^2$, której obliczenie jest czasowo (pamięciowo...) wymagające, a czasami (często) wręcz niemożliwe do wykonania.

W metodzie Eigenfaces stosuje się pewien trick a mianowicie macierz kowariancji (L) oblicza się, dokonując transpozycji macierzy A i A^T , w następujący sposób:

$$L = A_{MN^2}^T * A_{N^2M}$$

Prowadzi to do zmniejszenia wymiarów problemu do $M \times M$. Operacja ta nie powoduje straty informacji, a jednocześnie znacznie upraszcza analizę.

Rozpoznawanie twarzy

Metoda Eigenfaces – Algorytm

Następnie, ze zmodyfikowanej macierzy kowariancji L oblicza się wektory własne v i wartości własne λ , oraz wybiera się najważniejsze składniki główne. Liczba wektorów własnych potrzebnych do dobrego odzwierciedlenia zdjęcia jest znacznie mniejsza od liczby wszystkich wektorów własnych.

Wektory i wartości własne zmienionej macierzy kowariancji $A^T A$, można przedstawić za pomocą:

$$A A^T v_i = \lambda_i v_i$$

Mnożąc to równanie obustronnie przez A otrzymujemy:

$$A A^T A v_i = \lambda_i A v_i$$

Z ostatniego równania można wywnioskować, że $A v_i$ są wektorami własnymi macierzy $C = A A^T$.

Wynika z tego, że po analizie PCA w zmienionej przestrzeni, możliwe jest powrót do właściwych wektorów własnych u za pomocą:

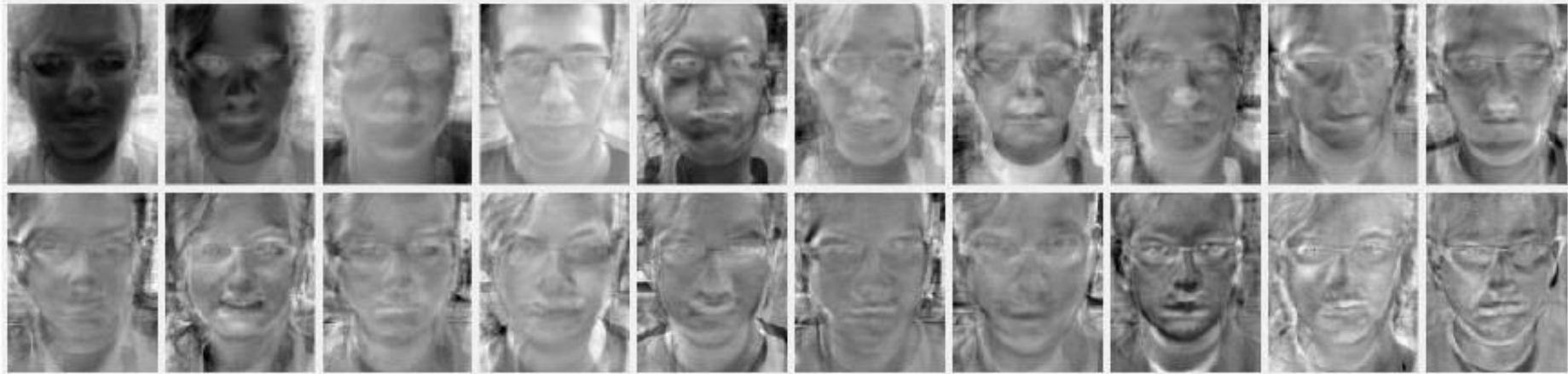
$$u_i = A v_i$$

Rozpoznawanie twarzy

Metoda Eigenfaces – Algorytm

Wektory składników głównych $u_{N \times 1}^2$ można przedstawić w postaci tzw. twarzy własnych (Eigen faces) o wymiarach $N \times N$.

Dla omawianego przykładu, zbiór dwudziestu wybranych twarzy własnych wygląda następująco:



Dla analizowanego przykładu, takie podejście redukuje rozmiar macierzy kowariancji poddanej PCA z 12000×12000 do 20×20 . Spośród 12000 wektorów własnych, 20 wektorów wystarcza (???) do zadawalającego przedstawienia obrazu twarzy.

Rozpoznawanie twarzy

Metoda Eigenfaces – Algorytm

Po wyznaczeniu twarzy własnych, należy wszystkie zdjęcia uczące przetransformować do przestrzeni cech:

$$\omega_k = u_k^T (\Gamma - \Psi)$$

Zdjęcia w przestrzeni cech są opisane jako wektory wag:

$$\Omega = [\omega_1 \cdots \omega_k]$$

z których każda waga ω_k niesie informację o udziale danej twarzy własnej w obrazie. Długość wektora wag jest zatem równa liczbie twarzy własnych.

Transformacja wszystkich zdjęć uczących daje bazę danych do identyfikacji nieznanymi twarzy.

Rozpoznawanie twarzy

Metoda Eigenfaces

Przykładowa twarz i jej wektor wag utworzona przy wykorzystaniu 20 twarzy własnych

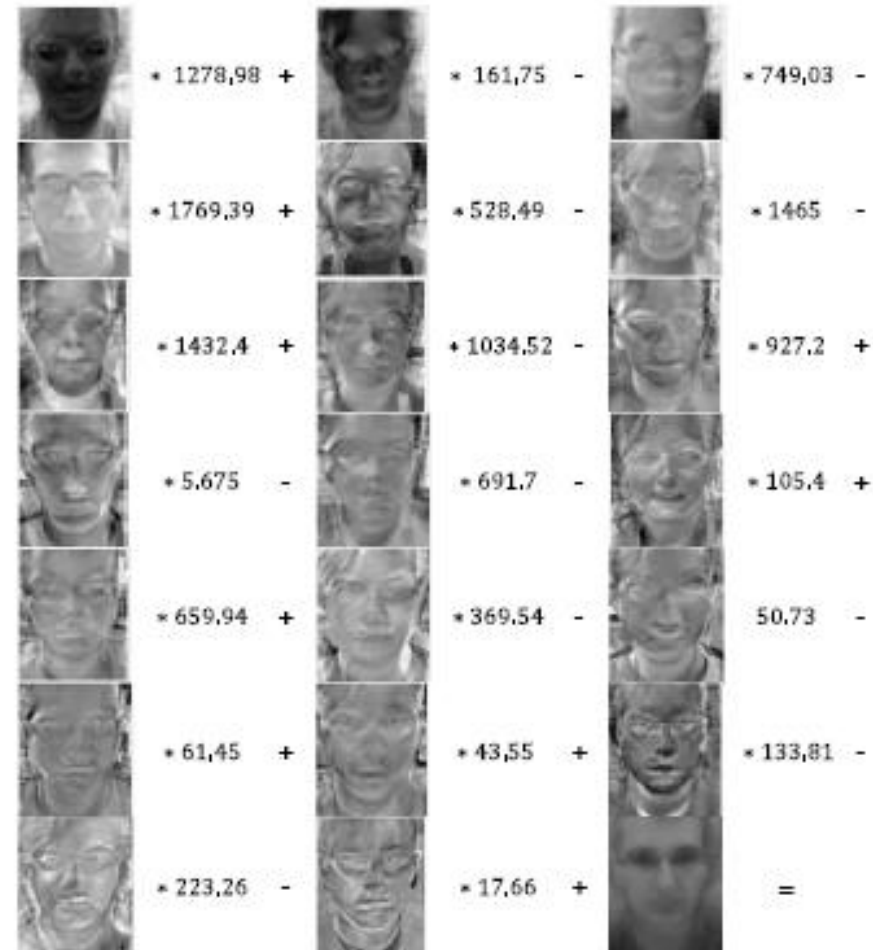


$$\Omega = [1278.98; 161.75; -749.03; -1769.39; 528.49; \\ -1465; -1432.4; 1034.52; -927.2; 5.675; -691.7; -105.4; \\ 659.94; 369.54; -50.73; -61.45; 43.55; 133.81; -223.26; \\ -17.66]$$

Monitorowanie i diagnostyka w systemach sterowania

Przykładowa twarz i jej wektor wag utworzona przy wykorzystaniu 20 twarzy własnych

Rekonstrukcja twarzy polega na odpowiednim przemnożeniu poszczególnych elementów wektora wag przez odpowiadające twarze główne oraz dodaniu twarzy średniej.

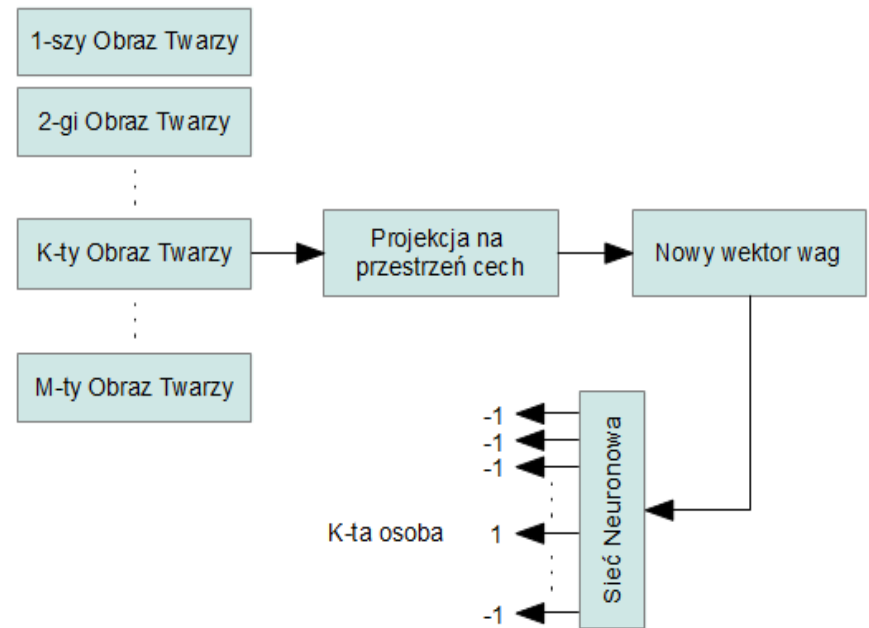


Rozpoznawanie twarzy

Metoda Eigenfaces + sieci neuronowe

Eigenfaces nie pozwala na identyfikację twarzy a jedynie na zmianie ich reprezentacji (głównie rozmiaru).

Celem rozpoznania twarzy należy użyć jakiegoś klasyfikatora np. metody najbliższego sąsiada (odległość Euclidesowa) lub sieci neuronowych.



Proces przygotowania danych do uczenia sieci neuronowej

źródło: Turk M.; Pentland A.; , "Eigenfaces for Recognition", Journal of Cognitive Neuro-science, 1991

Przykład

Dziękuję za uwagę