

Rozpoznawanie obrazów

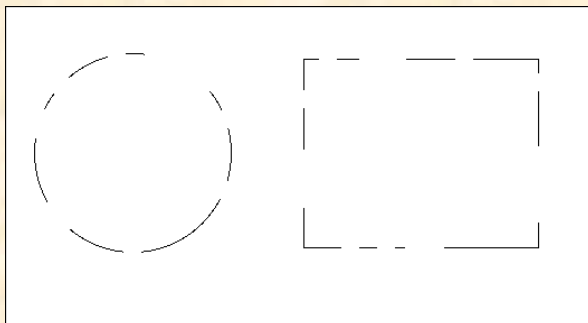
Końcowym etapem analizy obrazów może być nie tylko identyfikacja obiektów analizy oraz ich ilościowa charakterystyka ale również klasyfikacja obiektów i ich symboliczna interpretacja, te ostatnie często są określane terminem *rozpoznania obrazu*.

Z powyższych względów, faza rozpoznawania obrazów zwykle nie wykorzystuje gotowych procedur analizy danych a wymaga raczej specjalizowanych, bardziej zaawansowanych metod analizy, np. bazujących na sztucznej inteligencji.

Rozpoznawanie obrazów

Cechy zachowania inteligentnego:

- ⇒ możliwość wnioskowania na podstawie zbioru różnych, nieskojarzonych ze sobą danych,
- ⇒ zdolność do uczenia się na przykładach i generalizacji nabytej wiedzy (przez zastosowanie jej w innych zadaniach z innymi danymi),
- ⇒ zdolność rozpoznawania obiektów (interpretacji informacji) na podstawie niekompletnych danych.



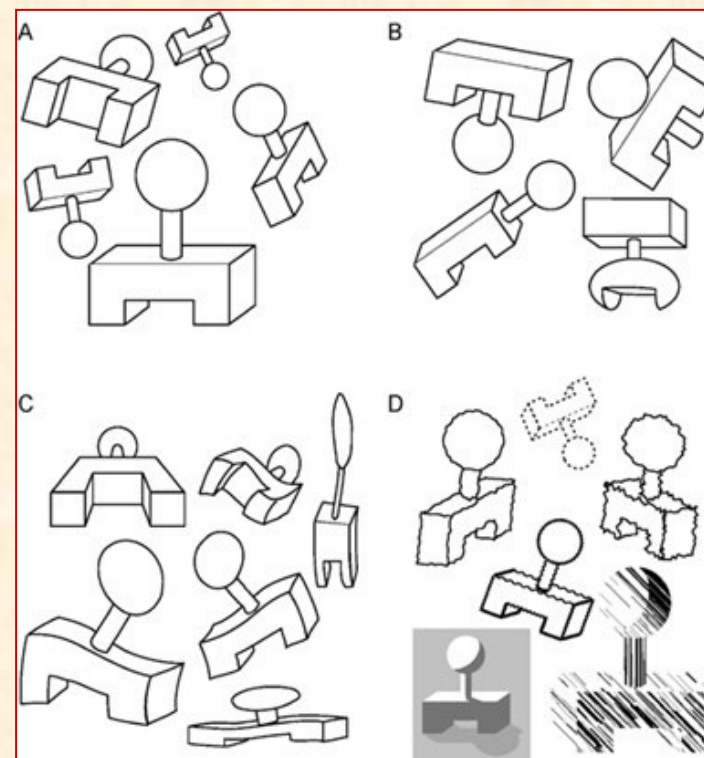
Przykład trudnego problemu rozpoznawania obrazu



Wyłanianie

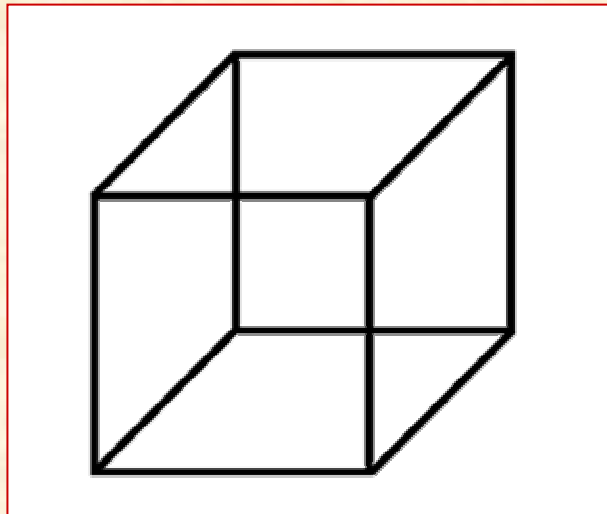
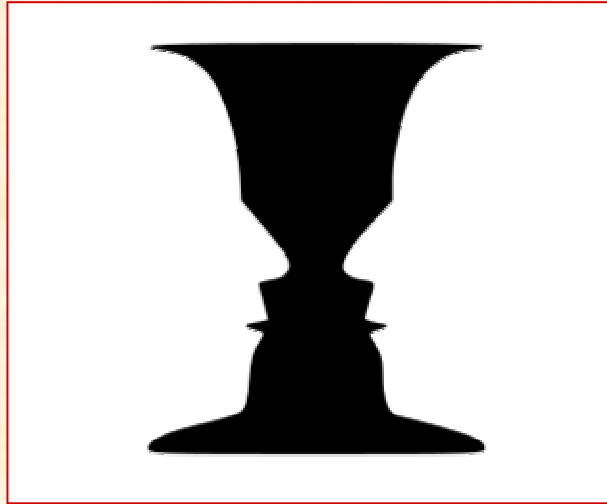


Przykład trudnego zadania rozpoznawania obrazu (wymagającego znamion zachowania inteligentnego)

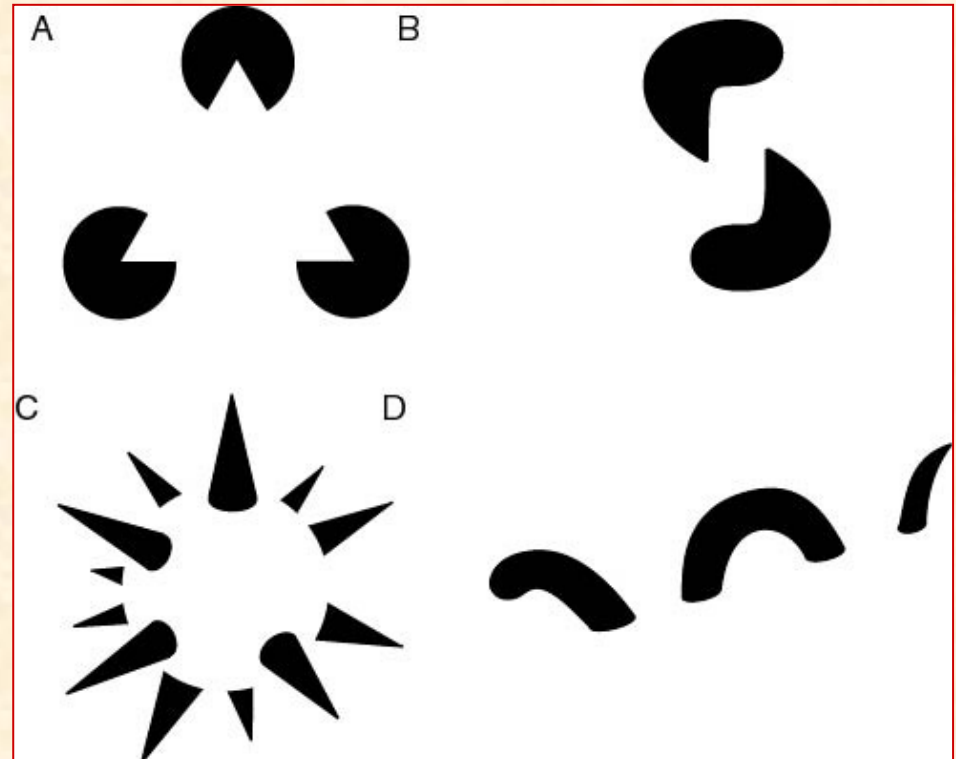


Niezmiennność

Teoria Gestalt (niem. *postać*)



Wieloznaczność



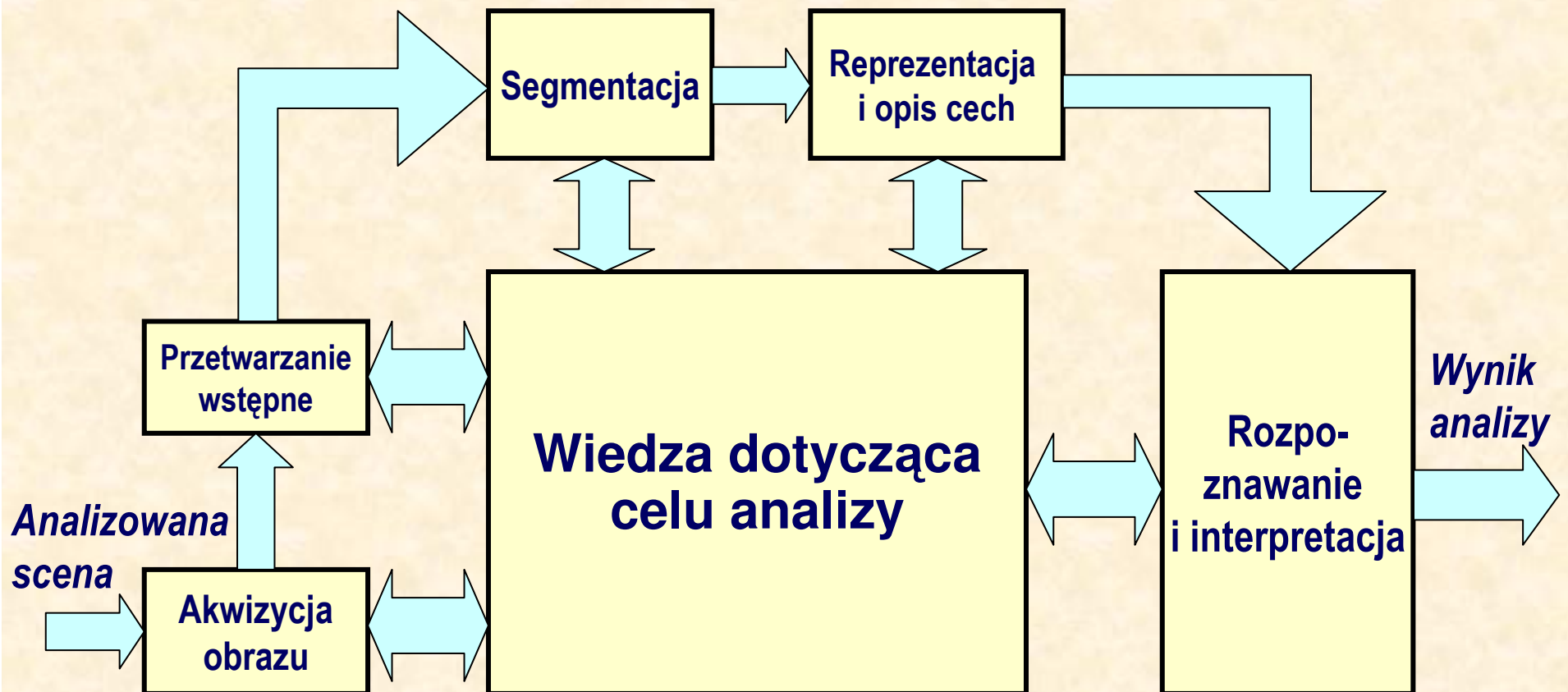
Reifikacja
(iluzoryczne obiekty)

System rozpoznawania obrazów

Elementy składowe kompletnego systemu rozpoznawania obrazów:

- ***przetwarzanie niskiego poziomu*** - obejmuje akwizycję obrazu, przetwarzanie wstępne, poprawę jakości obrazu, np. eliminacja zakłóceń, poprawa kontrastu, filtracja itd.,
- ***przetwarzanie średniego poziomu*** - dotyczy segmentacji obrazu oraz wydzielania i opisu cech obiektów obrazu, np. detekcja brzegów i konturów, przetwarzanie morfologiczne, itd.,
- ***przetwarzanie wysokiego poziomu*** - polega na klasyfikacji, rozpoznawaniu i interpretacji analizowanej sceny.

Elementy systemu rozpoznawania obrazów



Wzorce i klasy wzorców

Wzorzec to zbiór cech, który tworzy ilościowy i jakościowy opis obiektu; ściślej, wzorzec to wektor cech $\mathbf{x}=[x_1, x_2, \dots, x_N]$.

Klasa wzorców to zbiór wzorców charakteryzujących się podobnymi wektorami cech. Klasy wzorców oznaczmy $\omega_1, \omega_2, \dots, \omega_M$ gdzie indeks M jest numerem klasy.

Rozpoznawanie wzorców (nazywane też klasyfikacją) jest zadaniem polegającym na przyporządkowaniu wzorców do ich klas:

$$\mathbf{x} \rightarrow \omega$$

tj. przekształceniem przestrzeni wektorów cech X na przestrzeń klas wzorców Ω .

Klasyfikacja wzorców

Przestrzeń cech

Przestrzeń klas



Klasyfikacja wzorców

Przyporządkowanie $\mathbf{x} \rightarrow \omega$ powinno być bezbłędne dla jak największej liczby wzorców. Zagadnienie znalezienia najlepszego takiego przyporządkowania jest zadaniem optymalizacji statystycznej. Konkretnie sformułowanie tego zadania zależy od stopnia posiadanej wiedzy (modelu) o rozkładzie statystycznym zbioru cech, jak również granicach klas.

W przypadku, gdy rozkłady cech są trudne do zamodelowania lub wiedza o ich rozkładzie statystycznym jest niedostępna, klasyfikator, tj. przyporządkowanie $\mathbf{x} \rightarrow \omega$, może być zbudowany przez zastosowanie algorytmów uczących, samodzielnie wypracowujących reguły klasyfikacji na podstawie reprezentatywnego zbioru wzorców (wektorów cech).

Selekcja cech i ich właściwości

- **dyskryminacja** - cechy powinny przyjmować znacząco różne wartości dla obiektów z różnych klas, np. średnica owocu jest dobrą cechą dla rozróżnienia wiśni i grejfrutów,
- **niezawodność** - cechy powinny przyjmować podobne wartości dla wszystkich obiektów danej klasy, np. kolor jest złą cechą dla jabłek,
- **niezależność** - cechy wykorzystywane w danym systemie klasyfikacji powinny być nieskorelowane ze sobą, np. waga i wielkość owocu są cechami silnie skorelowanymi,
- **mała liczba** - złożoność systemu klasyfikacji rośnie szybko wraz z liczbą klasyfikowanych cech, np. należy eliminować cech skorelowane.

Selekcja cech i ich właściwości

W praktyce testuje się wybrany intuicyjnie zbiór cech (wzorzec), którego rozmiar zostaje zredukowany do akceptowalnej wielkości.

Selekcja cech może polegać na eliminacji cech o „najgorszych właściwościach”, przy zachowaniu wymaganej jakości systemu klasyfikacji.

Badanie jakości klasyfikacji dla wszystkich możliwych kombinacji podzbioru n cech ze zbioru wszystkich N cech jest w praktyce zbyt kosztowne obliczeniowo dla dużych N .

Selekcja cech i ich właściwości

Współczynnik korelacji cech x i y :

$$\sigma_{xy} = \frac{\frac{1}{P} \sum_{i=1}^P (x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y}$$

gdzie: P - liczna klasyfikowanych obiektów, μ , σ oznaczają odpowiednio wartości średnie i odchylenie standardowe danego zbioru cech. Jeżeli wsp. Korelacji jest bliski 1 (-1) cechy x i y uważa się za silnie skorelowane (np. jedną z nich można odrzucić).

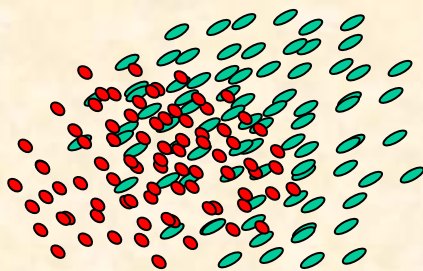
Selekcja cech i ich właściwości

Przykład miary separacji cechy x pomiędzy klasami j i k :

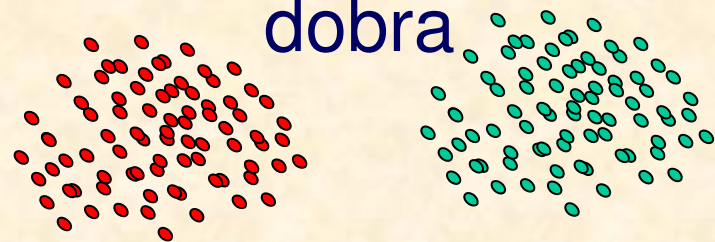
$$\hat{D}_{xjk} = \frac{|\mu_{xj} - \mu_{xk}|}{\sqrt{\sigma_{xj}^2 + \sigma_{xk}^2}}$$

Duża wartość tej miary świadczy o dobrej separacji cechy x pomiędzy klasami j i k .

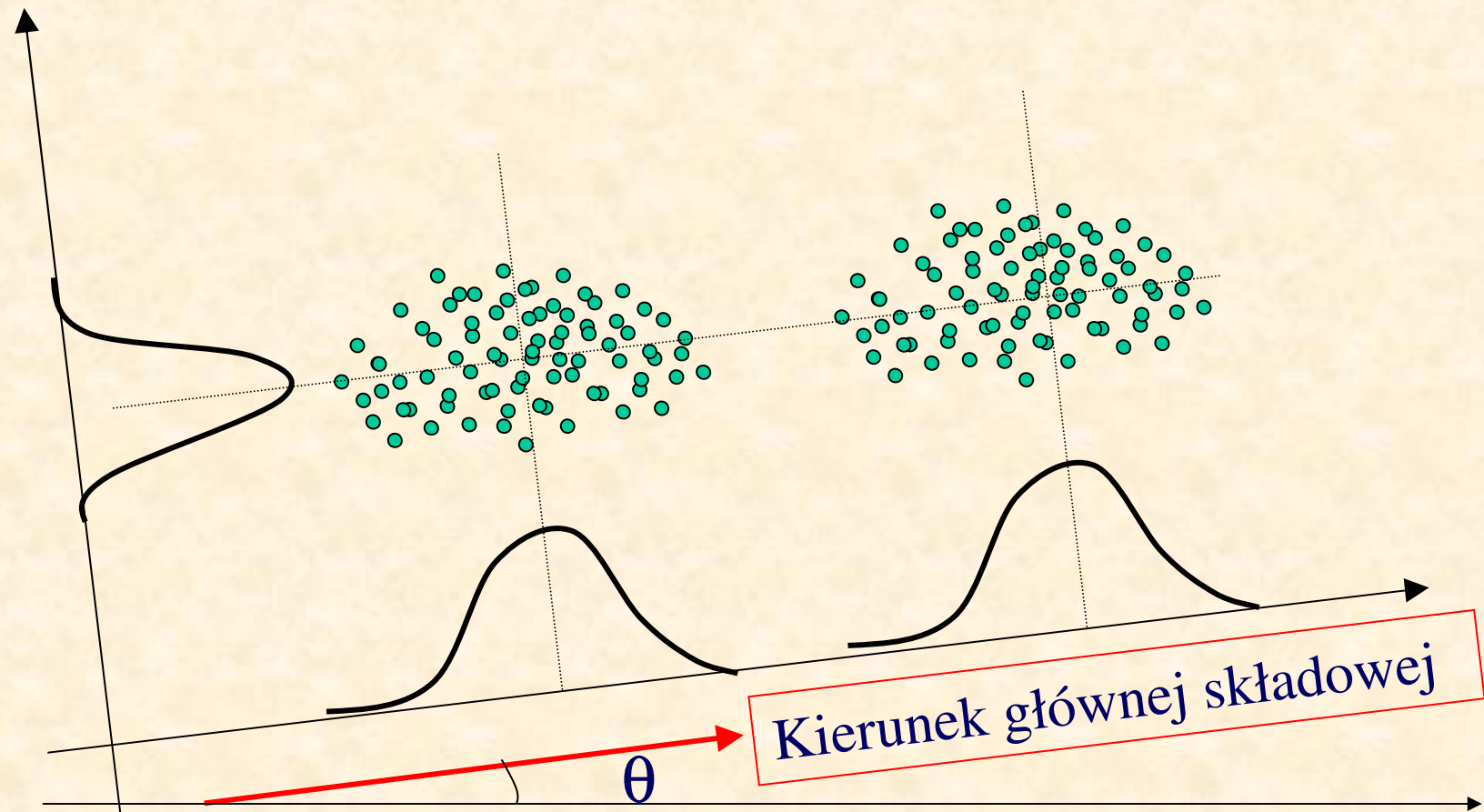
zła



dobra



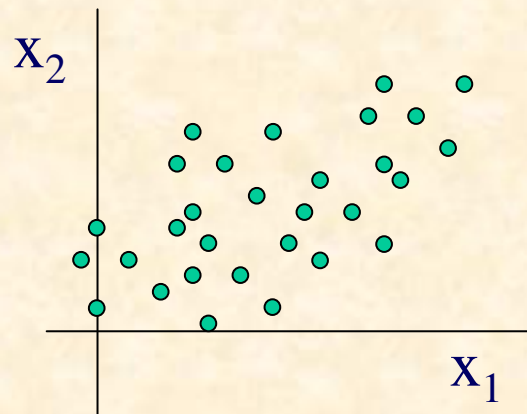
Redukcja wymiarowości



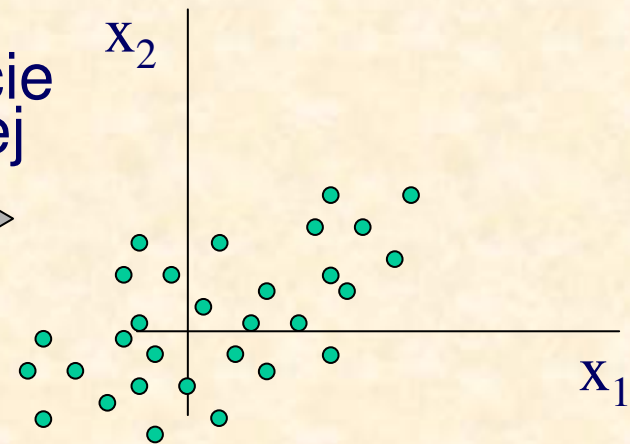
Kierunek głównej składowej

Należy tak dobrać θ aby separacja cech była jak najlepsza

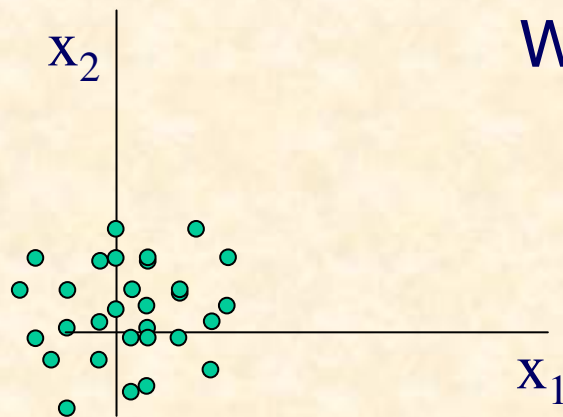
Wstępne przetworzenie cech



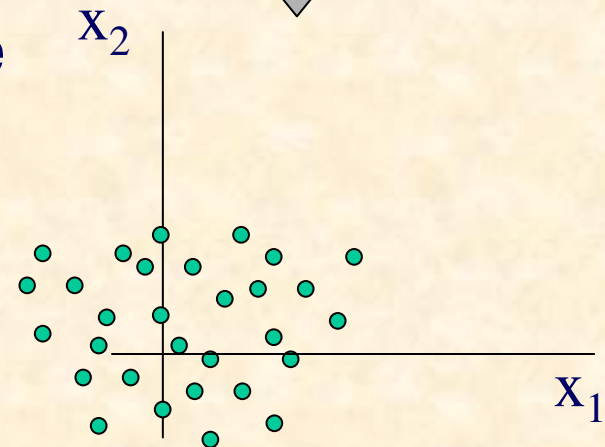
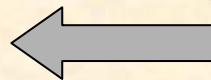
Usunięcie
średniej



dekorelacja



Wyrównanie
wariancji



Klasyfikatory statystyczne

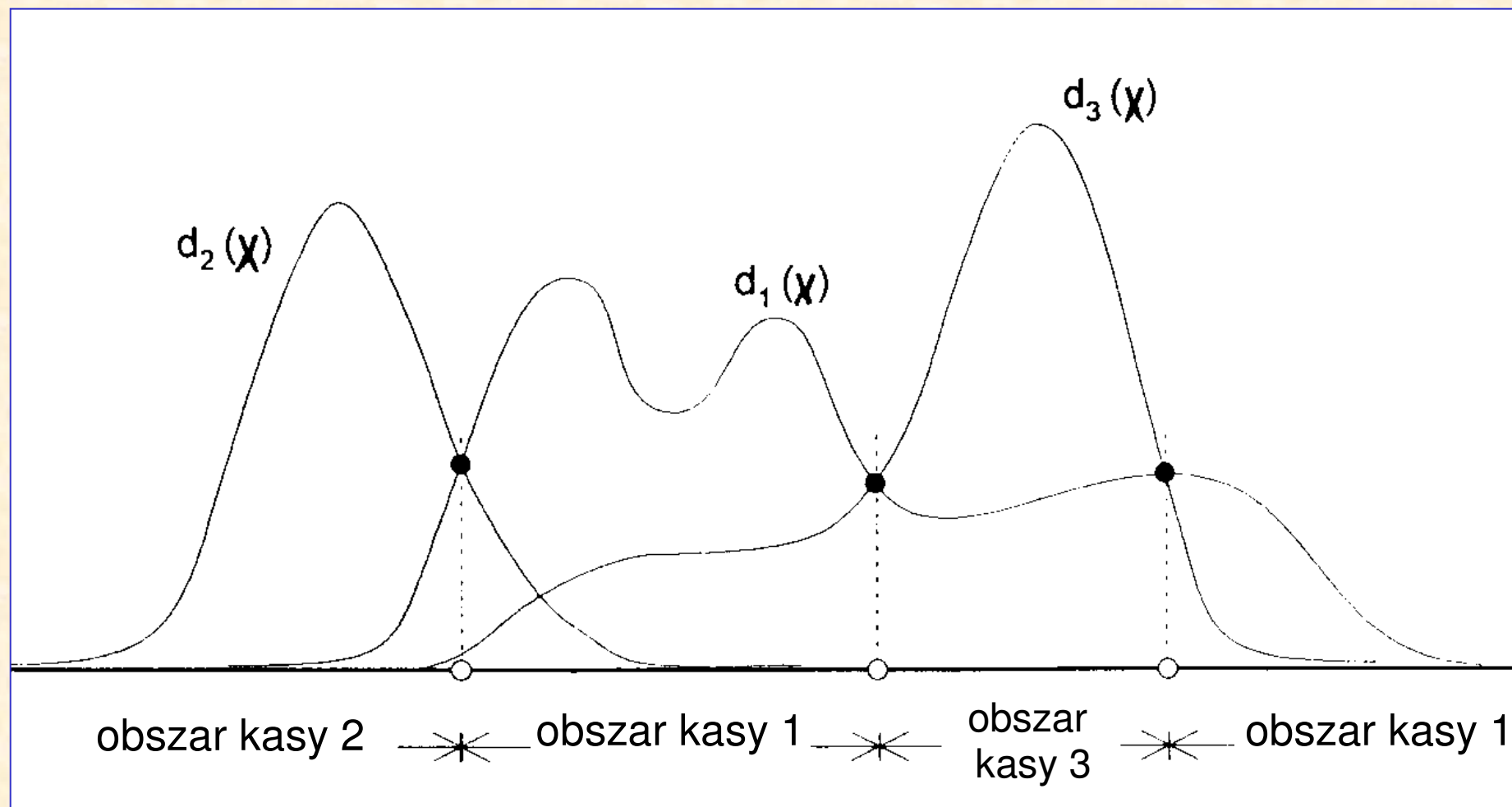
Jednym z możliwych podejść do zadania klasyfikacji jest budowanie tzw. *funkcji decyzyjnych*. **Liczba funkcji decyzyjnych jest równa liczbie klas wzorców.**

Niech $\mathbf{x}=[x_1, x_2, \dots, x_N]$ oznacza N-wymiarowy wektor cech. Spośród M klas $\omega_1, \omega_1, \dots, \omega_M$, poszukujemy M funkcji decyzyjnych $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_M(\mathbf{x})$ posiadających taką właściwość, że jeżeli wzorzec \mathbf{x} należy do klasy ω_i to zachodzi:

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \quad j = 1, 2, \dots, M; j \neq i$$

Funkcja decyzyjna $d_i(\mathbf{x})$, “wygrywa współzawodnictwo” do przyporządkowania wzorca \mathbf{x} klasie ω_i .

Funkcje decyzyjne



Funkcje decyzyjne

Granice decyzyjną pomiędzy dwiema dowolnymi klasami i oraz j ($i \neq j$) definiuje zależność:

$$d_{ij}(\mathbf{x}) = d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0$$

Zatem dla wzorców należących do i -tej klasy :

$$d_{ij}(\mathbf{x}) > 0$$

oraz dla wzorców należących do j -tej klasy:

$$d_{ij}(\mathbf{x}) < 0.$$

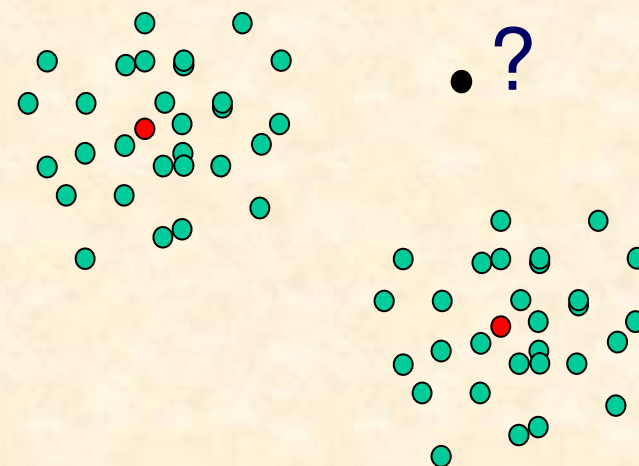
Klasyfikator minimalnoodległościowy (*ang. minimum distance classifier*)

Założmy, że każda klasa wzorów jest reprezentowana przez typowy wzorzec, zwany prototypem klasy:

$$m_i = \frac{1}{P_j} \sum_{x \in \omega_j} x_j, \quad j = 1, 2, \dots, M$$

gdzie P_j jest liczbą wzorców z klasy ω_j .

Jednym z możliwych sposobów klasyfikacji nowego wzorca x jest przydzielenie go do klasy, której prototyp znajduje się najbliżej tego wzorca.



Klasyfikator minimalnoodległościowy

Dla Euklidesowej miary odległości:

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \quad j = 1, 2, \dots, M$$

gdzie $\|\mathbf{z}\| = (\mathbf{z}^T \mathbf{z})^{1/2}$. Wzorzec \mathbf{x} jest przydzielony do klasy ω_j jeżeli $D(\mathbf{x}_j)$ jest odległością najmniejszą.

Klasyfikator minimalnoodległościowy

Na podstawie poprzednich zależności, funkcja decyzyjna jest postaci:

$$d_j(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j \quad j = 1, 2, \dots, M$$

Wektor \mathbf{x} jest przyporządkowywany do klasy ω_j jeżeli $d_j(\mathbf{x})$ jest największe.

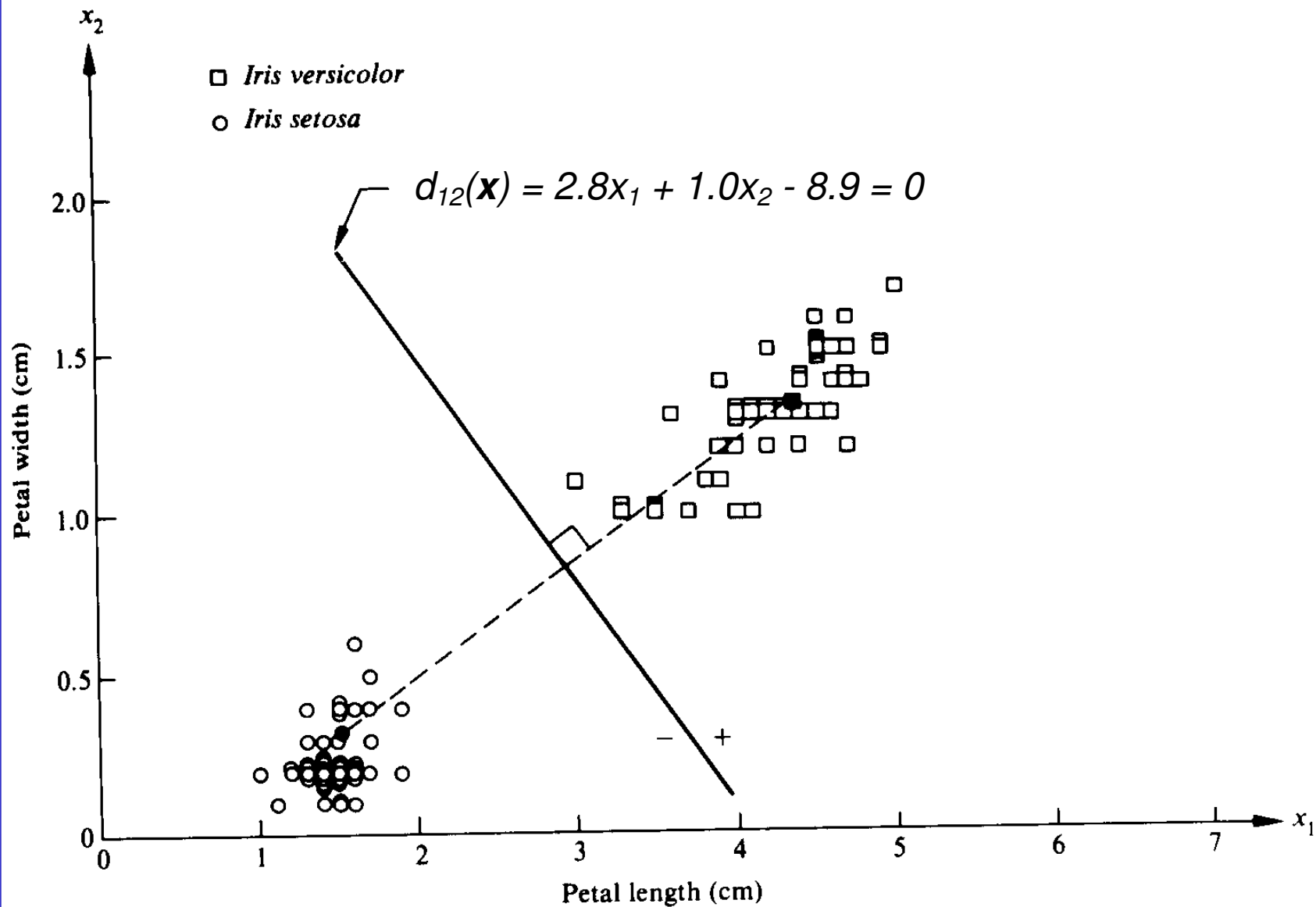
Klasyfikator minimalnoodległościowy

Granica decyzyjna pomiędzy klasami ω_i i ω_j dla klasyfikatora minimalnoodległościowego:

$$\begin{aligned}d_{ij}(\mathbf{x}) &= d_i(\mathbf{x}) - d_j(\mathbf{x}) = \\&= \mathbf{x}^T (\mathbf{m}_i - \mathbf{m}_j) - \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i - \mathbf{m}_j) = 0\end{aligned}$$

Obszar wyznaczony przez miejsca geometryczne spełniające to równanie dzieli na pół i jest prostopadły do linii łączącej \mathbf{m}_i i \mathbf{m}_j . Dla $N=2$ obszar ten redukuje się do linii prostej, dla $N=3$ jest płaszczyzną, oraz dla $N>3$ jest nazywany hiperpłaszczyzną.

Klasyfikator minimalnoodległościowy



Klasyfikator minimalnoodległościowy

Przykład 1: Rozpatrzmy dwie klasy ω_1 i ω_2 , posiadające wektory średnie cech $\mathbf{m}_1 = (4.3, 1.3)^T$ i $\mathbf{m}_2 = (1.5, 0.3)^T$. Funkcje decyzyjne dla tych klas są postaci:

$$d_1(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_1 - 0.5 \mathbf{m}_1^T \mathbf{m}_1 = 4.3x_1 + 1.3x_2 - 10.1$$

$$d_2(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_2 - 0.5 \mathbf{m}_2^T \mathbf{m}_2 = 1.5x_1 + 0.3x_2 - 1.17$$

Równanie granicy decyzyjnej:

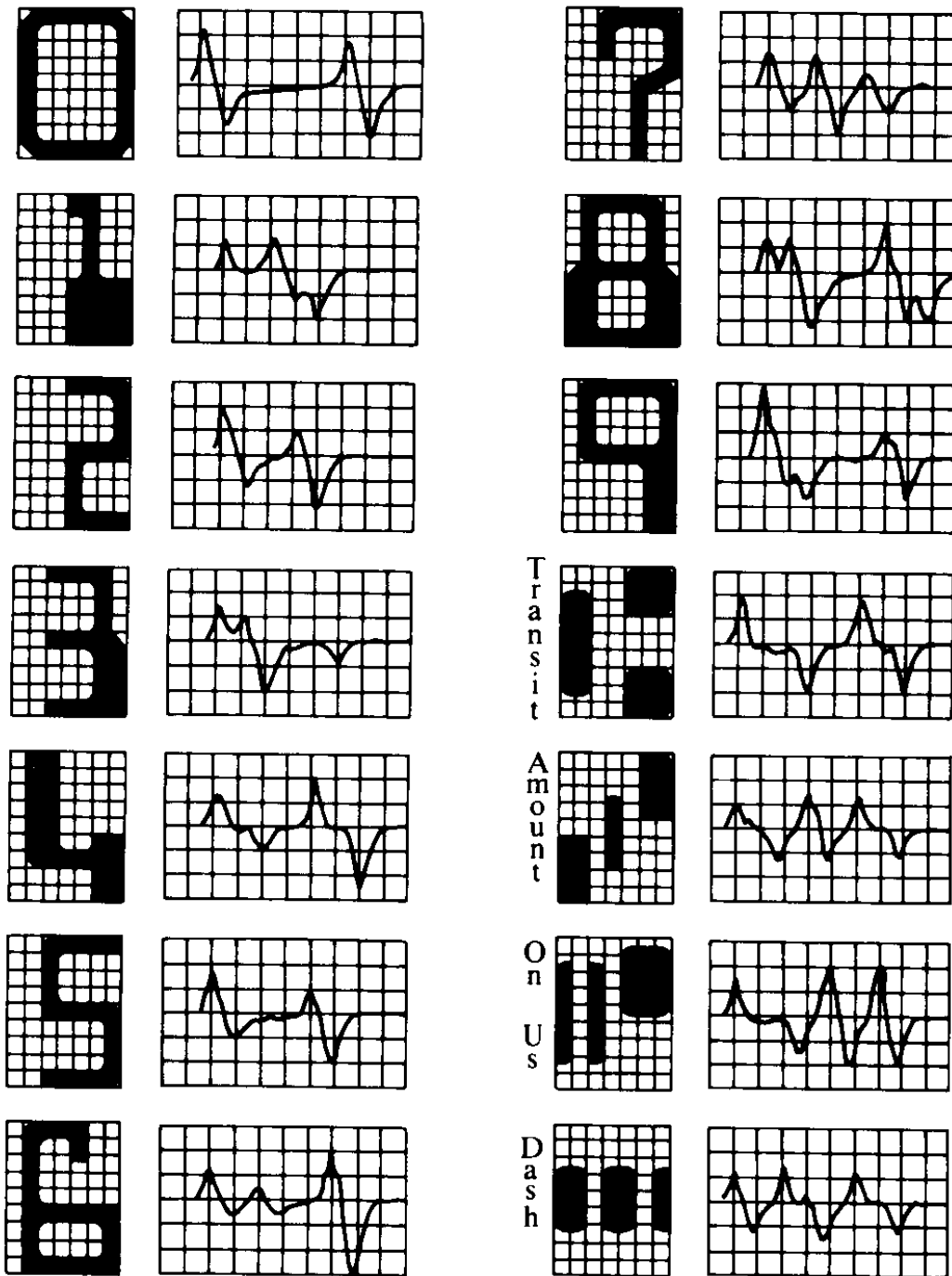
$$d_{12}(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x}) = 2.8x_1 + 1.0x_2 - 8.9 = 0$$

Nowy wektor cech jest klasyfikowany zależnie od znaku $d_{12}(\mathbf{x})$.

Przykład 2: Klasyfikacja znaków do numeracji czeków

Do klasyfikacji tych znaków zastosowano klasyfikator minimalnoodległościowy.

Próbki sygnałów odpowiadające każdemu ze znaków, tworzą zbiór 14 wektorów prototypowych. Klasyfikacja znaków polega na znalezieniu prototypu klasy najbliższego wektorowi cech utworzonego z rozpoznawanego znaku.



Klasyfikator minimalnoodległościowy

Klasyfikator minimalnoodległościowy sprawdza się w zastosowaniach, w których **odległość pomiędzy prototypami klas jest duża** w porównaniu z rozrzutem wzorców w klasie. Warunek ten jest jednak rzadko spełniany w praktyce chyba, że projektant systemu klasyfikacji zapewni spełnienie tego warunku.

Zauważmy również, że klasyfikator minimalnoodległościowy **wytwarza granice decyzyjne w postaci prostych i (hiper)-płaszczyzn**, a nie wszystkie problemy klasyfikacji należą do rozdzielných liniowo.

Rodzaje klasyfikatorów

Parametryczne: warunkowe gęstości prawdopodobieństwa cech są znane jako funkcje: nieznane są parametry tych funkcji (np. dla rozkładu Gaussowskiego μ i σ), np. klasyfikator Bayesa jest klasyfikatorem parametrycznym

Nieparametryczne: warunkowe gęstości prawdopodobieństwa są nieznane i muszą być estymowane na podstawie dużej liczby danych pomiarowych.

Uczenie klasyfikatorów

Uczenie (trening klasyfikatora): estymacja funkcji gęstości prawdopodobieństwa lub ich parametrów na podstawie pomiarów.

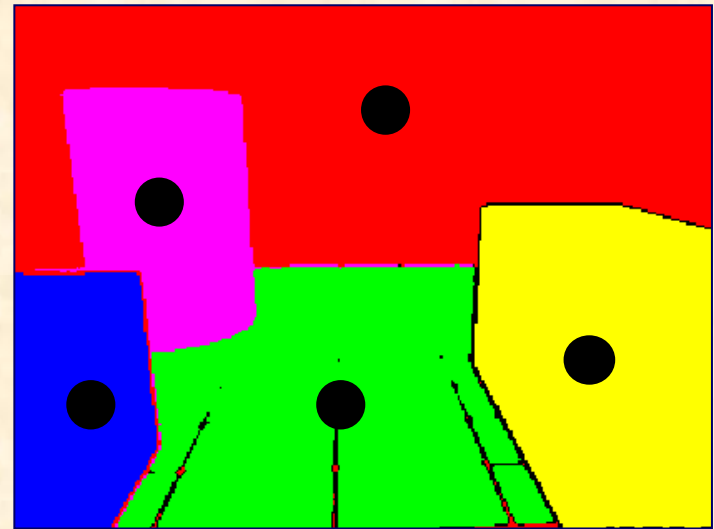
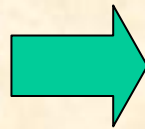
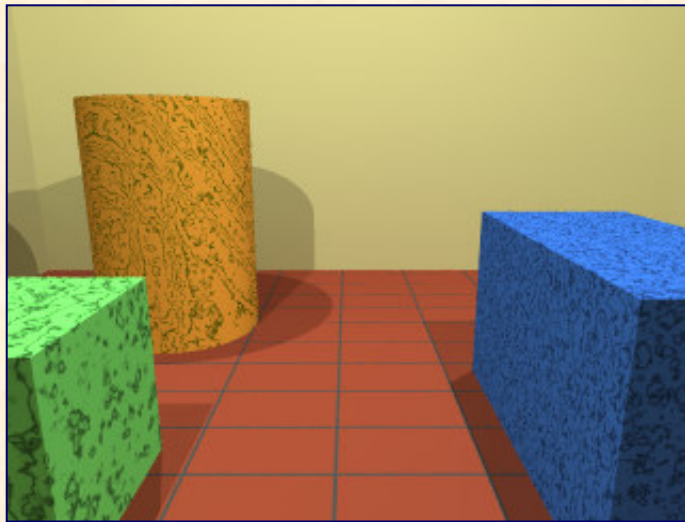
Uczenie nadzorowane: wektory cech zostały uprzednio sklasyfikowane za pomocą niezależnej, bezbłędnej metody.

Uczenie nienadzorowane: nieznane są klasy, do których należą zmierzone wektory cech. Klasy oraz ich liczba są określane przez grupowanie wzorców w przestrzeni cech (klasteryzacja).

Przykłady klasyfikatorów:

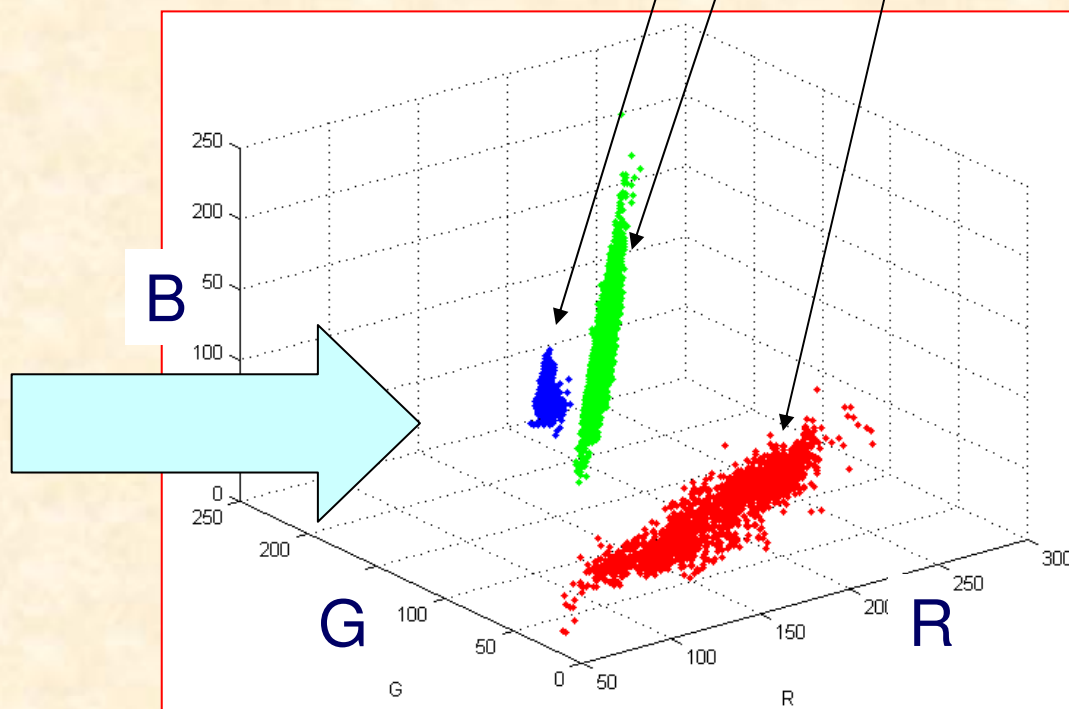
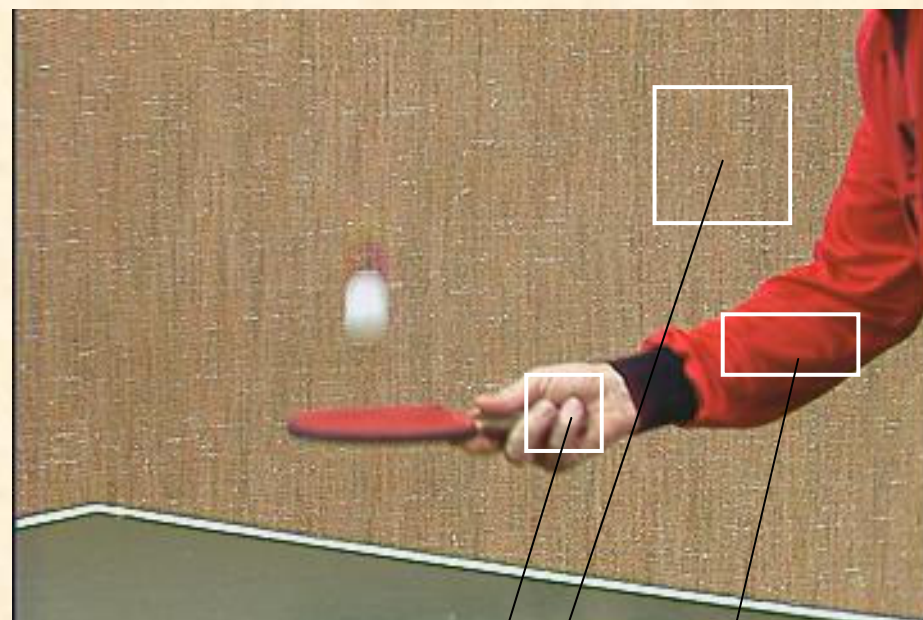
**Klasyfikatory k-NN,
Sztuczne sieci neuronowe**

Segmentacja obrazu – przykład

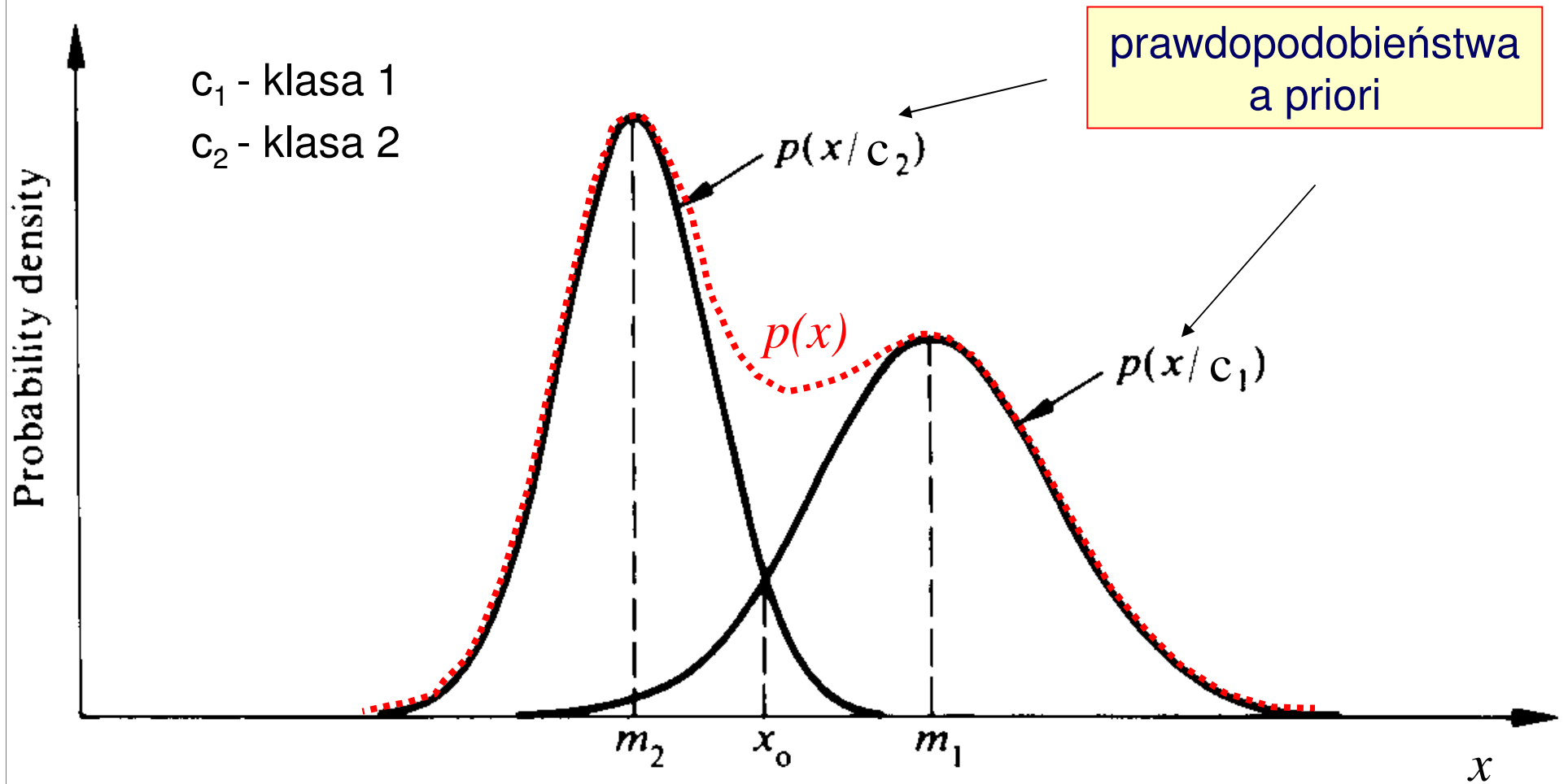


Segmentacja na podstawie składowych koloru $\{R,G,B\}$

Obraz trójwymiarowych
rozkładów (Gausa) dla
składowych koloru R,G,B



Klasyfikator Bayesa dla rozkładów Gaussa



Reguła Bayesa

Z podstaw rachunku prawdopodobieństwa wiadomo, że:

$$p(a/b) = \frac{p(a)p(b/a)}{p(b)}$$

zatem:

$$p(c_i / \mathbf{x}) = \frac{P(c_i)p(\mathbf{x}' / c_i)}{p(\mathbf{x})}$$

**prawdopodobieństwo
a priori**

**prawdopodobieństwo
a posteriori**

**reguła
Bayesa**

Estymator największej wiarygodności

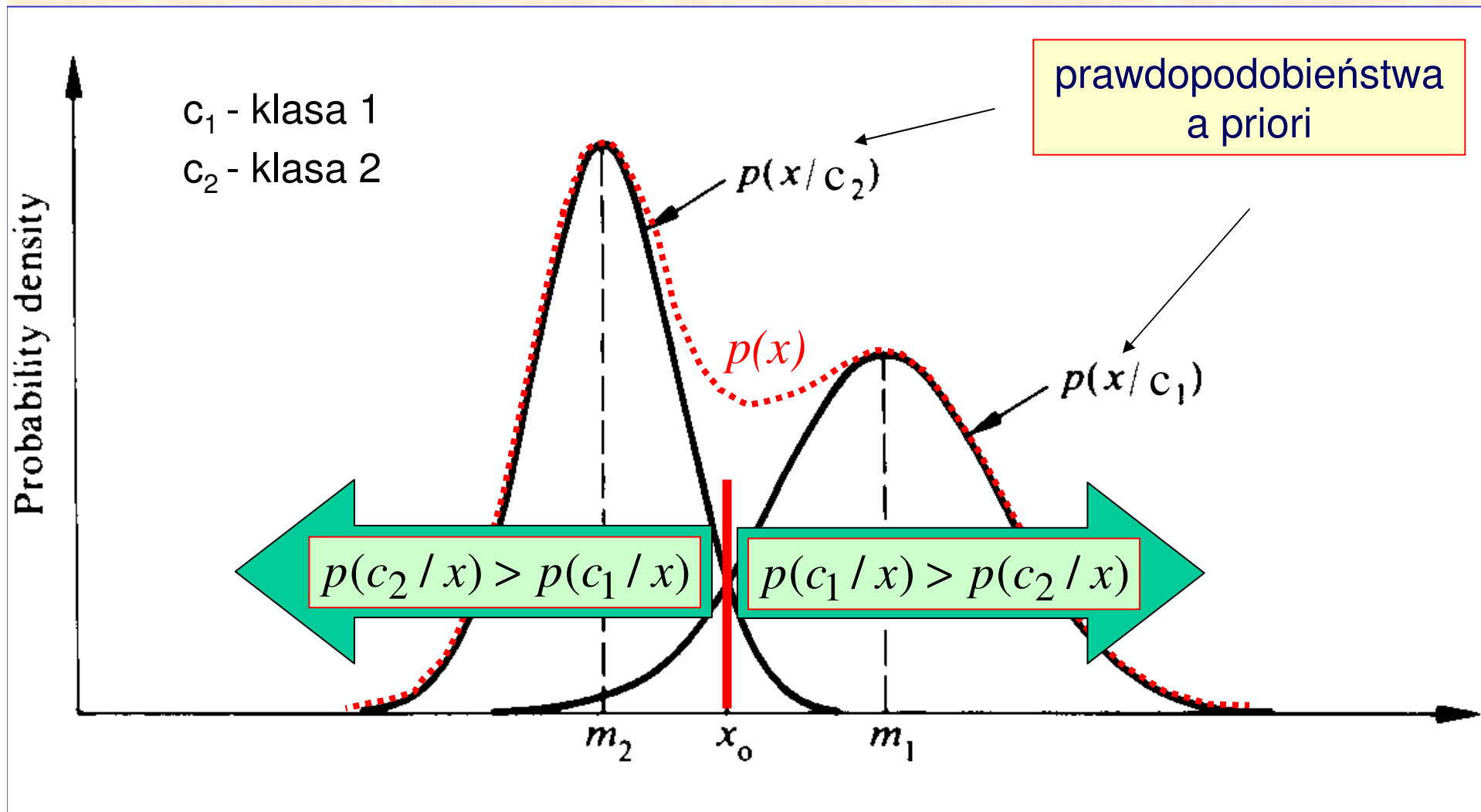
Punkt obrazu przydzielamy do tej i -tej klasy, dla której zachodzi:

$$L_i = \arg \max_i \left\{ p(c_j / \mathbf{x}) = \frac{P(c_i)p(\mathbf{x} / c_i)}{p(\mathbf{x})} \right\}, \quad i = 1, 2, \dots, N$$

tj.

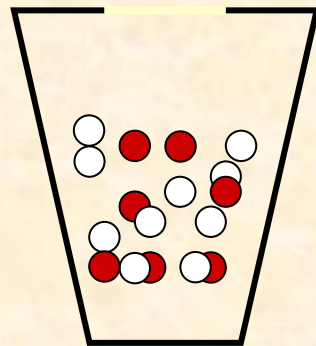
$$L_i > L_j, \quad j \neq i, \quad i = 1, 2, \dots, N$$

Klasyfikator Bayesa dla rozkładów Gaussa



Klasyfikator Bayesa - przykład

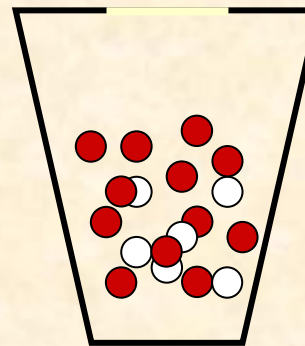
Urna A



20 – czerwone

20 – białe

Urna B



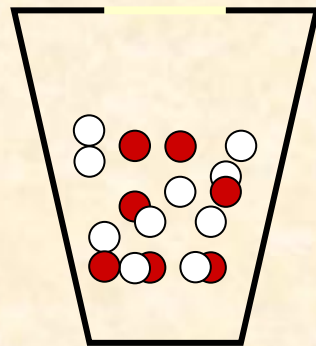
30 – czerwone

10 – białe

Losowo „w ciemno” wybieramy urnę a potem żeton w urnie.
Pytanie: Z której urny wybraliśmy żeton?

Klasyfikator Bayesa - przykład

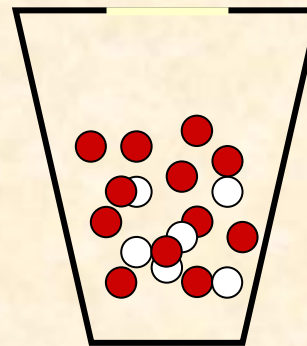
Urna A



20 – czerwone

20 – białe

Urna B



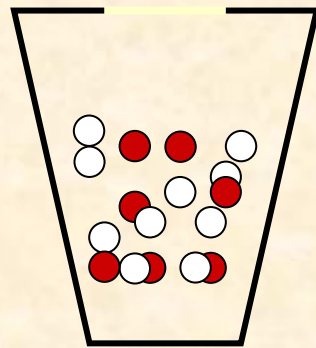
30 – czerwone

10 – białe

Zanim stwierdzimy jaki kolor żetonu wylosowaliśmy mamy podstawy twierdzić (wiedza a priori), że prawdopodobieństwo wybrania każdej z urn wynosi 0.5.

Klasyfikator Bayesa - przykład

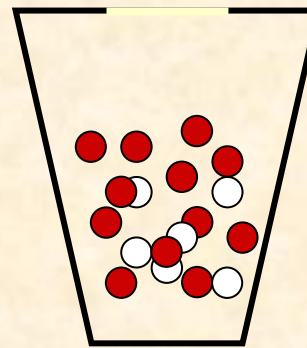
Urna A



20 – czerwone

20 – białe

Urna B



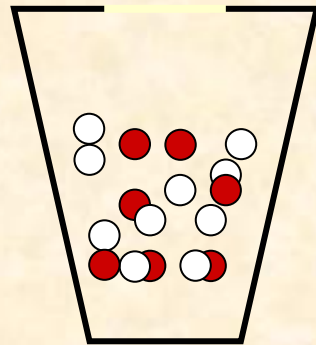
30 – czerwone

10 – białe

Przypuśćmy, że wylosowaliśmy czerwony żeton. Czy możemy teraz zweryfikować naszą hipotezę co do wyboru urny na podstawie nowej wiedzy a posteriori? Tak, odpowiedź mamy na podstawie tw. Bayesa.

Klasyfikator Bayesa - przykład

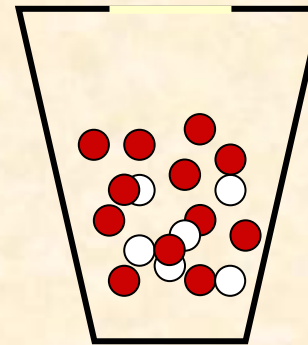
Urna A



20 – czerwone



20 – białe

Urna B



30 – czerwone

10 – białe

 To ta
urna!
najprawdopodobniej


$$p(A / \text{red}) = \frac{P(A)p(\text{red} / A)}{P(A)p(\text{red} / A) + P(B)p(\text{red} / B)} = \frac{0.5 \times 0.5}{0.5 \times 0.5 + 0.5 \times 0.75} = 0.4$$

$$p(B / \text{red}) = \frac{P(B)p(\text{red} / B)}{P(A)p(\text{red} / A) + P(B)p(\text{red} / B)} = \frac{0.5 \times 0.75}{0.5 \times 0.5 + 0.5 \times 0.75} = 0.6$$

Klasyfikator Bayesa dla rozkładów Gaussa

Dla rozkładu jednowymiarowego Gaussa
prawdopodobieństwo a priori:

$$p(x / c_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(x - m_i)^2}{2\sigma_i^2}\right] \quad i = 1, 2$$

Prawdopodobieństwo a posteriori:

$$p(c_i / x) = p(x / c_i)P(c_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(x - m_i)^2}{2\sigma_i^2}\right] P(c_i) \quad i = 1, 2$$

Rozkłady wielowymiarowe Gaussa

Dla rozkładu 2-wymiarowego $\mathbf{x}^T = [x_1 \ x_2]$ Gaussa prawdopodobieństwo a priori:

$$p(\mathbf{x} / c_i) = \frac{1}{(2\pi)^{1/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{(\mathbf{x} - \mathbf{m}_i)^T \Sigma_i^{-1} (\mathbf{x} - \mathbf{m}_i)}{2} \right] \quad i = 1, 2, \dots, N$$

Gdzie i jest numerem klas a N jest ich liczbą.
oraz macierz kowariancji:

$$\Sigma_i = \begin{bmatrix} \sigma_{x1}^2 & \sigma_{x1}\sigma_{x2} \\ \sigma_{x2}\sigma_{x1} & \sigma_{x2}^2 \end{bmatrix}_i$$

Rozkład dwuwymiarowy Gaussa - przykład

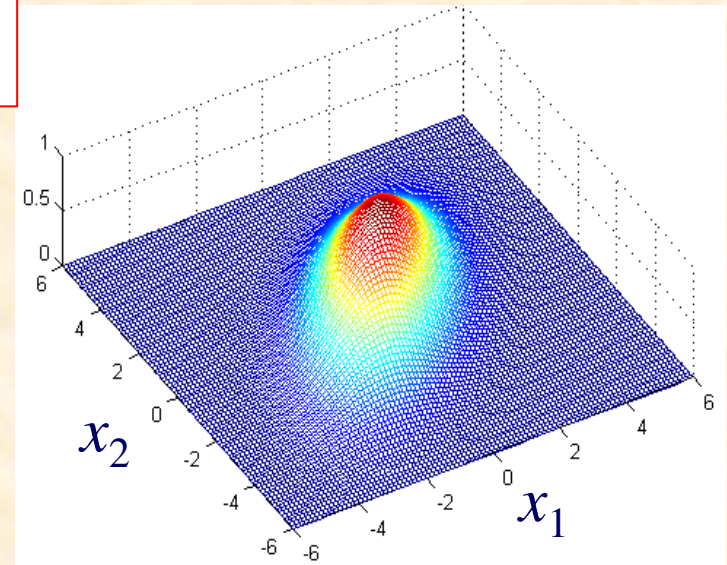
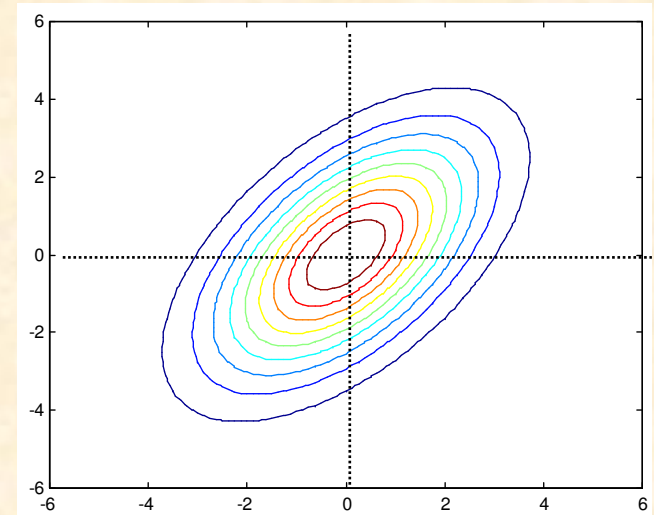
Niech dla $m_1=m_2=0$:

$$\Sigma = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \rightarrow \Sigma^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{8} \end{bmatrix} \rightarrow |\Sigma| = 8$$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{1/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} \left(\mathbf{x}^T \Sigma^{-1} \mathbf{x} \right) \right] = A \exp(B)$$

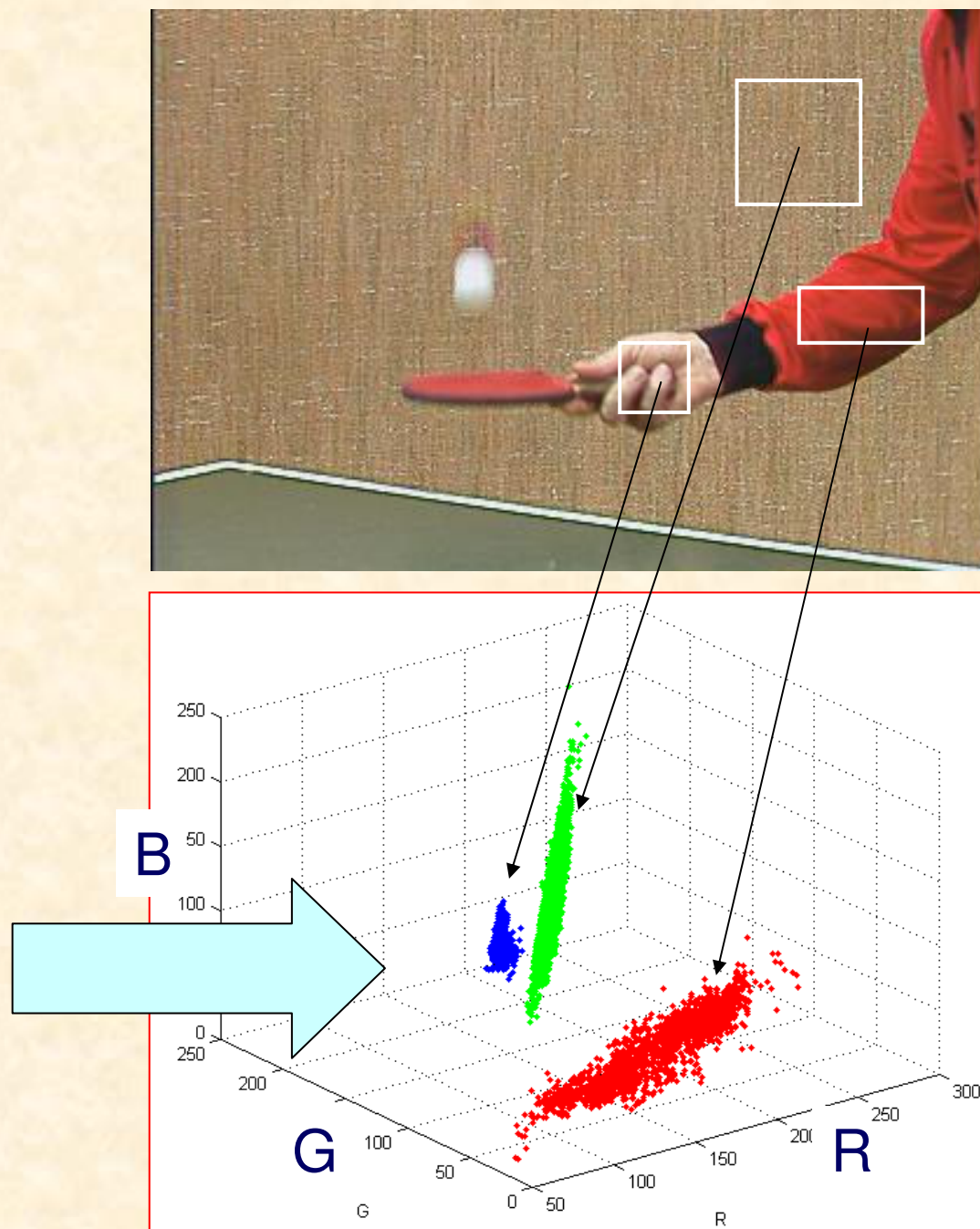
$$A = \frac{1}{(2\pi)^{1/2} 8^{1/2}} = \frac{1}{4\sqrt{\pi}}$$

$$B = -\frac{1}{2} \left(\frac{1}{2} x_1^2 - \frac{1}{2} x_1 x_2 + \frac{3}{8} x_2^2 \right)$$



Segmentacja na podstawie składowych koloru $\{R,G,B\}$

Obraz trójwymiarowych
rozkładów dla składowych
koloru R,G,B; można je
opisać wielowymiarowymi
rozkładami Gaussa



Wybór cech obrazu



Wektor cech:

$$\theta = \{x, y, R, G, B, u, v\}$$

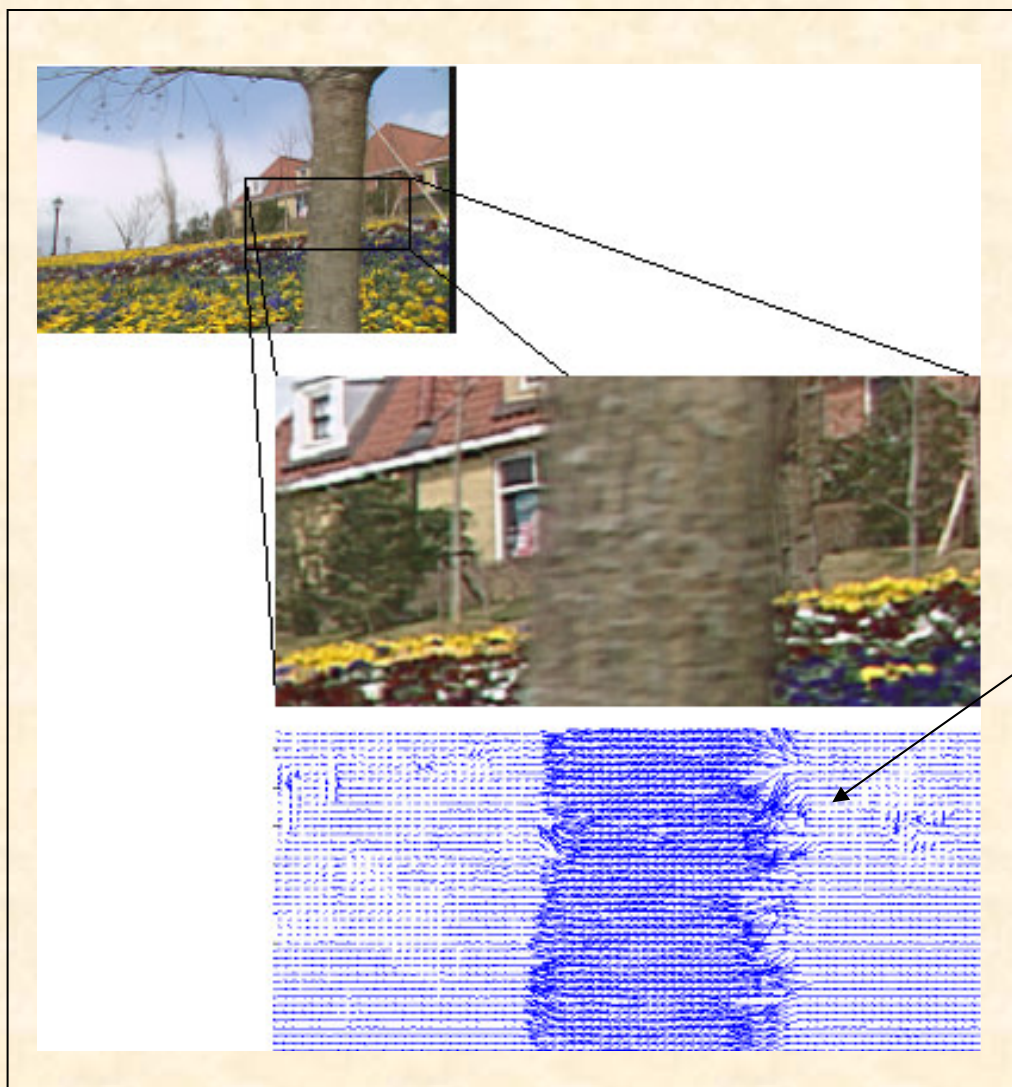
współrzędne

kolor

ruch

$$u = \frac{dx}{dt}$$
$$v = \frac{dy}{dt}$$

Segmentacja na podstawie ruchu

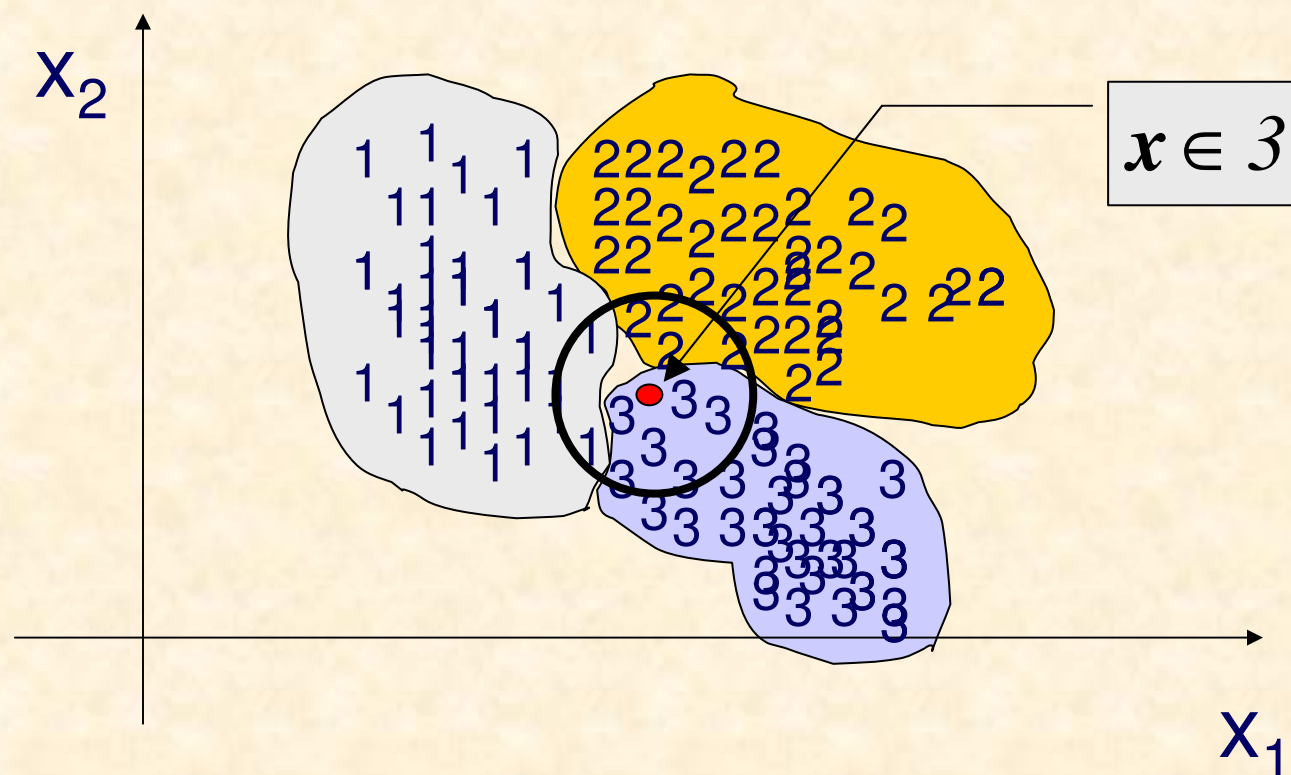


Segmentacja
na podstawie
składowych
ruchu

$\{u, v\}$

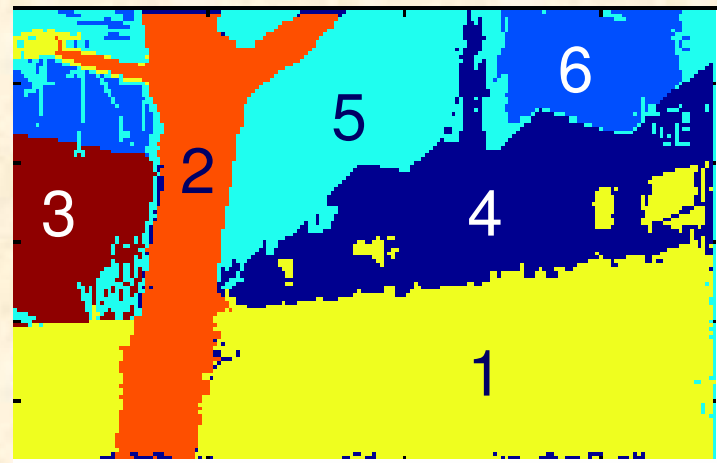
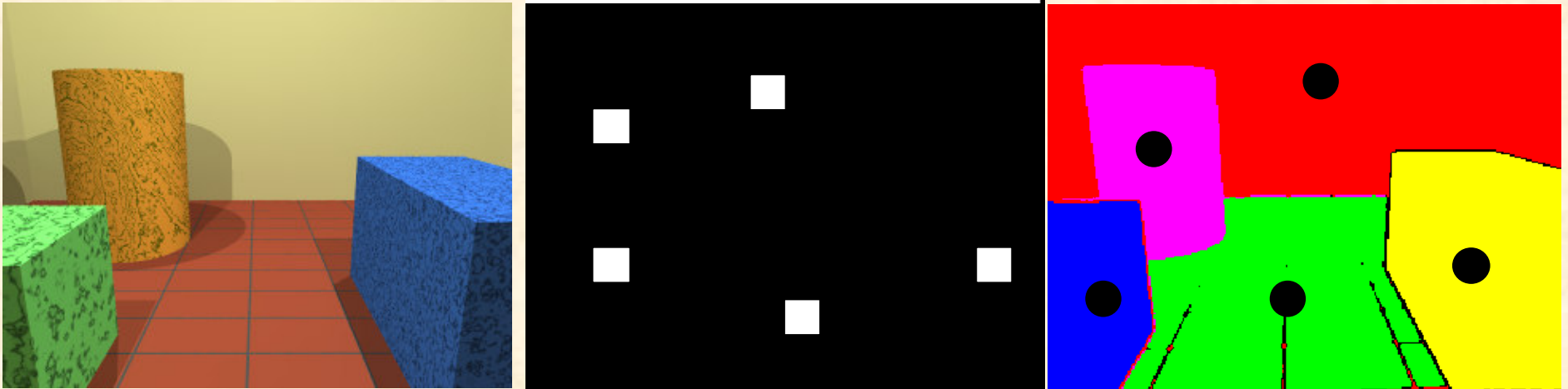
Pole wektorowe
przepływu
optycznego
(ang. *optical flow*)

Algorytm k -najbliższych sąsiadów

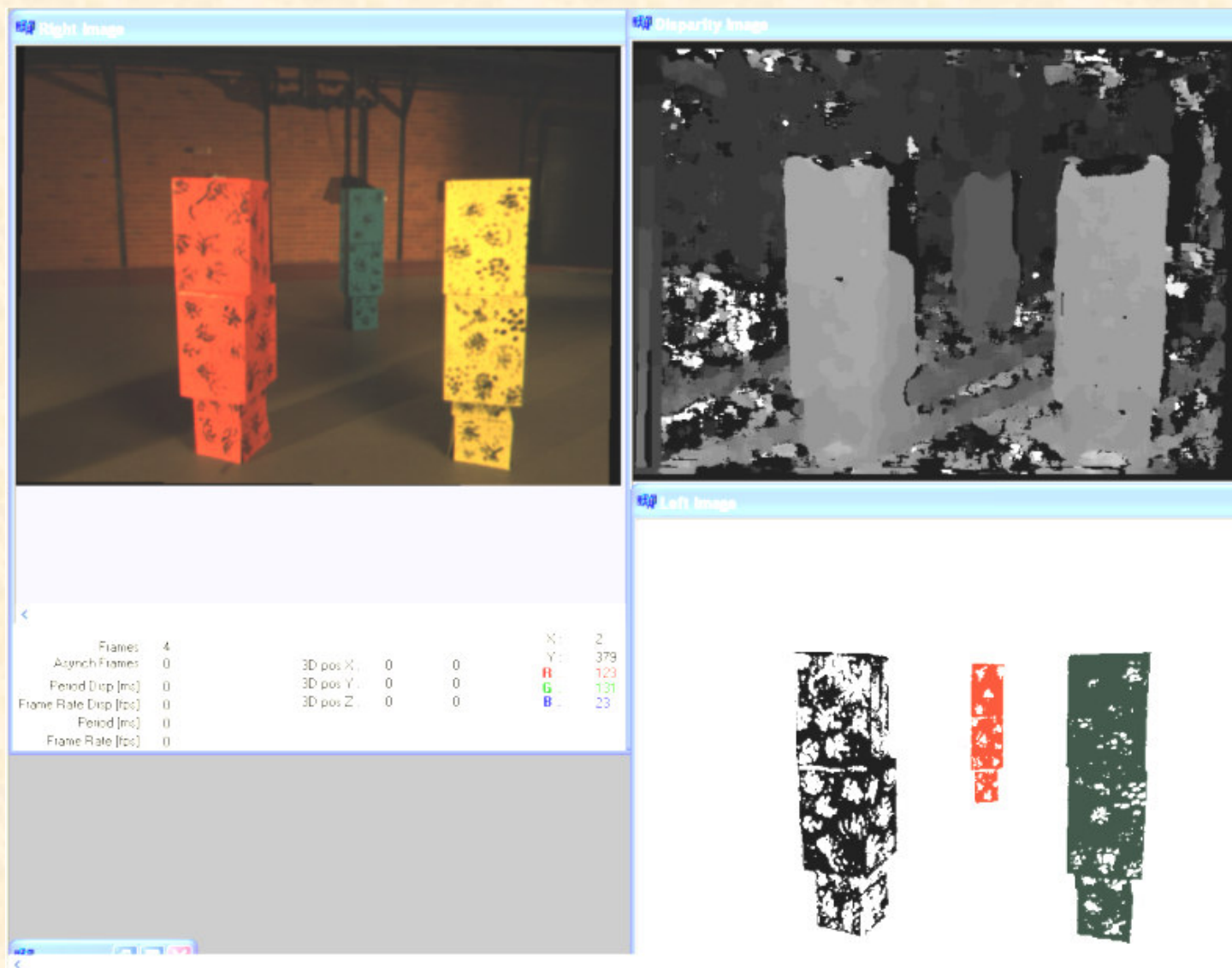


Dla każdego punktu przestrzeni cech jest badanych k -najbliższych sąsiadów. Punkt jest przydzielany do tej klasy do której należy największa liczba wzorców spośród k -sąsiadów (funkcja decyzyjna?).

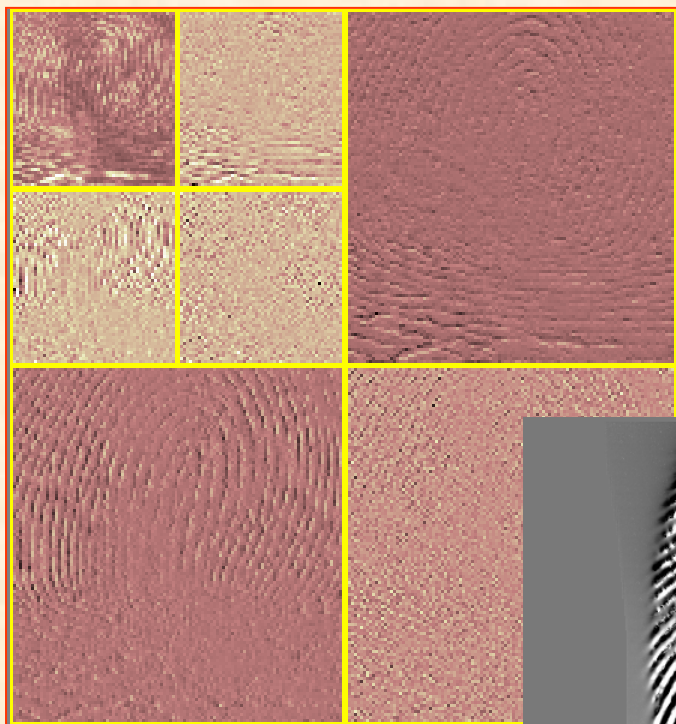
Wyniki segmentacji - przykłady



Wyniki segmentacji - przykłady



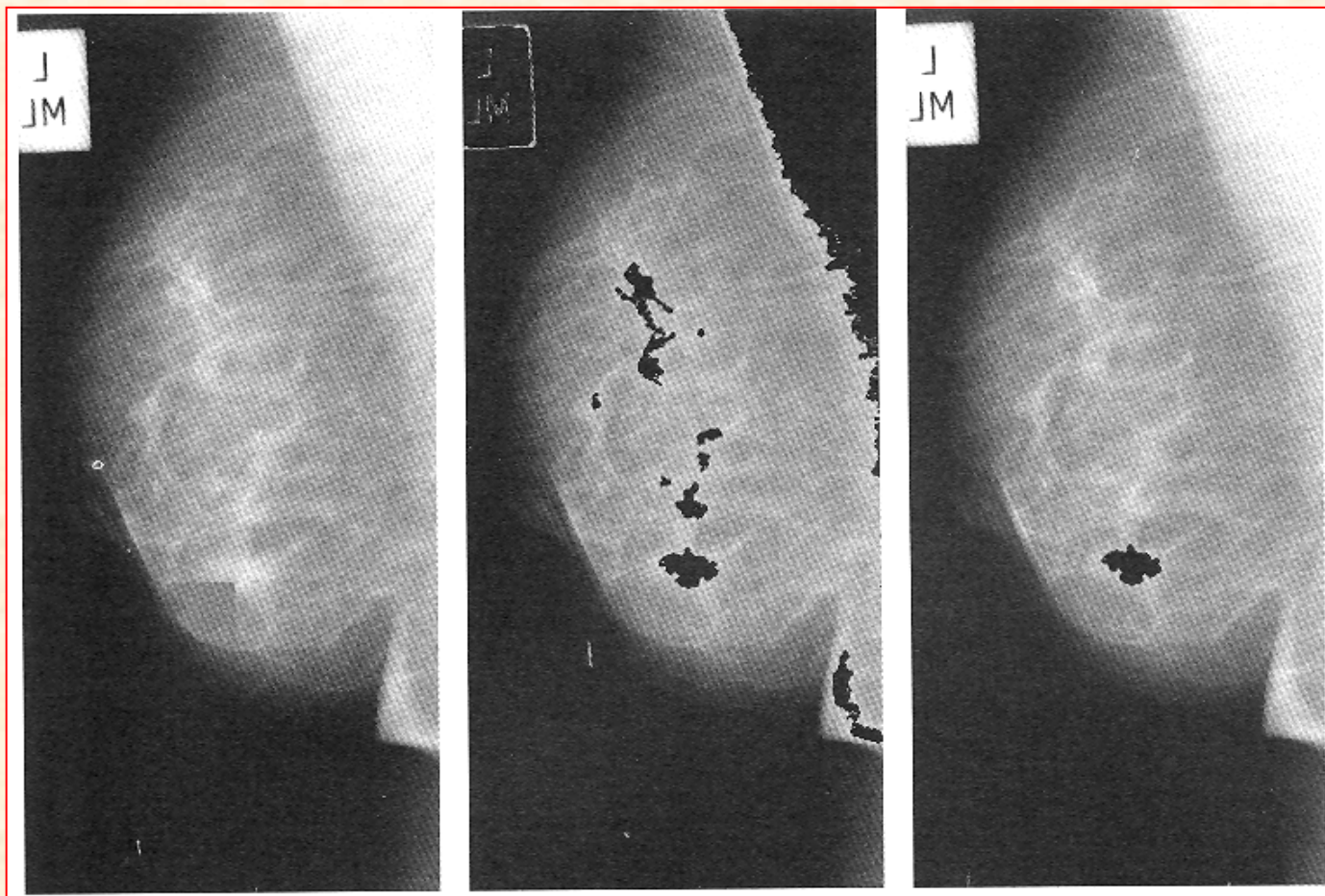
Przykłady zadań rozpoznawania obrazów: Odciski palców



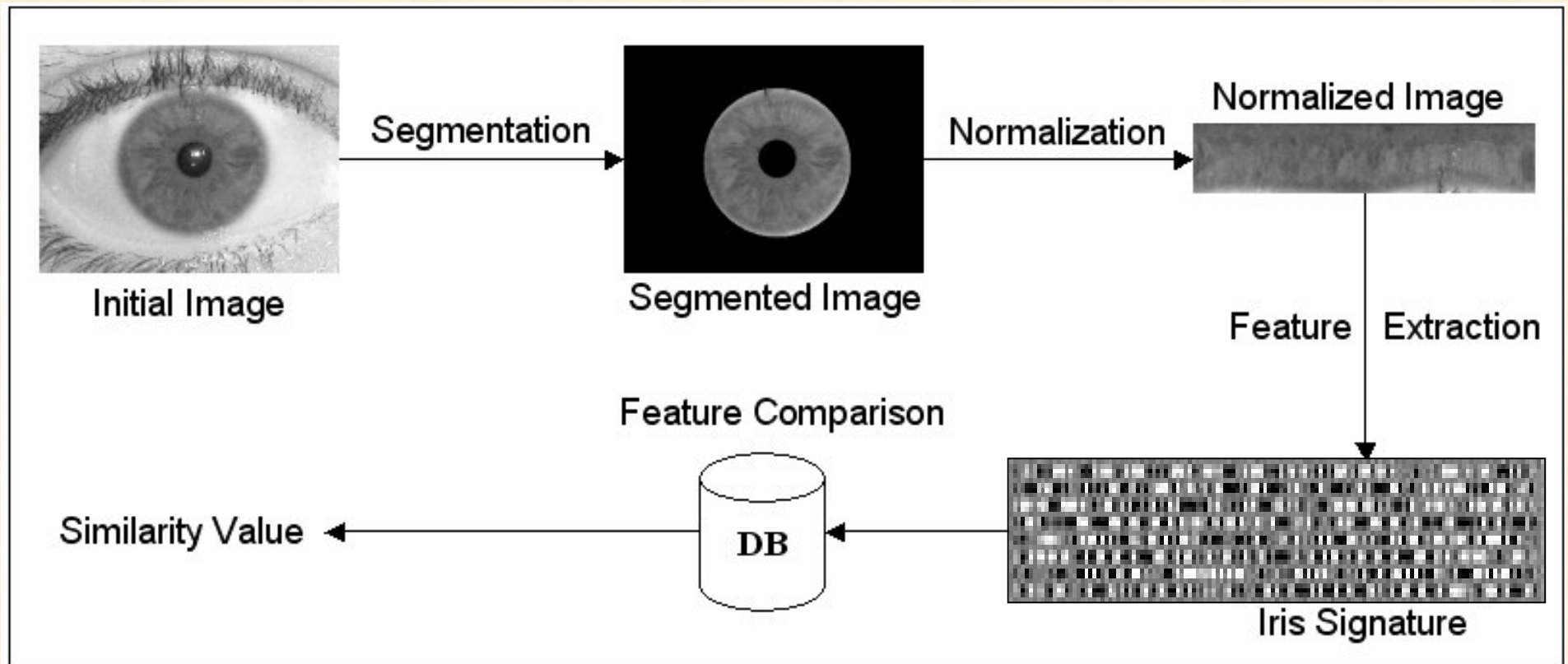
Baza obrazów z
odciskami palców
FBI 1992



Przykłady zadań rozpoznawania obrazów: Diagnostyka obrazowa



Przykłady zadań rozpoznawania obrazów: rozpoznawanie tęczówki oka (biometria)



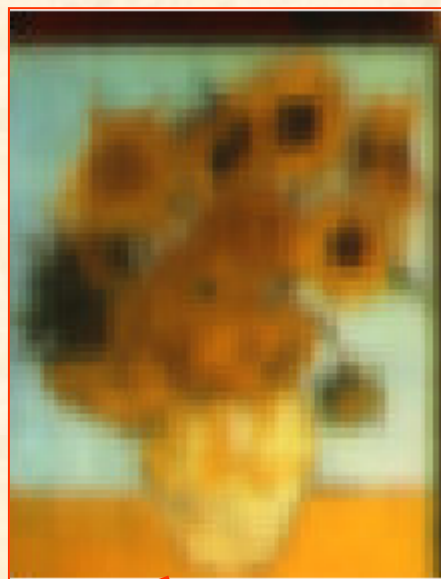
Patentowany algorytm Daugmana

Przykłady zadań rozpoznawania obrazów: Bazy obrazów

„Idea” obrazu
poszukiwanego lub



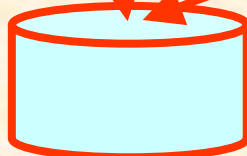
kopia obrazu
poszukiwanego



Obraz
odszukany

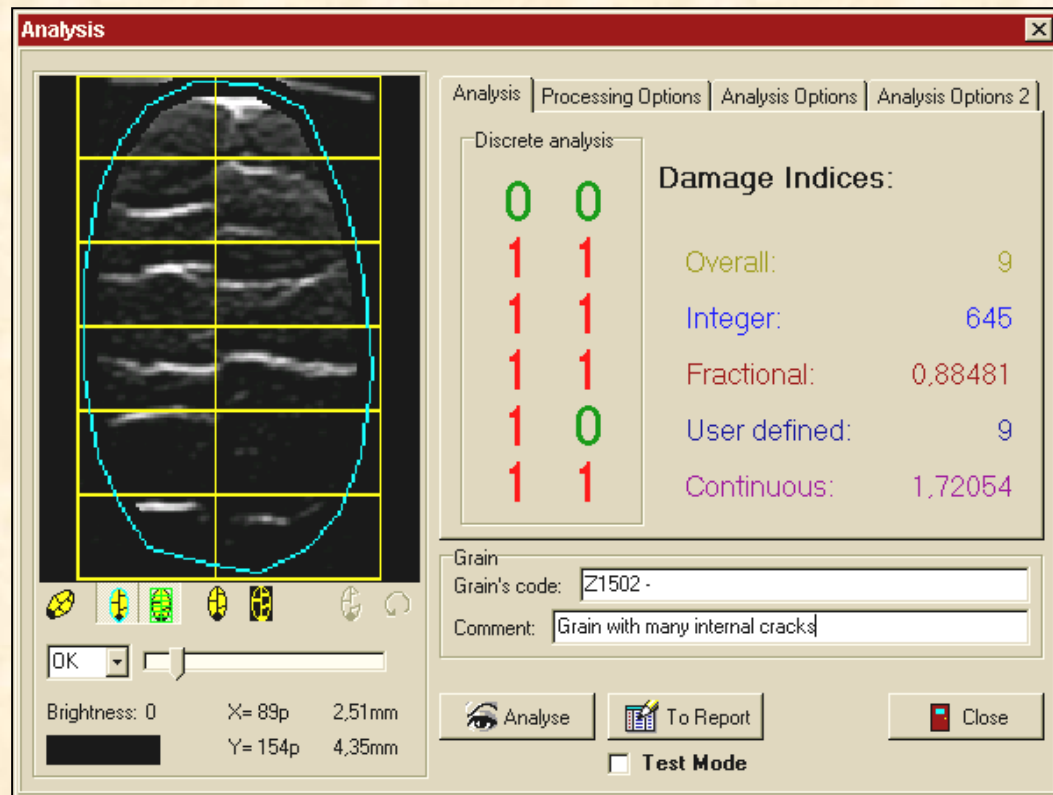
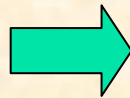


DWT



*C.E. Jacobs, A. Finkelstein, D.H. Salesis,
„Fast multiresolution image quering”, 1999*

Analiza obrazów RTG ziaren pszenicy



P. Strumillo, J. Niewczas, P. Szczypiński, P. Makowski, W. Wozniak,
“Computer system for analysis of X-ray images of wheat grains”,
International Agrophysics, 1999, vol. 13, No. 1, pp. 133–140.

Automatyczna detekcja tekstu w scenie

