Zaawansowane algorytmy wizyjne
AGH, Katedra Automatyki i Robotyki

Rozpoznawanie kształtów

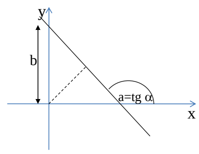
Transformata Hough'a, uogólniona transformata Hough'a, metody metryczne dopasowania kształtów, rozpoznawanie obiektów złożonych

Konspekt stanowi uzupełnienie wykładu,
nie pokrywa całości materiału przedstawionego na wykładzie
i obowiązującego do zaliczenia

Transformata Hougha

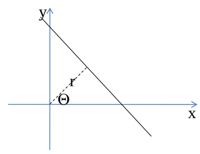
Opis prostej przez parametry (a, b)

y = a x + b



Opis prostej przez parametry (Θ, r)

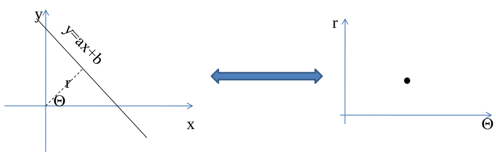
x cos Θ + y sin Θ = r



(a,b) ⟷ (Θ, r): *a = -cos Θ / sin Θ; b = r/sin Θ*

Transformata Hougha

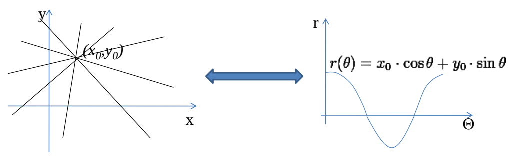
(x, y) – przestrzeń obrazu
(Θ, r) – przestrzeń parametrów prostej



prosta w przestrzeni (x,y) → punkt w przestrzeni (Θ, r)

Transformata Hougha

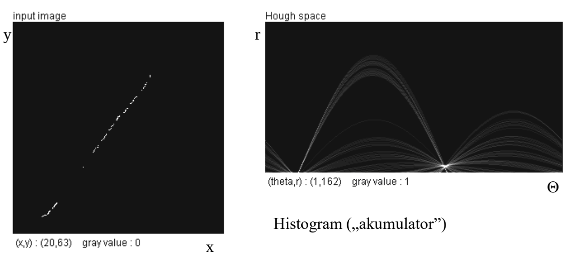
(x, y) – przestrzeń obrazu
(Θ, r) – przestrzeń parametrów prostej



prosta w przestrzeni (x,y) → punkt w przestrzeni (Θ, r)

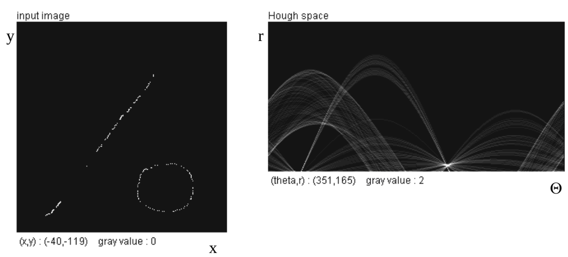
punkt w przestrzeni (x,y) → sinusoida w przestrzeni (Θ, r) odpowiadająca *pękowi* prostych

Transformata Hougha

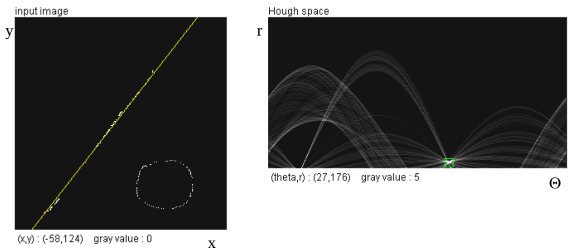


Histogram („akumulator”)

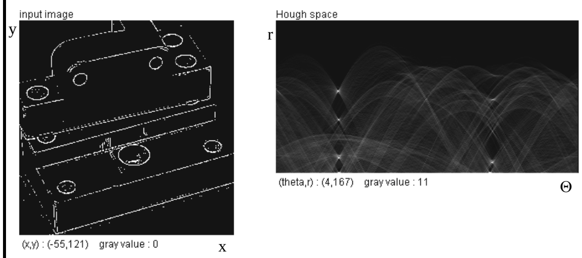
Transformata Hougha



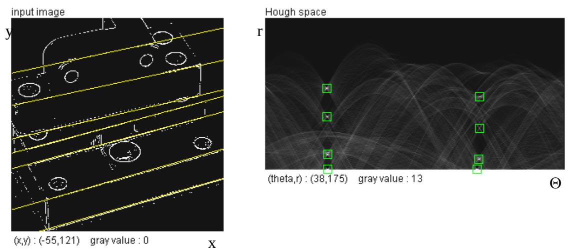
Transformata Hougha



Transformata Hougha

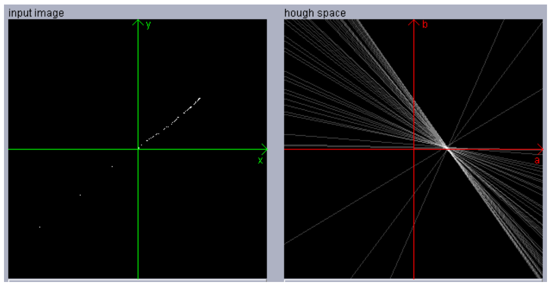


Transformata Hougha



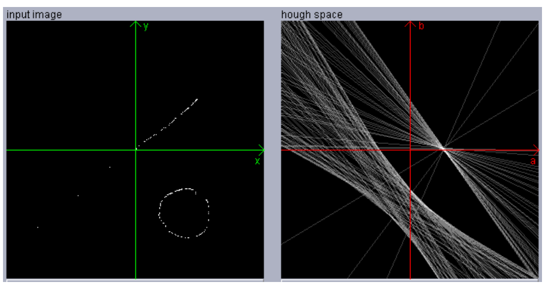
Transformata Hougha

Wersja dla parametrów prostej: współczynnik kierunkowy + wyraz wolny



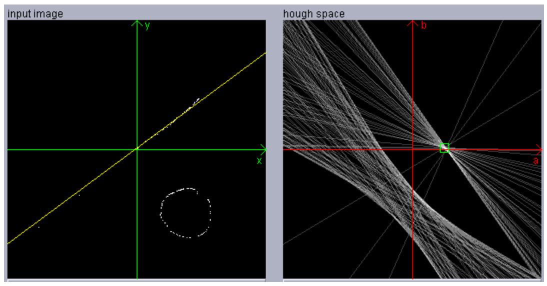
Transformata Hougha

Wersja dla parametrów prostej: współczynnik kierunkowy + wyraz wolny



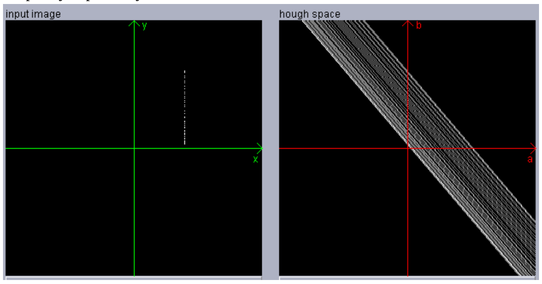
Transformata Hougha

Wersja dla parametrów prostej: współczynnik kierunkowy + wyraz wolny



Transformata Hougha

Wersja dla parametrów prostej: współczynnik kierunkowy + wyraz wolny
W tej wersji transformaty nie można uzyskać współczynników dla prostych pionowych

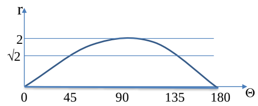


Transformata Hougha – przykład

Przykłady

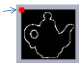
1. Narysuj transformatę Hough'a w przestrzeni (Θ,r) dla $\Theta \in [0\ldots180°]$ dla obrazu złożonego z dwóch pikseli (x,y)=(0,0) oraz (x,y)=(0, 2).

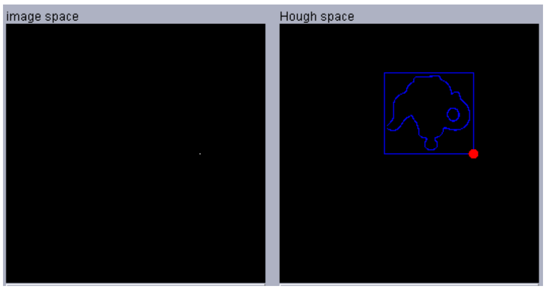
Wartości powinny być opisane co 45°.



Uogólniona transformata Hougha (transf. Ballarda)

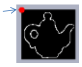
Punkt odniesienia

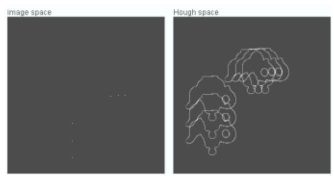
Obraz odniesienia: 



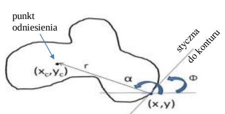
Uogólniona transformata Hougha

Punkt odniesienia

Obraz odniesienia: 



Uogólniona transformata Hougha



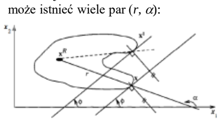
Utworzenie R-table:

Dla każdego punktu (x,y) należącego do konturu należy obliczyć Φ oraz r i α

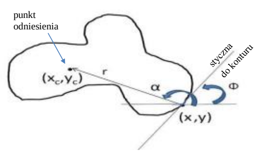
Pary (r,α) są zapamiętywane w tabeli odniesienia (reference table, R-table):

i	ϕ_i	R_{ϕ_i}
1	0	$(r_{11}, \alpha_{11})\ (r_{12}, \alpha_{12}) \ldots (r_{1n}, \alpha_{1n})$
2	$\Delta\phi$	$(r_{21}, \alpha_{21})\ (r_{22}, \alpha_{22}) \ldots (r_{2m}, \alpha_{2m})$
3	$2\Delta\phi$	$(r_{31}, \alpha_{31})\ (r_{32}, \alpha_{32}) \ldots (r_{3k}, \alpha_{3k})$
...	...	...

Dla danej wartości Φ może istnieć wiele par (r, α):



Uogólniona transformata Hougha



Lokalizacja obiektu na obrazie:

Dla każdego punktu obrazu (x, y) należy obliczyć Φ, który wskaże odpowiedni wiersz w R-table.

Należy dodać do akumulatora wszystkie punkty wskazane w tym wierszu przez pary (r, α):

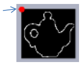
$x_c = x + r\cos(\alpha)$,
$y_c = y + r\sin(\alpha)$.

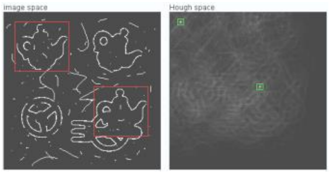
Maksima akumulatora wskazują lokalizacje obiektu

Uogólniona transformata Hougha

Punkt odniesienia

Obraz odniesienia:





Rozpoznawanie kształtów przez dopasowanie

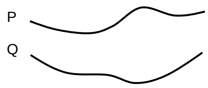
Definicja metryki

Funkcja d: $X \times X \to R^+$ jest metryką, jeżeli spełnia warunki:

- $d(x, y) = 0 \Leftrightarrow x = y$
- $d(x, y) = d(y, x)$ - warunek symetrii
- $d(x, y) + d(y, z) \geq d(x, z)$ – warunek trójkąta

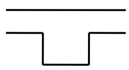
Przykłady metryki:
- Euklidesowa
- Uliczna (taksówkowa, Manhattan)
- Metryka maksimum (Czebyszewa)

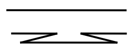
Odległość Frecheta



Odległość Frecheta – minimum po wszystkich monotonicznie rosnących parametryzacjach z maksymalnej odległości między $P(\alpha(t))$ i $Q(\beta(t))$

$$\delta_F(P, Q) = \min_{\substack{\alpha[0,1]\to[0,N] \\ \beta[0,1]\to[0,M]}} \{\max_{t\in[0,1]} d(P(\alpha(t)), Q(\beta(t)))\}$$





Definicja metryki Hausdorffa

Funkcja d: $X \times X \to R^+$ jest metryką, jeżeli spełnia warunki:

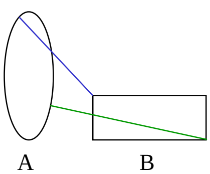
- $d(x, y) = 0 \Leftrightarrow x = y$
- $d(x, y) = d(y, x)$ - warunek symetrii
- $d(x, y) + d(y, z) \geq d(x, z)$ – warunek trójkąta

Metryka Hausdorffa jest określona na zbiorach:

$d_H(A,B) = \max \{d_{H+}(A,B), d_{H-}(A,B)\}$,

$d_{H+}(A,B) = \sup \{d(x,B) : x \in A\}$,

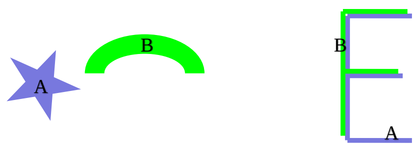
$d_{H-}(A,B) = \sup \{d(y,A) : y \in B\}$.



Odległość Hausdorffa

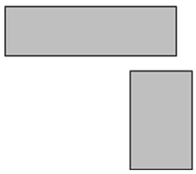
$d_{H+}(A,B) = \max \{d(x,B) : x \in A\}$,
$d_{H-}(A,B) = \max \{d(y,A) : y \in B\}$,

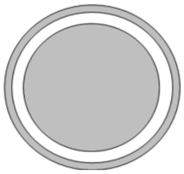
$d_H(A,B) = \max (d_{H+}(A,B), d_{H-}(A,B))$



Odległość Hausdorffa – przykłady

a) b)



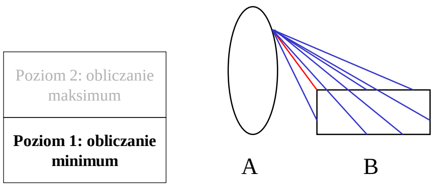


Własności metryki Hausdorffa istotne z punktu widzenia rozpoznawania obiektów

- Jednoznaczność: $d_H(A, B) = 0 \Leftrightarrow A=B$
- Informacja o parametrach transformacji (rozpoznawanie obiektów złożonych)
- Możliwość uwzględnienia dowolnych transformacji obiektu
- Przewidywalność – interpretacja intuicyjna
- Możliwość wykorzystania d_{H+} i d_{H-} w rozpoznawaniu obiektów częściowo przesłoniętych lub nieprawidłowo wyodrębnionych

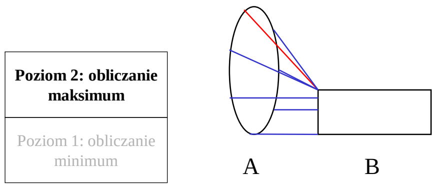
Poziom obliczeń: 1

$$d_{H+}(A,B) = \sup\{d(x,B) : x \in A\}$$

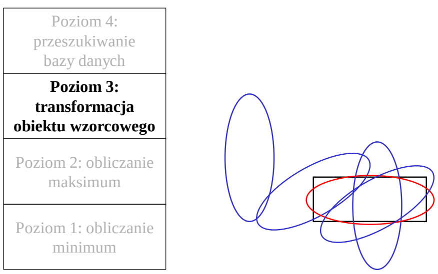


Poziom obliczeń: 2

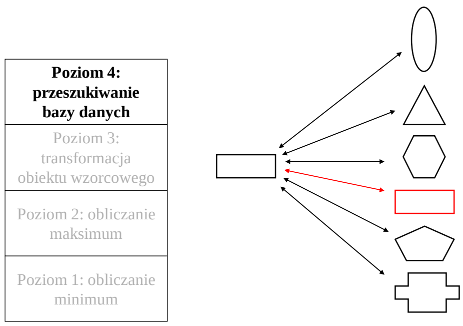
$$d_{H+}(A,B) = \sup\{d(x,B) : x \in A\}$$



Poziom obliczeń: 3



Poziom obliczeń: 4

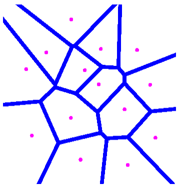


Metody przyspieszenia obliczeń metryki Hausdorffa

poziom 4: poszukiwanie obiektu	Eliminacja części biblioteki obiektów wzorcowych + optymalna kolejność przeszukiwania		wysoki poziom (końcowa faza obliczeń)
poziom 3: poszukiwanie transformacji	Eliminacja części przestrzeni transformacji obiektu	Optymalizacja odl. Hausdorffa po parametrach transformacji	
poziom 2: poszukiwanie maksimum	Metoda eliminacji podzbiorów konturu Eliminacja wnętrza obiektu		
poziom 1: poszukiwanie minimum	Metoda triangulacji Delaunay'a	Metoda *distance transform*	niski poziom (początkowa faza obliczeń)

Poziom 1:
Poszukiwanie minimum

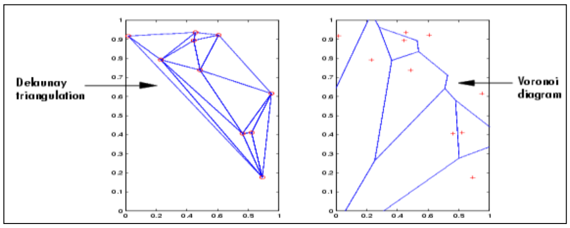
Diagram Woronoja i triangulacja Delaunay'a



Komórka Woronoja dla punktu P_i – zbiór punktów leżących bliżej P_i niż jakiegokolwiek innego elementu zbioru P

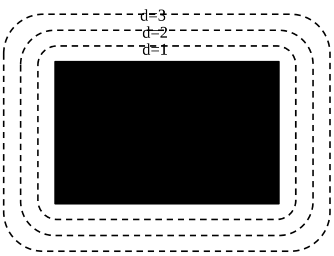
Diagram Woronoja – podział płaszczyzny (przestrzeni) na komórki Woronoja

Diagram Woronoja i triangulacja Delaunay'a



tri = delaunay(X,Y);
k = dsearch(X,Y,tri,x1,y1); % indeks punktu najbliższego (x1, y1)

Transformata odległościowa (distance transform)



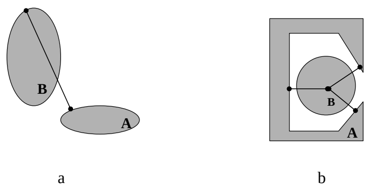
Poziom 2:
Poszukiwanie maksimum

a) eliminacja wnętrza obiektu

b) eliminacja podzbiorów konturu

Metoda eliminacji wnętrza obiektu

Odległość Hausdorffa nie zawsze jest realizowana na brzegu obiektu



$d_{H+}(B,A) = \max \{d(x,A) : x \in B\}$

Poziom 3: Transformacja obiektu wzorcowego

Optymalizowane parametry

Dla obiektów 2D o unikalnym kształcie można przyjąć 4 parametry transformacji:

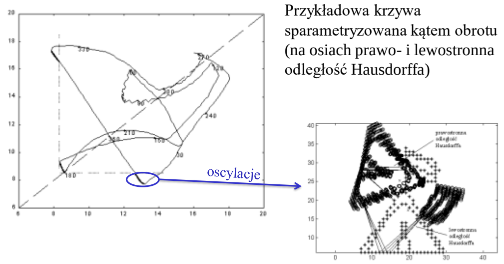
- Translacja x
- Translacja y
- Skalowanie (jednokładność)
- Obrót

Podejście A: dokładny przegląd całej przestrzeni z algorytmami obcinania (Rucklidge)

Podejście B: użycie metod optymalizacji

Problemy optymalizacyjne w przestrzeni transformacji

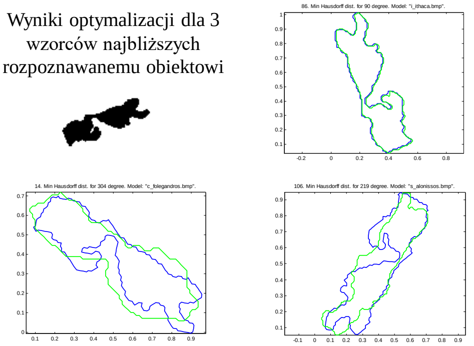
1. Odległość Hausdorffa ma nieciągłą pochodną względem parametrów tranformacji => użycie metod gradientowych jest niewskazane
2. Występuje wiele ekstremów lokalnych, zwłaszcza przy zmianie kąta obrotu



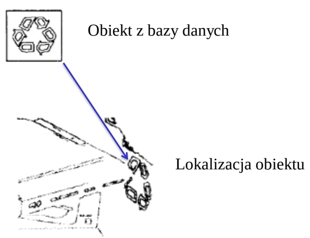
Przykładowa krzywa sparametryzowana kątem obrotu (na osiach prawo- i lewostronna odległość Hausdorffa)

Nałożenie obiektów dla optymalnej transformacji

Wyniki optymalizacji dla 3 wzorców najbliższych rozpoznawanemu obiektowi



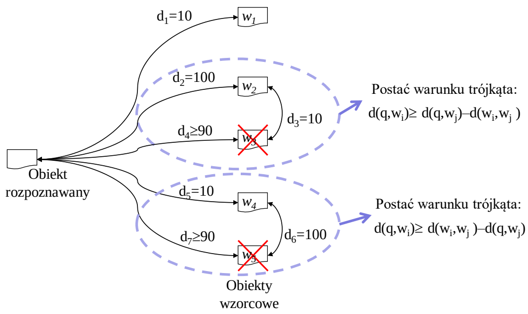
Alternatywna metoda – przeszukiwanie całej przestrzeni



Możliwość zlokalizowania obiektu lub jego fragmentu bez potrzeby segmentacji obrazu

Poziom 4:
Algorytmy nawigacji w bazie danych

Eliminacja wzorców



Nawigacja w bazie danych – eliminacja wzorców

Wykorzystanie cech: współczynników kształtu i niezmienników momentowych.

Cechy dwóch obiektów istotnie się różnią => kształt obiektów jest różny

Cechy dwóch obiektów są zbliżone lub identyczne => nic nie wiadomo

Zatem:

Cechy nie nadają się dobrze do rozpoznawania kształtów jako samoistne kryterium (chyba że mamy do czynienia z niewielkim zbiorem klas)

Cechy nadają się do wyeliminowania części wzorców z obliczeń

Wykorzystanie cech do eliminacji wzorców

Przykładowe cechy:

$$M_1 = (M_{20}+M_{02}) \,/\, m_{00}^2 ,$$

$$M_2 = ((M_{20}+M_{02})^2 + 4\, M_{11}^2) \,/\, m_{00}^4 ,$$

$$W_{BB} = \frac{S}{\sqrt{2\pi \iint r^2 ds}}$$

gdzie:
m_{00} – moment zwykły rzędu (0,0), czyli suma jasności wszystkich pikseli obrazu,
M_{ij} – moment centralny, wyrażony wzorem:

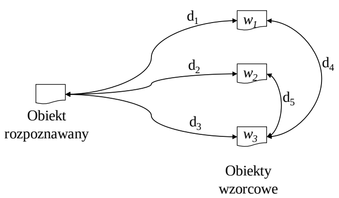
$$M_{pq} = \sum_{i=1}^{I} \sum_{j=1}^{J} (i - i_0)^p (j - j_0)^q f_{ij}$$

i_0, j_0 – współrzędne środka ciężkości
f_{ij} – jasność piksela (i,j),

Nawigacja w bazie danych – efektywna kolejność

Obliczyliśmy odległości rozpoznawanego obiektu od w_1 i w_2.

Pytamy czy w_3 jest dobrym kandydatem



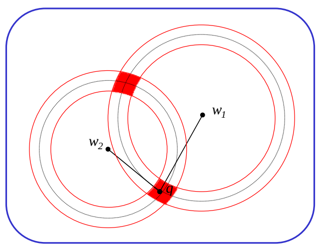
jeżeli $d_3 \to 0$ to

$|d_4 - d_1| \to 0$

i

$|d_5 - d_2| \to 0$

Nawigacja w bazie danych – efektywna kolejność



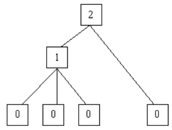
$$j = \arg\min_{Niesprawdz.} \max_{i \in Spr.} \frac{1}{r(i)} \left| d(w_i, w_j) - r(i) \right|$$

Rozpoznawanie obiektów złożonych przez dopasowanie

Modelowanie obiektów złożonych

Pojecie grafu i drzewa
graf G = {V, E},
krawędź – nieuporządkowana para wierzchołków

Drzewo – graf spójny acykliczny
Graf o n wierzchołkach jest drzewem ⇔
• jest spójny i nie zawiera cykli
• jest spójny i ma n-1 krawędzi
• każde 2 wierzchołki są połączone dokładnie 1 drogą
• usunięcie dowolnej krawędzi spowoduje utratę spójności grafu



Modelowanie obiektów złożonych

Wierzchołki grafu – podobiekty
Krawędzie grafu – relacje między obiektami

Typy relacji:
• relacje geometryczne (na lewo, powyżej, itp.)
• relacje topologiczne – określają wzajemne przyleganie obiektów i zawieranie się jednych obiektów w innych
• przynależności do obiektu nadrzędnego (mereologiczne)

Hipergrafy i modelowanie obiektów 3D

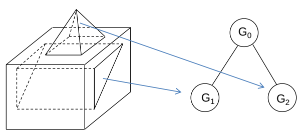
Przykładowe podejścia do modelowania obiektów 3D:

wierzchołki grafu	**krawędzie grafu**	**hiperkrawędzie**	**wielopoziomowość**
ściany figury	krawędzie figury	wierzchołki figury	zagregowanie elementów figury

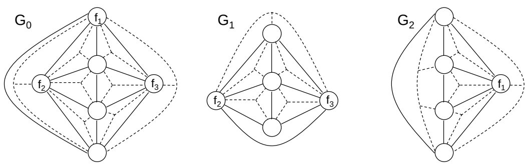
wierzchołki grafu	**krawędzie grafu**	**hiperkrawędzie**	**wielopoziomowość**
ściany figury	krawędzie figury	zagregowanie elementów figury	brak

wierzchołki grafu	**krawędzie grafu**	**hiperkrawędzie**	**wielopoziomowość**
wierzchołki figury	krawędzie figury	brak	brak

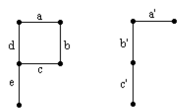
Hipergrafy i modelowanie obiektów 3D



wierzchołki grafu	**krawędzie grafu**	**hiper-krawędzie**	**wielopozio-mowość**
ściany figury	krawędzie figury	wierzchołki figury	zagregowanie elementów figury



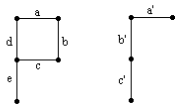
Metody *matchingu* grafów



W tym przykładzie:

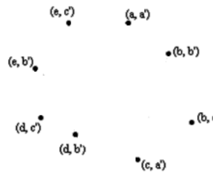
Podobiekty: odcinki
Atrybuty podobiektów: poziomy/pionowy
Relacje: połączone/niepołączone

Metody *matchingu* grafów

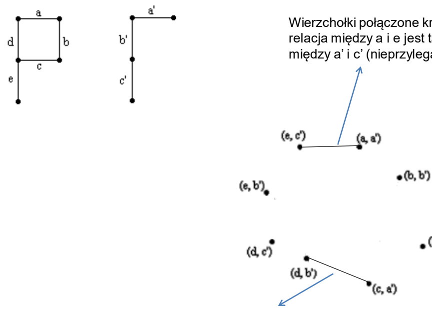


Graf dopasowań:

Dla każdej pary podobiektów o takich samych atrybutach tworzony jest wierzchołek grafu dopasowań



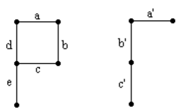
Metody *matchingu* grafów



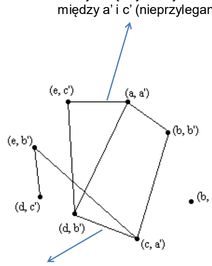
Wierzchołki połączone krawędzią, bo relacja między a i e jest taka sama jak między a' i c' (nieprzyleganie)

Wierzchołki połączone krawędzią, bo relacja między c i d jest taka sama jak między a' i b' (przyleganie)

Metody *matchingu* grafów



Wierzchołki połączone krawędzią, bo relacja między a i e jest taka sama jak między a' i c' (nieprzyleganie)

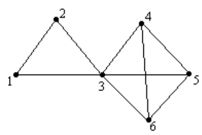


Wierzchołki połączone krawędzią, bo relacja między c i d jest taka sama jak między a' i b' (przyleganie)

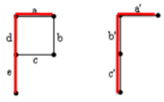
Metody *matchingu* grafów

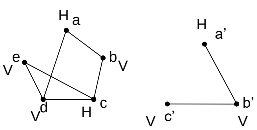
Graf pełny – graf w którym każda para wierzchołków jest połączona
Podgraf pełny (klika) – podgraf (podzbiór grafu), który jest grafem pełnym
Maksymalny podgraf pełny – nie jest zawarty w żadnym innym podgrafie pełnym
Największy podgraf pełny – podgraf pełny o największej liczbie wierzchołków

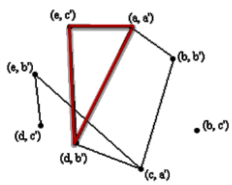
Największa klika jest równocześnie maksymalną kliką, ale nie odwrotnie.



Metody *matchingu* grafów

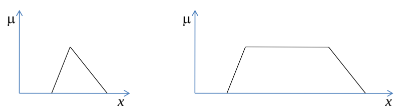




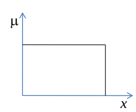


Modelowanie relacji rozmytych

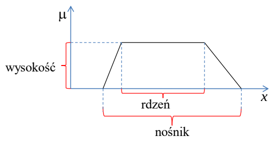
Zbiór rozmyty A – zbiór uporządkowanych par $\{x, \mu_A(x)\}$,
$\mu_A(x)$: X → [0, 1]



Dla zbioru nierozmytego:



Modelowanie relacji rozmytych



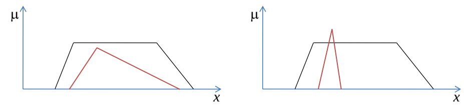
Nośnik *(support)* zbioru rozmytego - zbiór elementów *x*, dla których $\mu_A(x) > 0$
Rdzeń *(core)* zbioru - zbiór elementów x, dla których $\mu_A(x) = 1$
Wysokość zbioru - największa wartość funkcji przynależności

wysokość=1 => zbiór znormalizowany

Modelowanie relacji rozmytych

Równość zbiorów rozmytych: $A = B \leftrightarrow \forall_{u \in U} \mu_A(u) = \mu_B(u)$

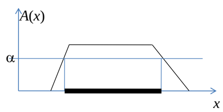
Zawieranie zbiorów rozmytych: $A \subset B \leftrightarrow \forall_{u \in U} \mu_A(u) \leq \mu_B(u)$



Zbiory α-poziomicowe

Zbiory o poziomie α (α-poziomicowe)

A – zbiór rozmyty
A_α – zbiór o poziomie α:
$A_\alpha(x) := \{x: A(x) \geq \alpha\}$



Zbiory o poziomie α umożliwiają reprezentowanie zbioru rozmytego przez zbiory nierozmyte:

$$A = \bigcup_{\alpha} \alpha A_\alpha$$

Grafy rozmyte

Grafy rozmyte
W grafie rozmytym wierzchołkom i krawędziom przypisane są wartości funkcji przynależności.
Funkcja przynależności dla wierzchołka nie może być mniejsza niż dla dochodzącej do niego krawędzi.

Graf α-poziomicowy
Graf α-poziomicowy to zbiór tych wierzchołków i krawędzi grafu G, dla których wartość funkcji przynależności jest większa lub równa α

W przypadku modelowania obiektów złożonych rozmycie modeluje niejednoznaczność cech podobiektów i relacji między nimi

Matching grafów rozmytych