

Zastosowania obliczeń inteligentnych do wyszukiwania w obrazowych bazach danych

Tatiana Jaworska

Jaworska@ibspan.waw.pl

www.ibspan.waw.pl/~jaworska



Istniejące systemy - Google

Szukaj w Google - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

Garou- A m... GONGO - e... Définition : ... Szukaj w G... Tatiana Jaw... +

images.google.com/search?tbs=sbi:AMhZZiuTjiNI ☆ CBIR

+Ty Wyszukiwarka **Grafika** Wideo Mapy Wiadomości Gmail Więcej - Zaloguj

Google ja_07_2011.jpg

Wyszukiwarka

Wszystko

Grafika

Mapy

Filmy

Wiadomości

Zakupy

Więcej

Wyszukiwanie obrazem

Podobne wizualnie

Więcej rozmiarów

Kiedykolwiek

Ostatnia godzina

Ostatnie 24 godziny

Ostatni tydzień

Ostatni miesiąc

Ostatni rok

Zakres dat...

Rozmiar obrazu: 1152 × 1596

Nie znaleziono innych rozmiarów tego obrazu.

[Obrazy podobne wizualnie](#) - Zgłoś grafiki

ja_07_2011 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz

Tati1204 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz

Istniejące systemy - Google

Szukaj w Google - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

Garou - Firs... GONGO - e... Définition : ... Szukaj w G... Tatiana Jaw... +

images.google.com/search?tbs=sbi:AMhZZitFwHZ ☆ CBIR

+Ty Wyszukiwarka **Grafika** Wideo Mapy Wiadomości Gmail Więcej Zaloguj

Google Tati1204.jpg x

Wyszukiwarka 1 wynik (0.97 s)

Strony zawierające podobne obrazy

Wszystko

Grafika

Mapy

Filmy

Wiadomości

Zakupy

Więcej

Wyszukiwanie obrazem

Podobne wizualnie

Więcej rozmiarów

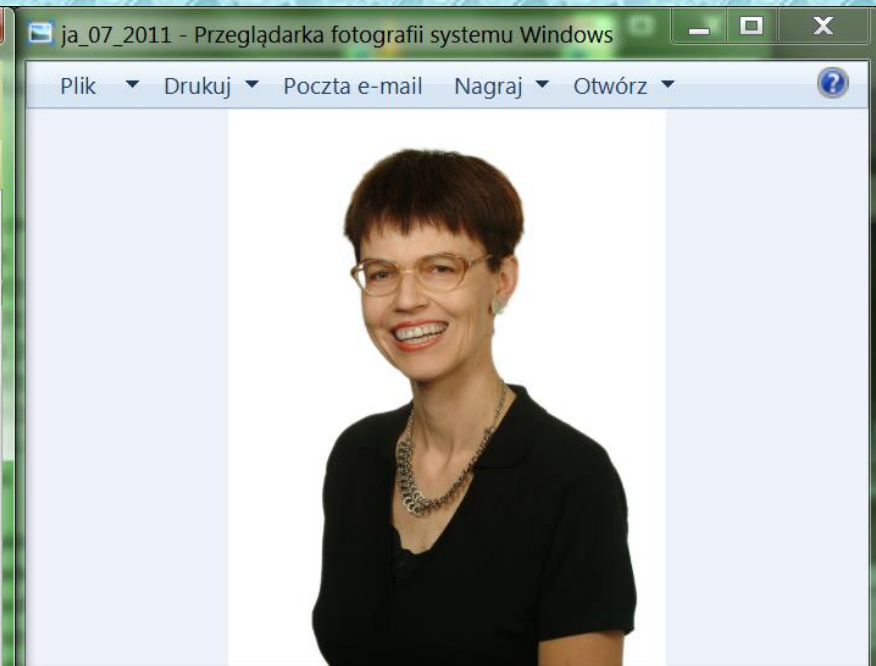
Kiedykolwiek

Ostatnia godzina




ja_07_2011 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz



Tati1204 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz



Istniejące systemy - Google

Szukaj w Google - Mozilla Firefox

plik Edycja Widok Historia Zakładki Narzędzia Pomoc

Garou Je vais t'aimer - ... GONGO - encyclopedi... Définition : hype - Dict... Szukaj w Google

www.google.com/search?hl=pl&safe=active&sa=G&gbv=2&tbs=simg:CAESVhpUCxCo1NgEGgQIAA

+Ty Wyszukiwarka **Grafika** Wideo Mapy Wiadomości Gmail Więcej

Google

Wyszukiwarka Około 199,000 wyników (0.32 s)

- Wszystko
- Grafika
- Mapy
- Filmy
- Wiadomości
- Zakupy
- Więcej

Podobne wizualnie



- Podobne wizualnie**
- Więcej rozmiarów
- Wszystkie wyniki**
- Według tematyki
- Dowolny rozmiar**
- Duże
- Średnie
- Ikony
- Większe niż...

okapi1 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz

Navigation icons: back, forward, home, search, refresh, close.

Istniejące systemy - Google

Szukaj w Google - Mozilla Firefox

Edycja Widok Historia Zakładki Narzędzia Pomoc

On n'a pas toute la so... Match Up Results Définition : mots de tr... Szukaj w Google

www.google.com/search?hl=pl&safe=active&sa=G&gbv=2&tbs=simg:CAESWhpYCxCwJkC

Wyszukiwarka Grafika Wideo Mapy Wiadomości Gmail Więcej -

Google

Wyszukiwarka

Około 816,000 wyników (0.51 s)

Podobne wizualnie



dom_core1 - Przeglądarka fotografii systemu Windows

Plik Drukuj Poczta e-mail Nagraj Otwórz



Navigation icons: search, print, back, forward, refresh, close.

parametry tekstury
pobierz teksturę

Wybierz wybraną warstwę

Istniejące systemy – CIRES i FIDS


CIRES: Results - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc


extrait d'un c... Match Up Re... Définition : e... CIRES: Results shapiro, lind... +

amazon.ece.utexas.edu/~qasim/cgi/process.cgi Malibu KIA dealer

Query Image




Feedback Image Set




Relevance feedback type: Cluster. Weights: Perceptual Grouping = 0.33, Color = 0.33, Texture = 0.33, L, A, B channels.
For relevance feedback, please select the check boxes below each image, and then select the feedback type. Note that NS = "Not Sure".


Retrieved Images



Yes NS No Yes NS No Yes NS No Yes NS No Yes NS No



Yes NS No Yes NS No Yes NS No Yes NS No Yes NS No



Yes NS No Yes NS No Yes NS No Yes NS No Yes NS No


demo: Fids - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc



extrait d'... Match U... Définitio... demo: ... shapiro, l... +

www.cs.washington.edu/research Malibu KIA dealer

Fids demo



Put In Cart
Check Out



Put In Cart
Check Out

demo: Fids - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

extrait d'... Match U... Définitio... demo: ... shapiro, l... +

www.cs.washington.edu/research Malibu KIA dealer

Fids demo



Put In Cart
Check Out



Put In Cart
Check Out

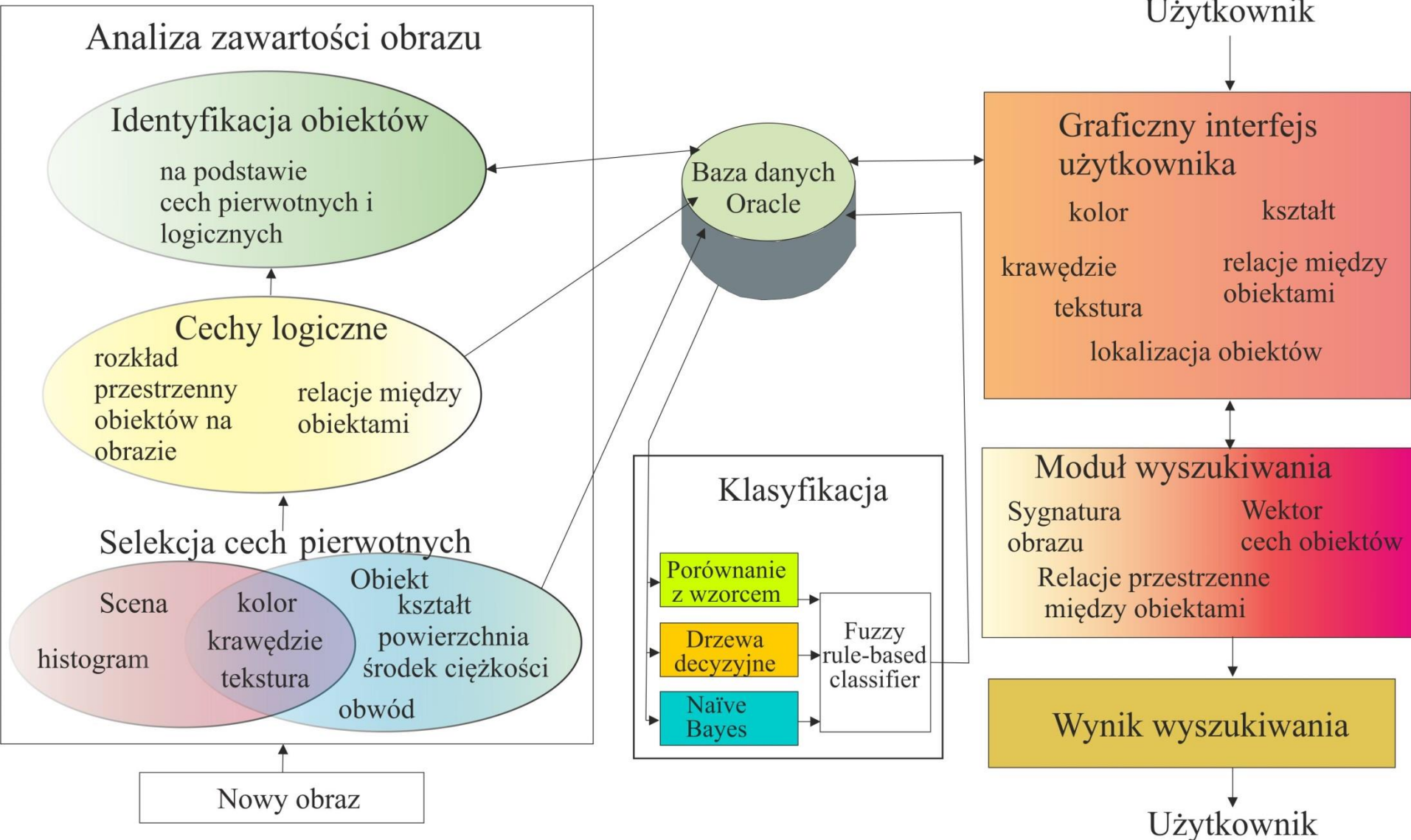
Podstawowe pojęcia

- Content-Based Image Retrieval (CBIR) – obrazowa baza danych
 - moduł odpowiedzialny za przetwarzanie obrazu
 - baza danych
 - interfejs użytkownika GUI
 - silnik wyszukiwania

Schemat ogólny obrazowej bazy danych

Przetwarzanie wsadowe

Przetwarzanie bezpośrednie



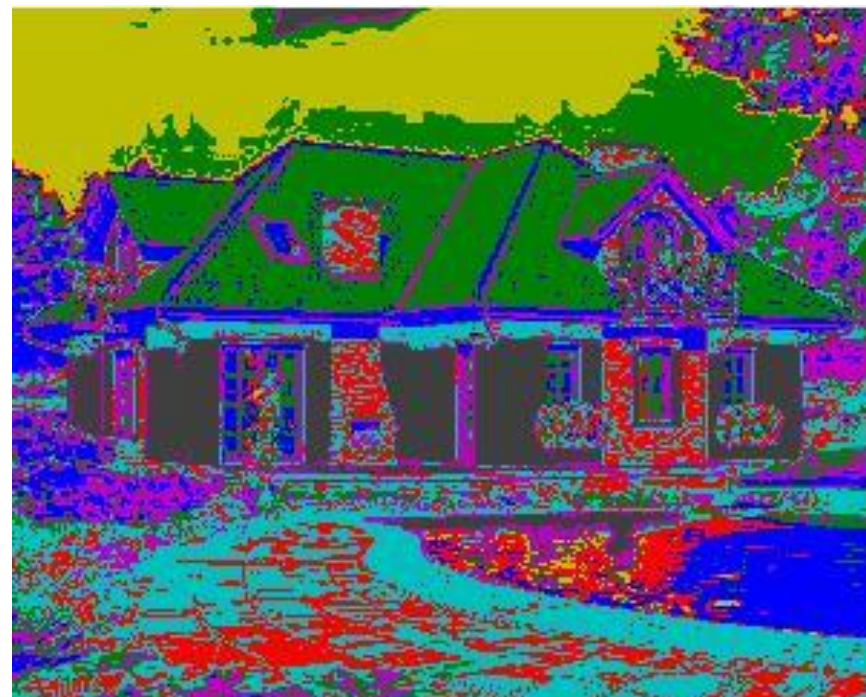
Segmentacja obrazu metodą c-środków



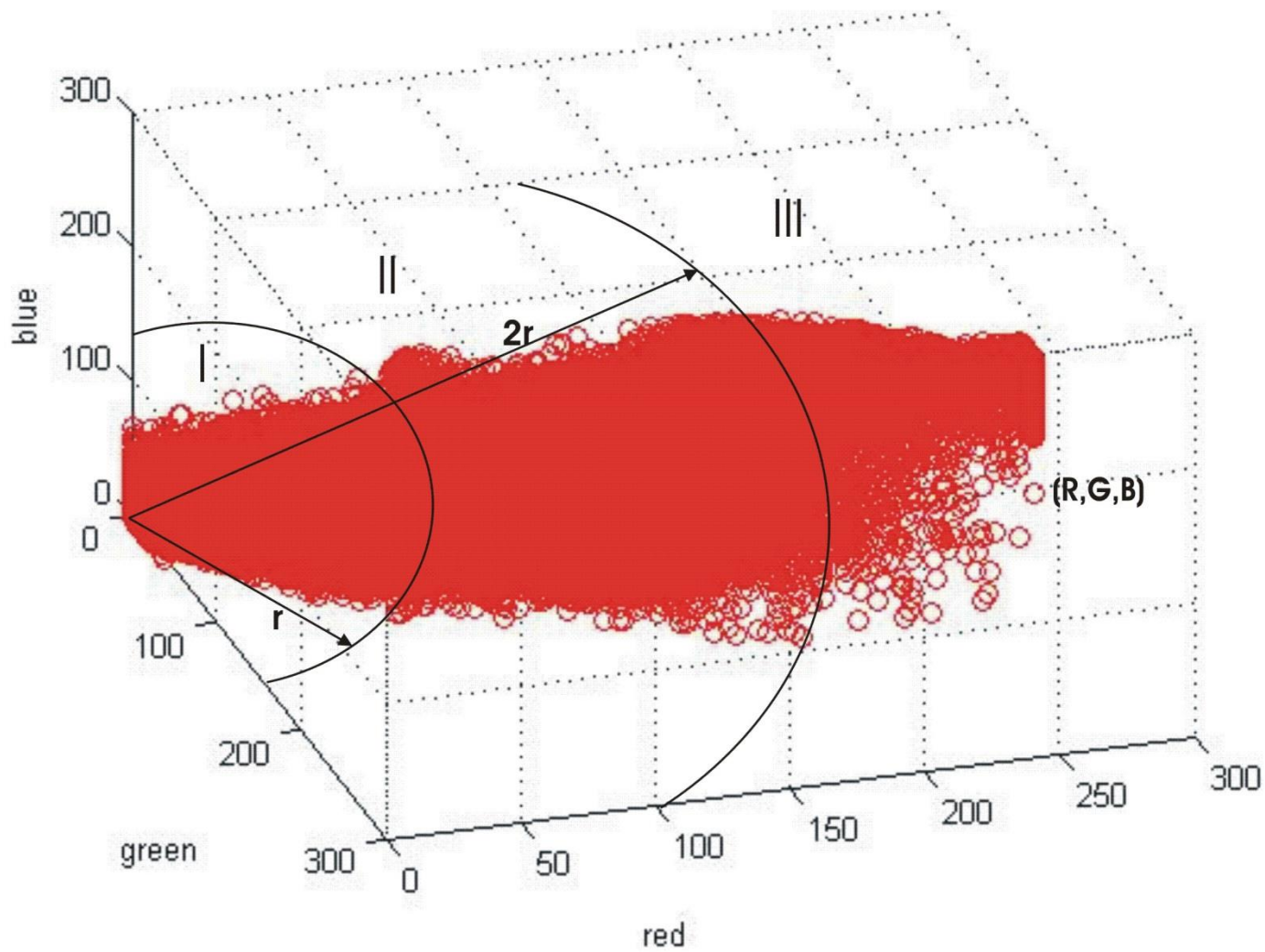
$C = 5$



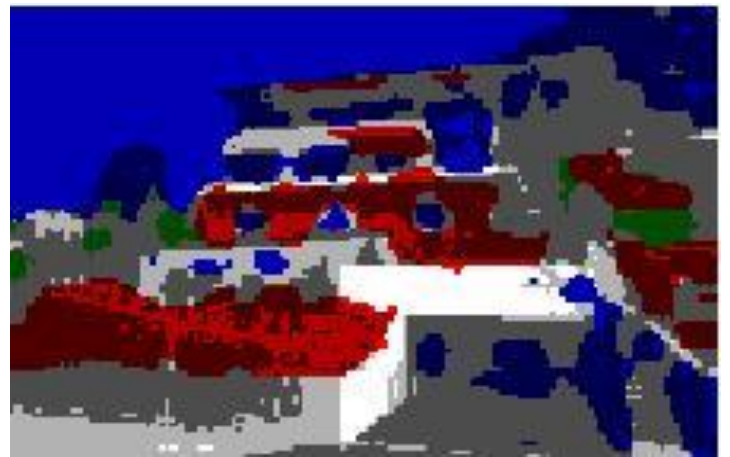
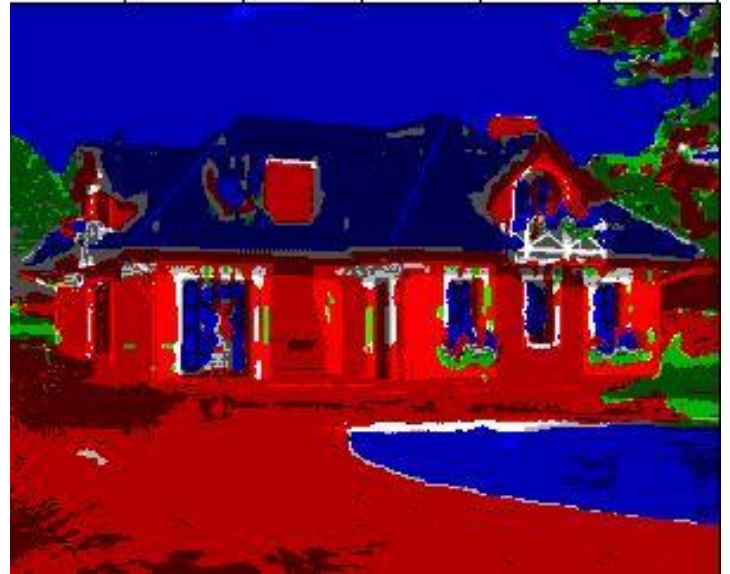
$C = 12$



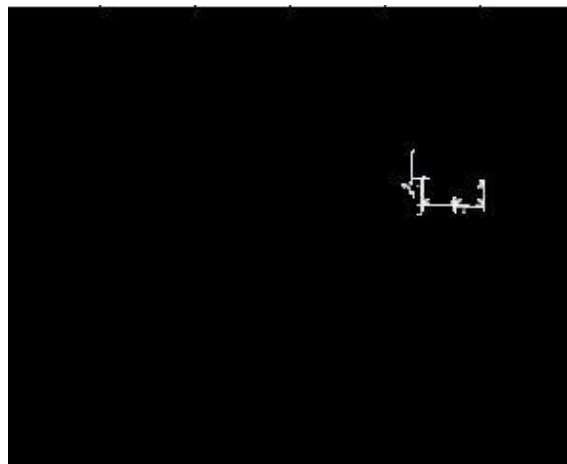
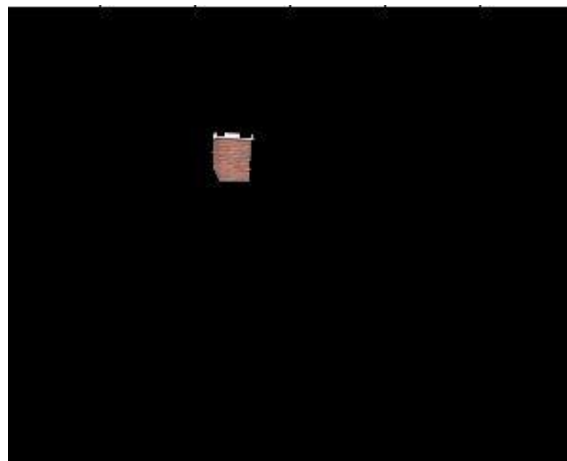
Idea algorytmu opartego na kolorach



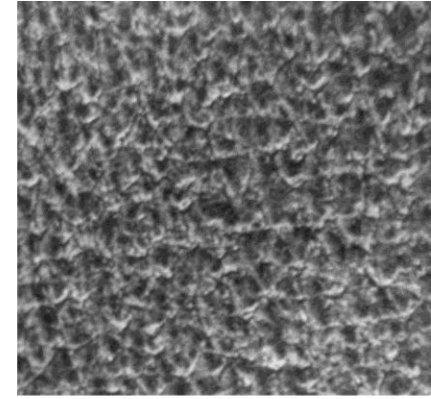
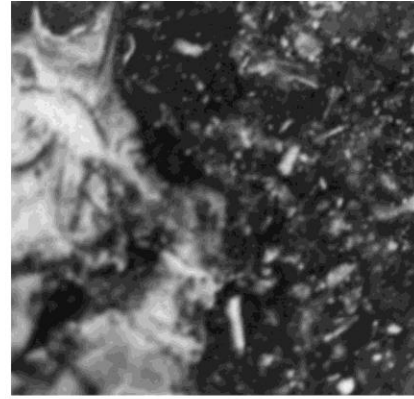
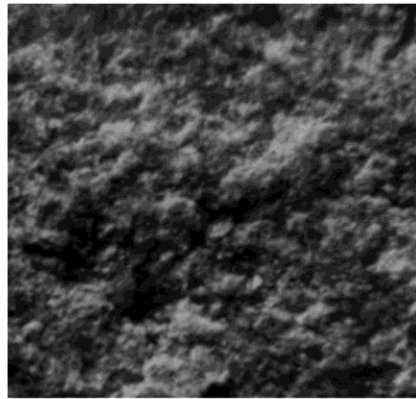
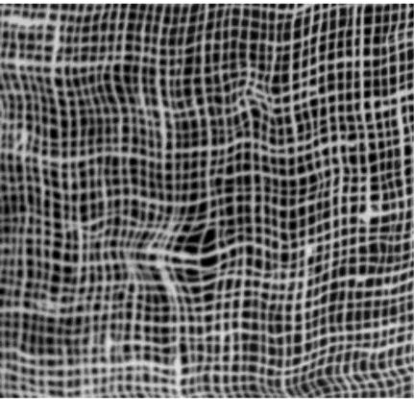
Segmentacja obrazu oparta na kolorach



Wydzielenie obiektów



Tekstury

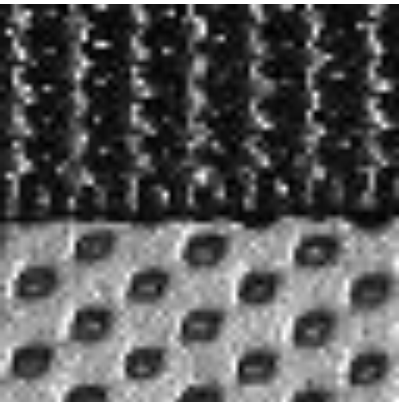


a)

b)

c)

d)



- a) tkanina
- b) kamień
- c) marmur
- d) korek

Tekstura może być opisywana jako:

- Dwuwymiarowy proces Markowa
- Dwuwymiarowy histogram
- Korelacja transformat Fouriera tekstur
z wcześniej przygotowanymi maskami

Falki dwuwymiarowe

Dwuwymiarowe transformacje falkowe obrazu sprowadzają się do rzutowania obrazu na ustalony ortonormalny układ falek. Ze względu na dyskretny charakter obrazów, będziemy się posługiwać falkami dyskretnymi.

Dwuwymiarowy separowany układ falek ortonormalnych utworzony konwencjonalnie na podstawie układu jednowymiarowego ma postać:

$$\{\psi_{j,k}(x), \psi_{l,m}(y)\}_{(j,k,l,m) \in \mathbb{Z}^4}$$

gdzie $\psi_{j,k}(x)$ jest k -tą falką jednowymiarową na poziomie j -tym analizy wielorozdzielczej.

Wykorzystując założenia analizy wielorozdzielczej można pokazać, że wyżej zaproponowany układ jest nadmiarowy i można go zredukować do następującego układu:

$$\{\psi_{j,k,m}^1(x, y), \psi_{j,k,m}^2(x, y), \psi_{j,k,m}^3(x, y)\}_{(j,k,m) \in \mathbb{Z}^3}$$

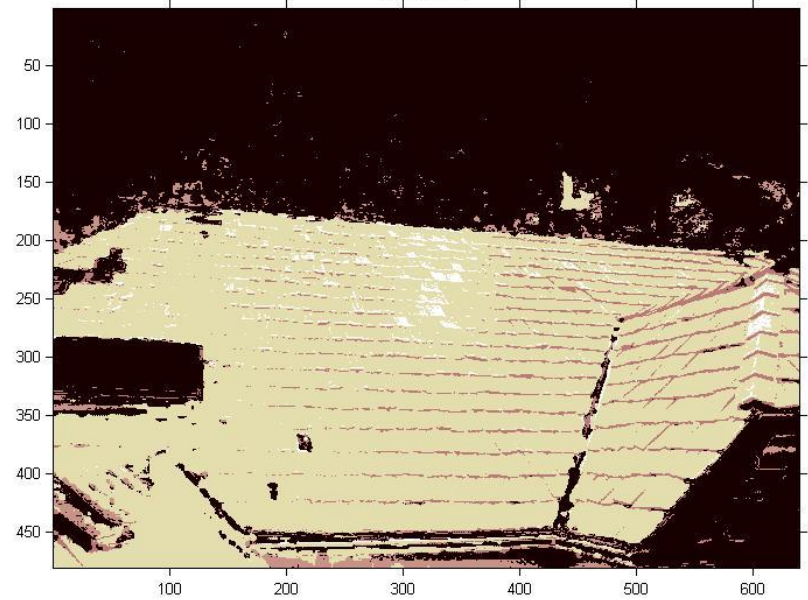
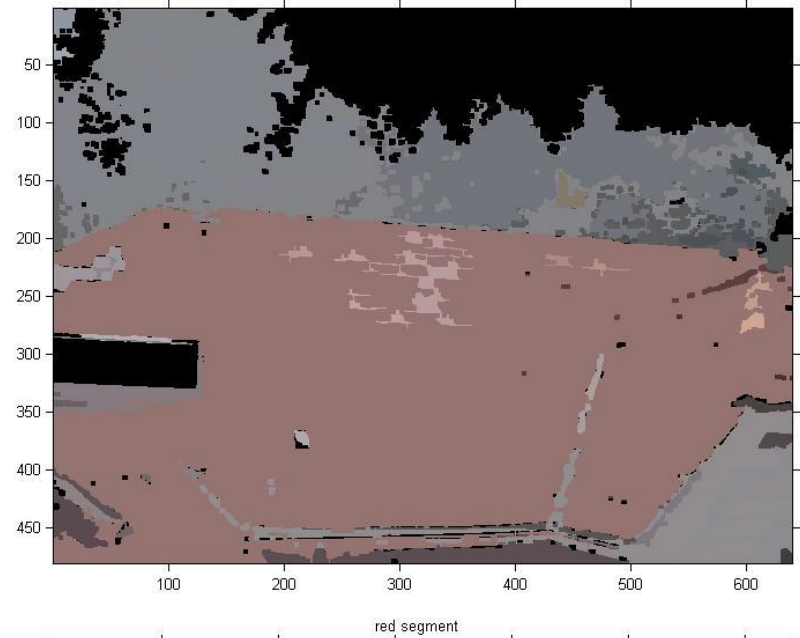
gdzie:

$$\psi^1(x, y) = \phi(x) \psi(y)$$

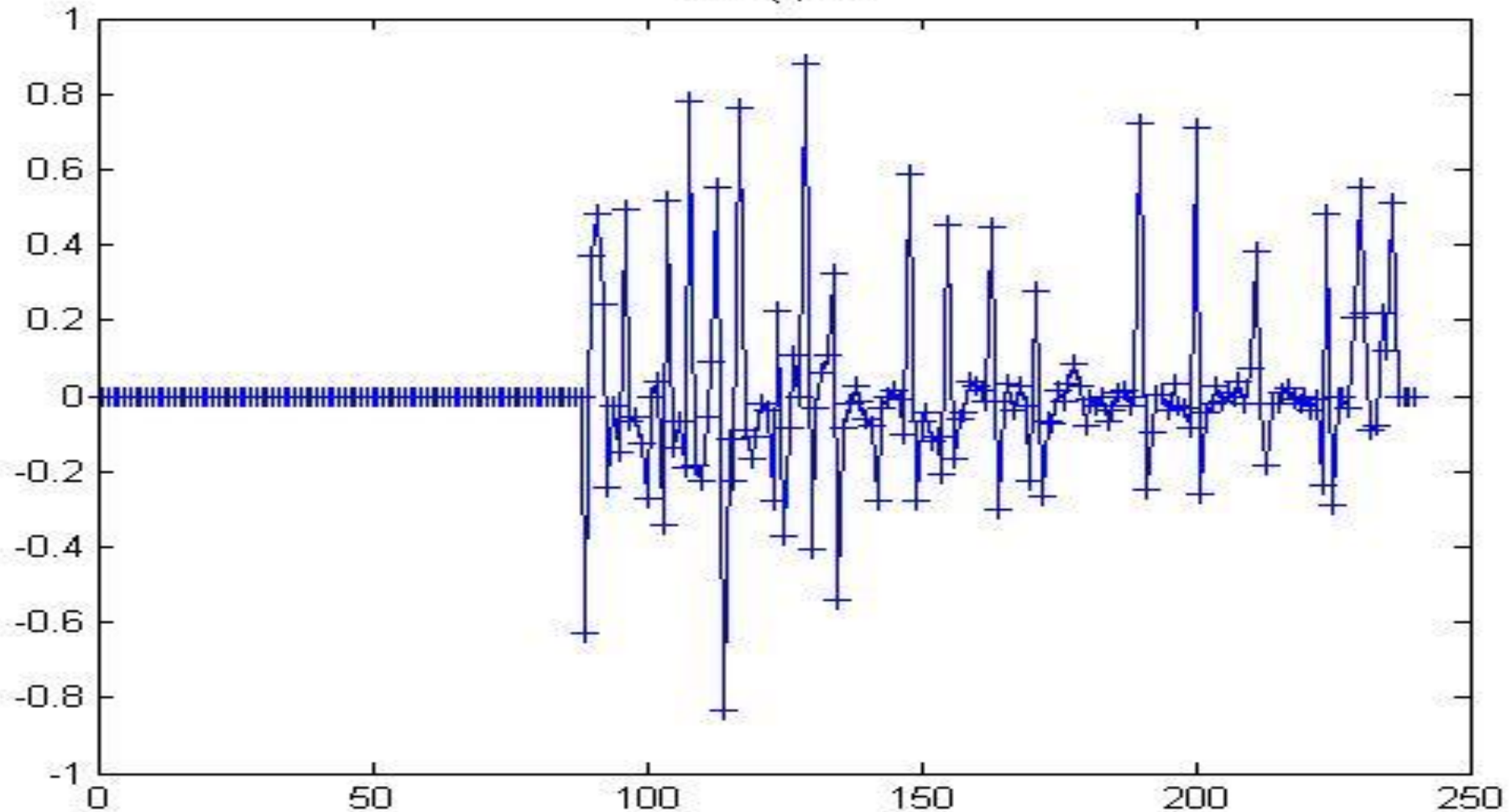
$$\psi^2(x, y) = \psi(x) \phi(y)$$

$$\psi^3(x, y) = \psi(x) \psi(y)$$

Obiekt z fakturą



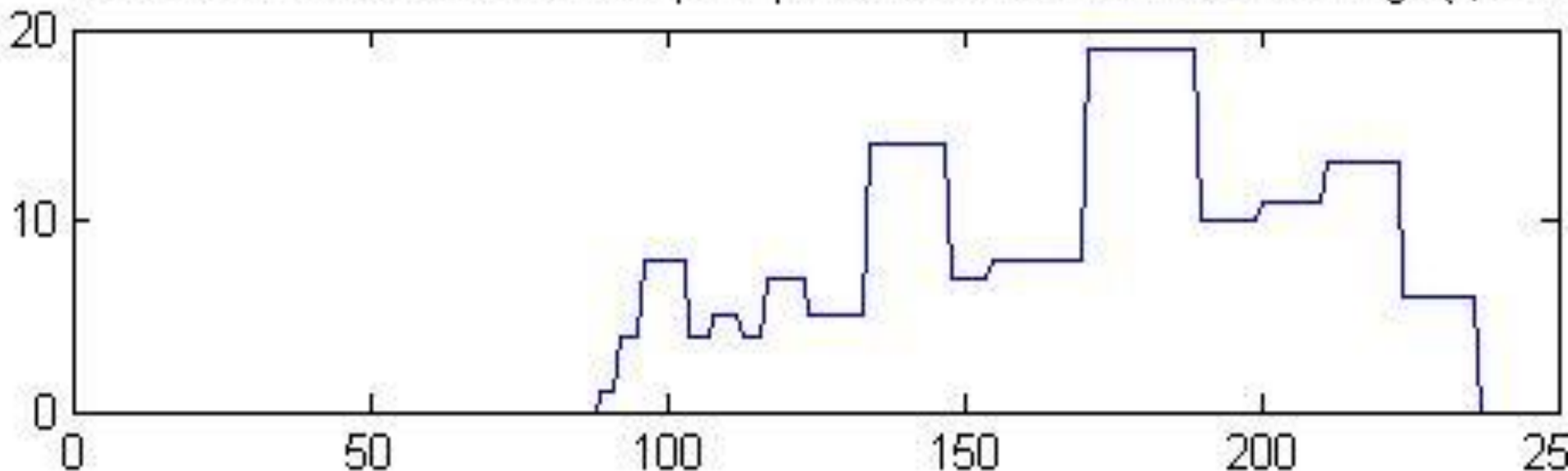
Mapa odległości oparta na transformacji falkowej



Przekrój przez detale horyzontalne
dla 100-tej kolumny (falka Haara, $j=1$)

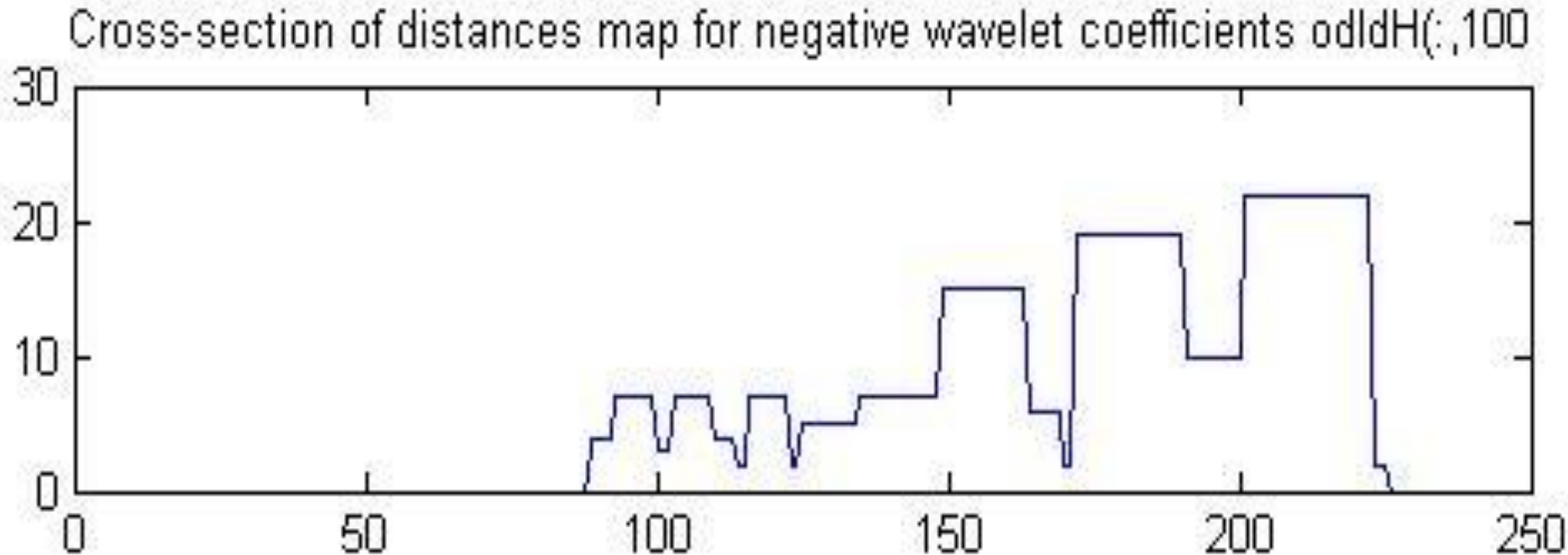
Odległości liczone na podstawie dodatnich współczynników falkowych

Cross-section of distances map for positive wavelet coefficients `odlgH(:,100)`



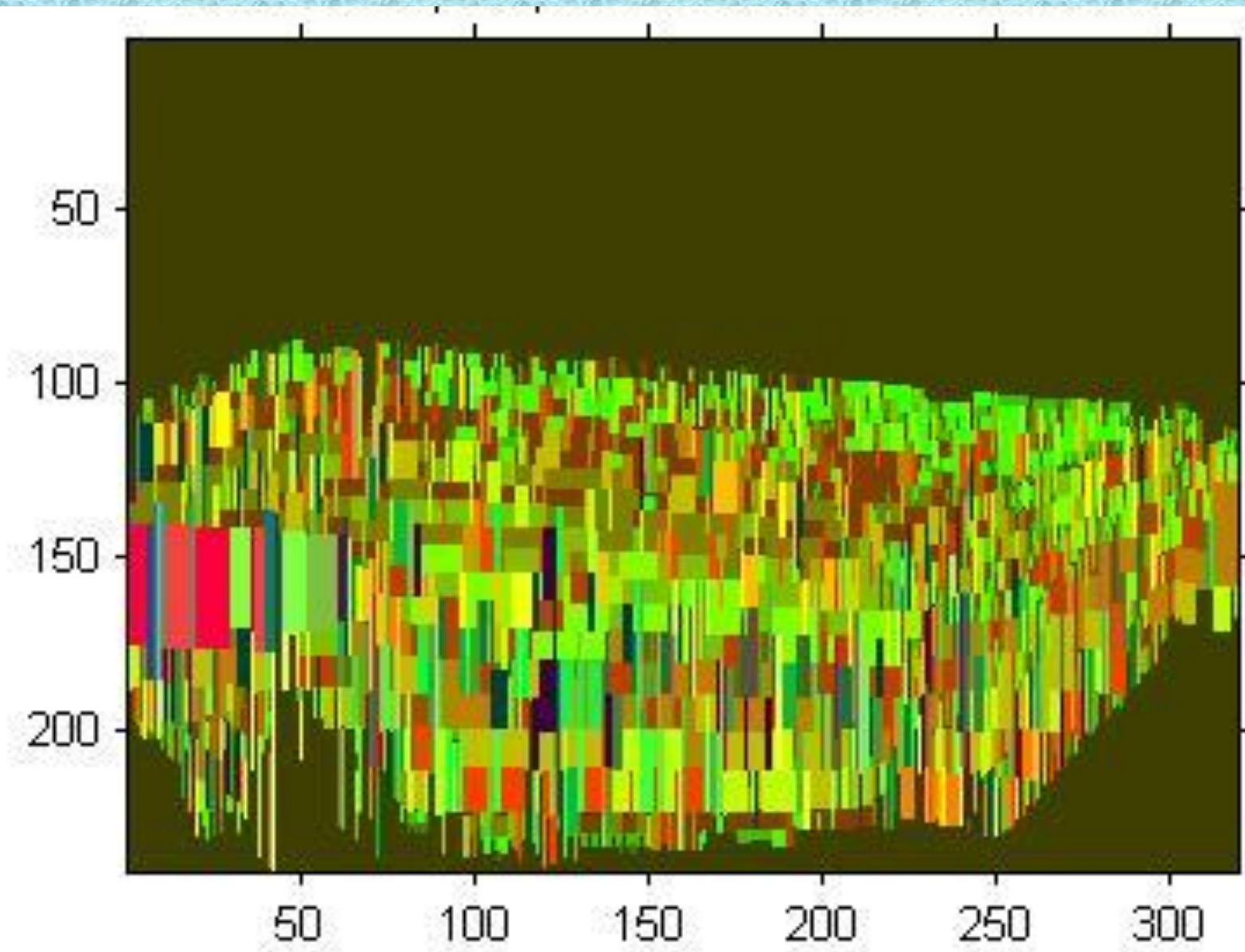
Odległości wyznaczone dla progu odcięcia 1% wartości maksymalnej dla całej macierzy detali horyzontalnych dodatnich. Przekrój przez 100-ną kolumnę.

Odległości liczone na podstawie ujemnych współczynników falkowych

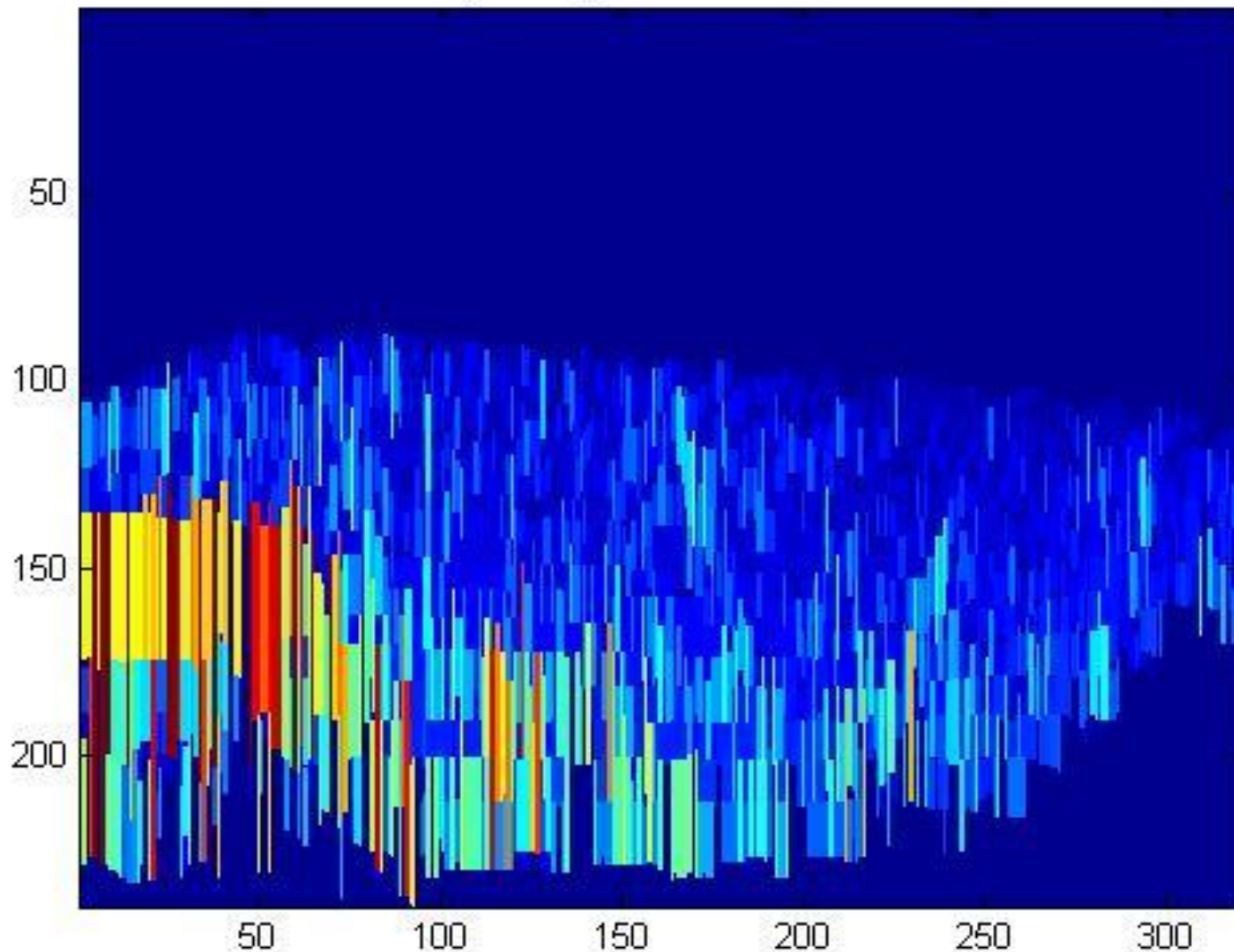


Odległości wyznaczone dla progu odcięcia 1% wartości minimalnej dla całej macierzy detali horyzontalnych ujemnych.
Przekrój przez 100-ną kolumnę

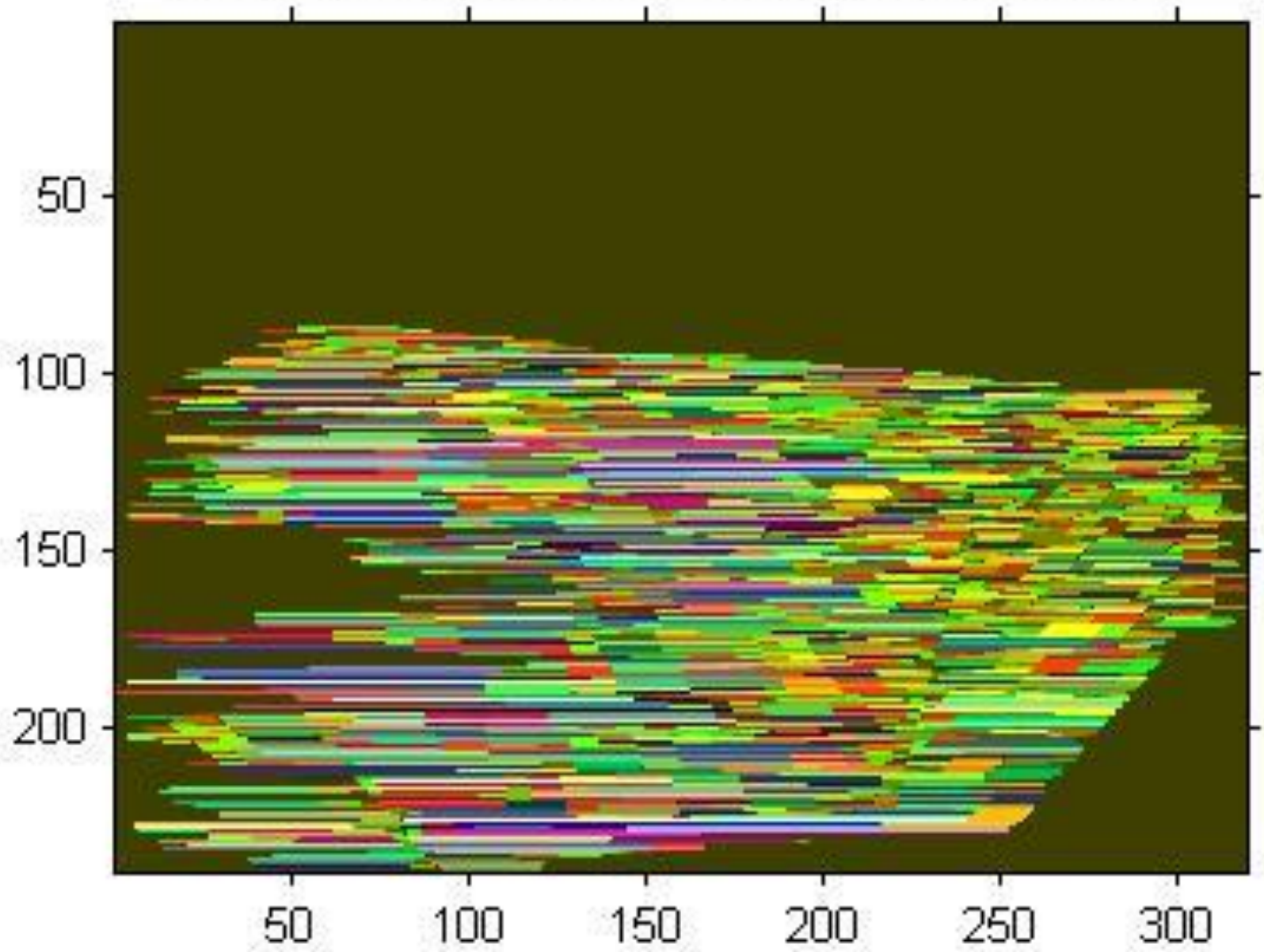
Mapa odległości wyznaczonych na podstawie horyzontalnych współczynników dodatnich



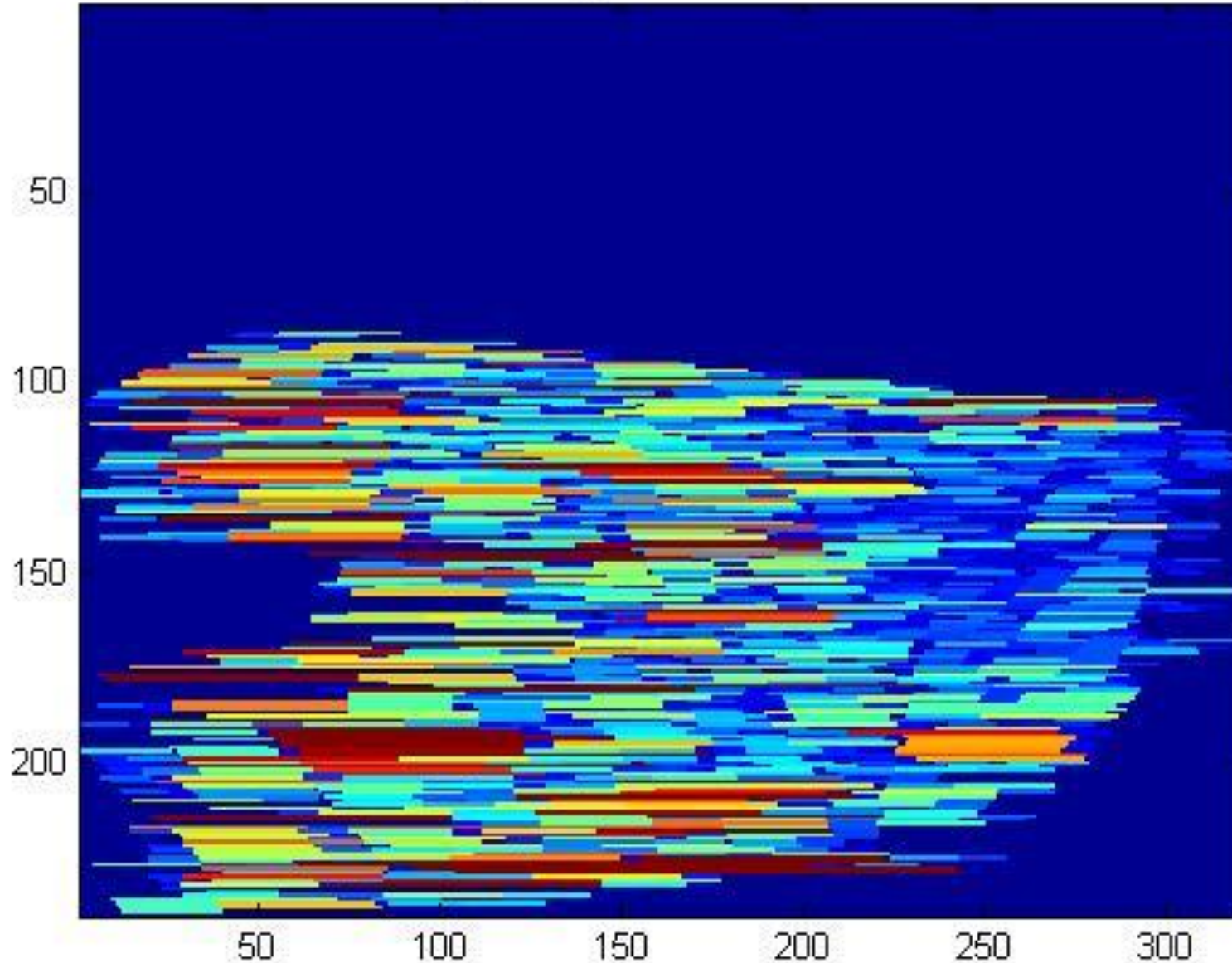
Mapa odległości wyznaczonych na podstawie horyzontalnych współczynników ujemnych



Mapa odległości wyznaczonych na podstawie wertykalnych współczynników dodatnich

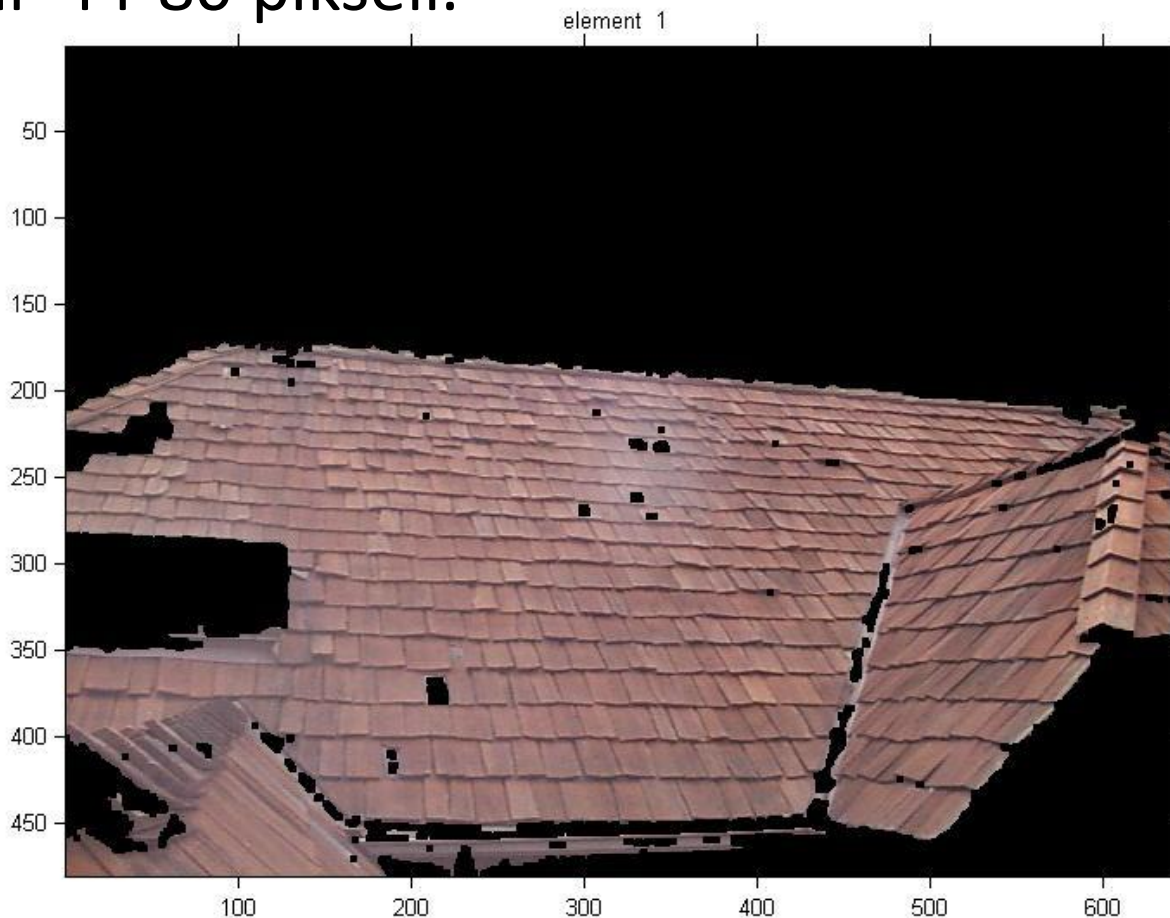


Mapa odległości wyznaczonych na podstawie wertykalnych współczynników ujemnych

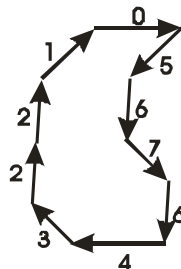
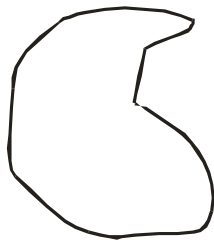
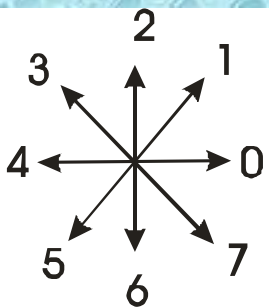


Parametry tekstury

Na podstawie wykonanych map odległości można oszacować rozmiar pionowy dachówki na 15-20 falek czyli 30-40 pikseli, natomiast rozmiar pionowy na and 22-40 falek czyli 44-80 pikseli.



Różne deskryptory kształtu



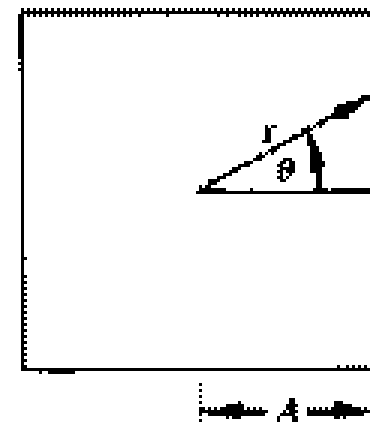
21056...

Definicja

Obiekt

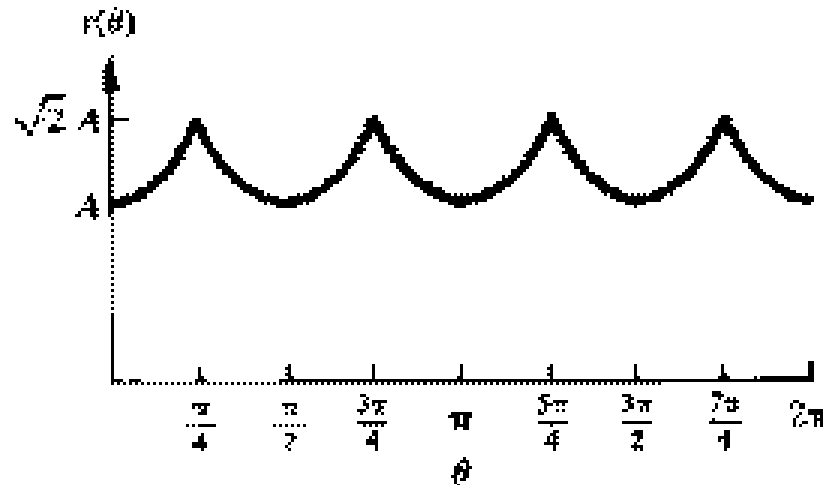
Sposób
kodowania

Kod



$$r(\theta) = \sqrt{[x(\theta) - \bar{x}]^2 + [y(\theta) - \bar{y}]^2}$$

$$a_n = \frac{1}{N} \sum_{\theta=0}^{N-1} r(\theta) \exp\left(\frac{-j2\pi n\theta}{N}\right)$$



Wielomiany Zernike'a

Ciąg wielomianów ortogonalnych nad dyskiem jednostkowym

$$V_{pq}(x, y) = R_{pq}(r)e^{-iq\theta}, \quad r \in [0, 1]$$

$$V_{pq}(r, \theta) = R_{pq}(r) = \sum_{s=0}^{(p-q)/2} (-1)^s \frac{(p-s)!}{s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} r^{p-2s}$$

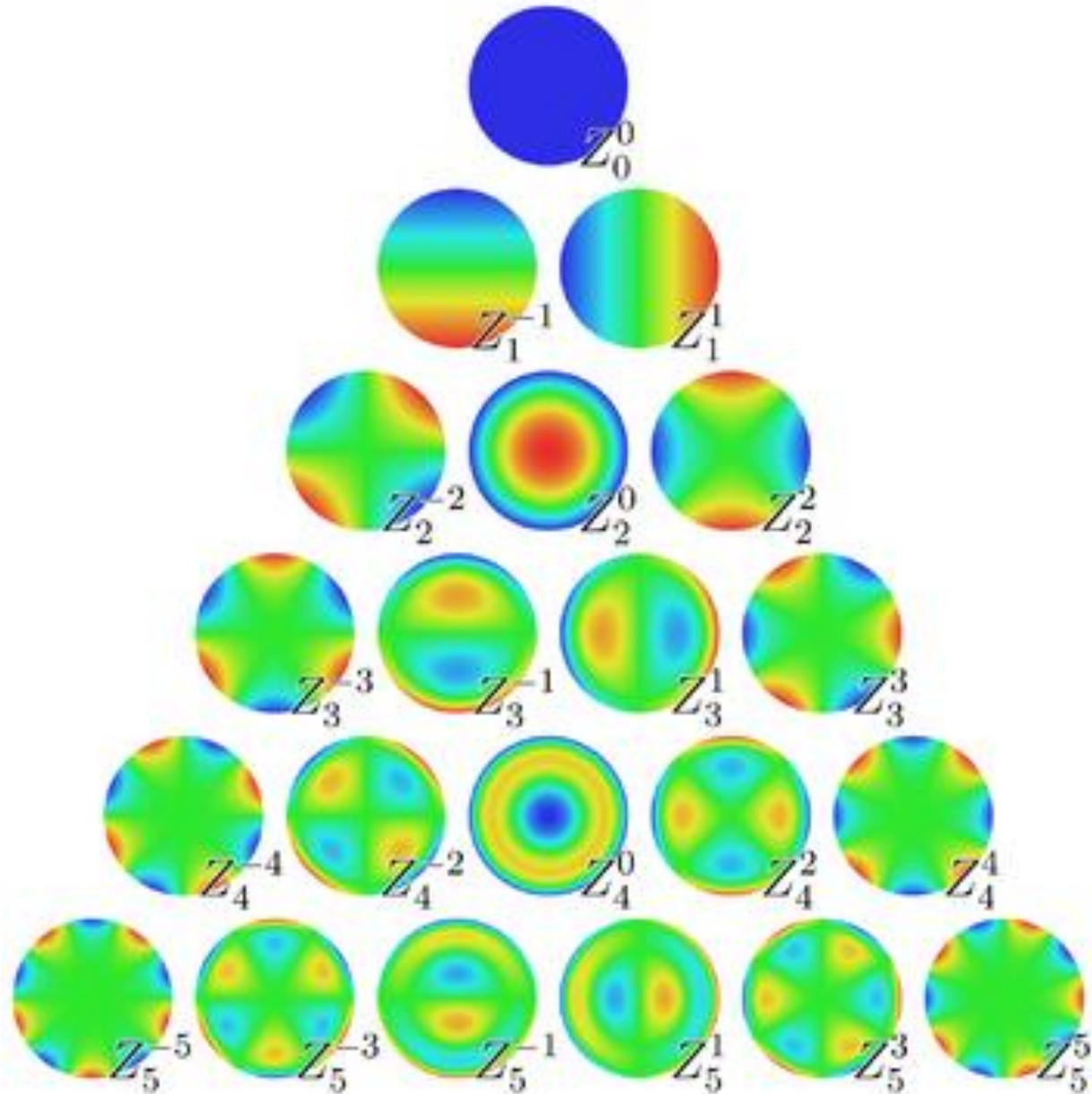
gdzie: $r = \sqrt{x^2 + y^2}$ jest długością wektora od środka dysku do punktu (x, y) . Wielomiany o wartościach rzeczywistych

$$V_{pq}(x, y) = \sum_{k=q}^p B_{pqk} r^k e^{iq\theta}$$

gdzie $p - |q|$ jest parzyste, $0 \leq |q| \leq p$ i $p \geq 0$.

Niech $s \rightarrow (p - k)/2$, wtedy wielomiany mogą być reprezentowane jako

Funkcje bazowe Zernike'a



Momenty Zernike'a

Wtedy B_{pqk} są współczynnikami tego wielomianu:

$$B_{pqk} = \frac{(-1)^{\frac{p-k}{2}} \left(\frac{p+k}{2}\right)!}{\left(\frac{p-k}{2}\right)! \left(\frac{k+q}{2}\right)! \left(\frac{k-q}{2}\right)!}$$

Dwuwymiarowe momenty Zernike'a rzędu p z powtórzeniami q dla obrazu o funkcji intensywności $f(x, y)$

$$Z_{pq} = \frac{p+1}{\pi} \int \int_{x^2+y^2 \leq 1} V_{pq}^*(x, y) f(x, y) dx dy$$

gdzie $V_{pq}^*(x, y) = V_{p, -q}(x, y)$

Właściwości momentów Zernike'a

1. Tak zdefiniowane momenty Zernike'a są niezmiennicze względem rotacji.
2. Aby uzyskać niezmienniczość względem przesunięcia trzeba centrum obiektu umieścić w środku układu współrzędnych.
3. Natomiast niezmienniczość skali uzyskujemy poprzez skalowanie obiektu.

Przykład dopasowania dwóch obiektów na podstawie pierwszych 10 momentów Zernike'a



elementy1

File

obraz elementu - pytania

50
100
150
200
250

50 100 150 200 250 300 350 400

match

element dopasowywany

50
100
150
200
250
300
350
400
450
500

100 200 300 400 500 600 700

Dopasuj elementy

Stopień niedopasowania -0.0066514

20
40
60
80
100
120
140
160

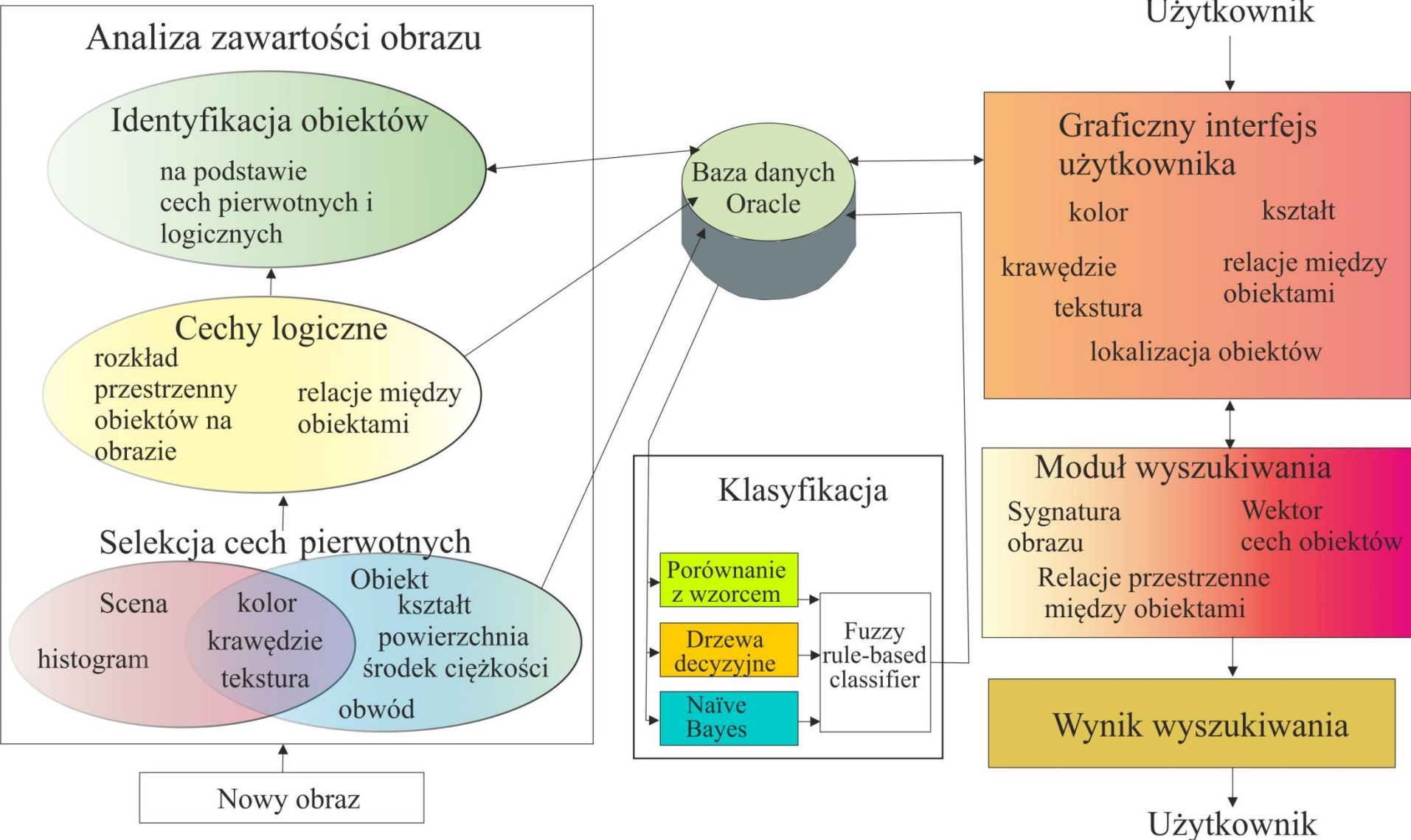
20 40 60 80 100 120

Detailed description: The image shows a software interface for object matching. It features three coordinate grids. The top-left grid, titled 'obraz elementu - pytania', shows a white outline of a house on a black background with axes from 50 to 400. The top-right grid, titled 'element dopasowywany', shows a white outline of a window on a black background with axes from 50 to 700. The bottom-left grid shows a white outline of a house on a black background with axes from 20 to 160. A 'Dopasuj elementy' button is located to the right of the bottom-left grid. Below the button, a yellow box displays the text 'Stopień niedopasowania -0.0066514'. The interface also includes a menu bar with 'File' and a toolbar with icons for zooming and panning.

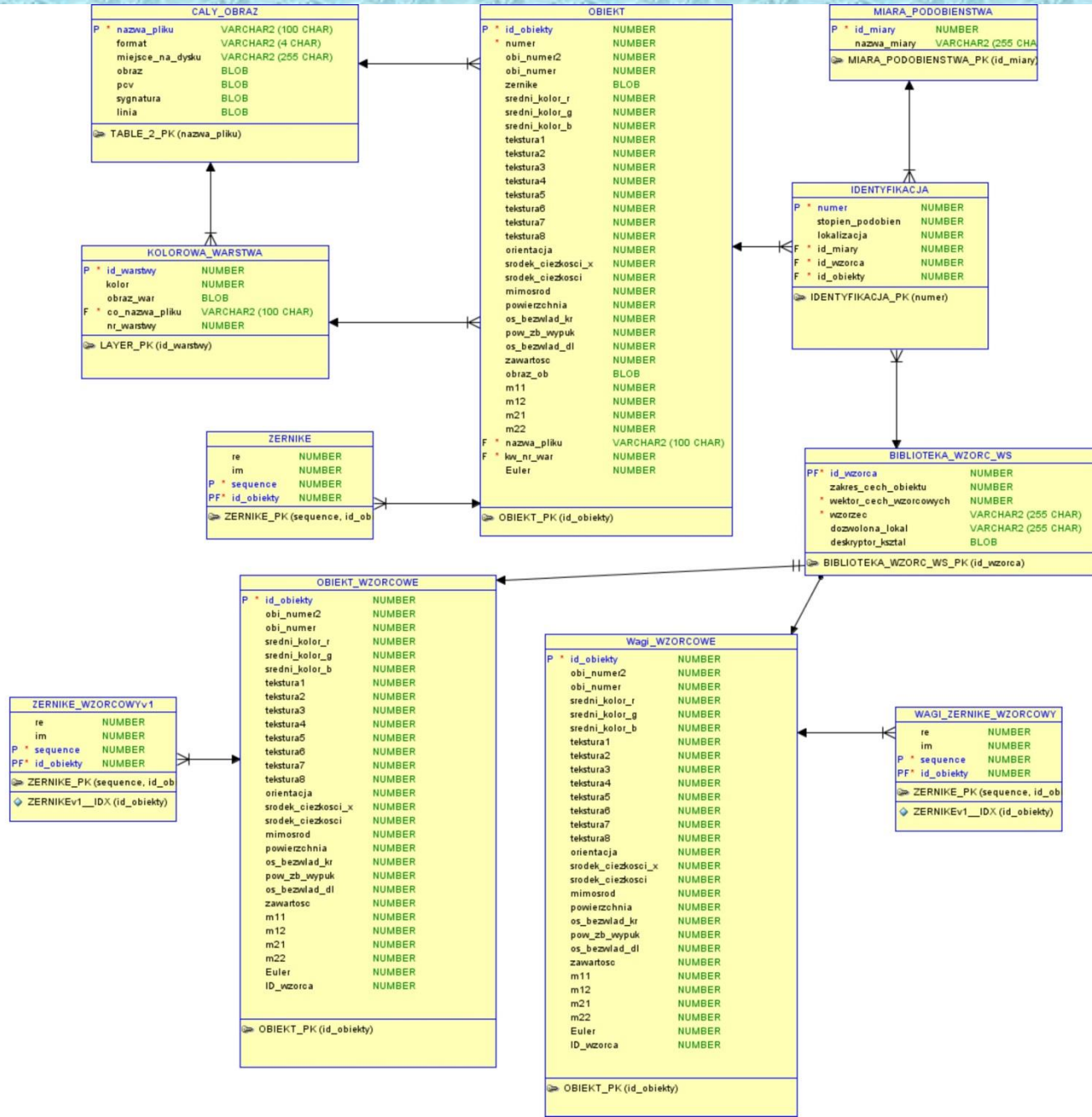
Schemat ogólny obrazowej bazy danych

Przetwarzanie wsadowe

Przetwarzanie bezpośrednie



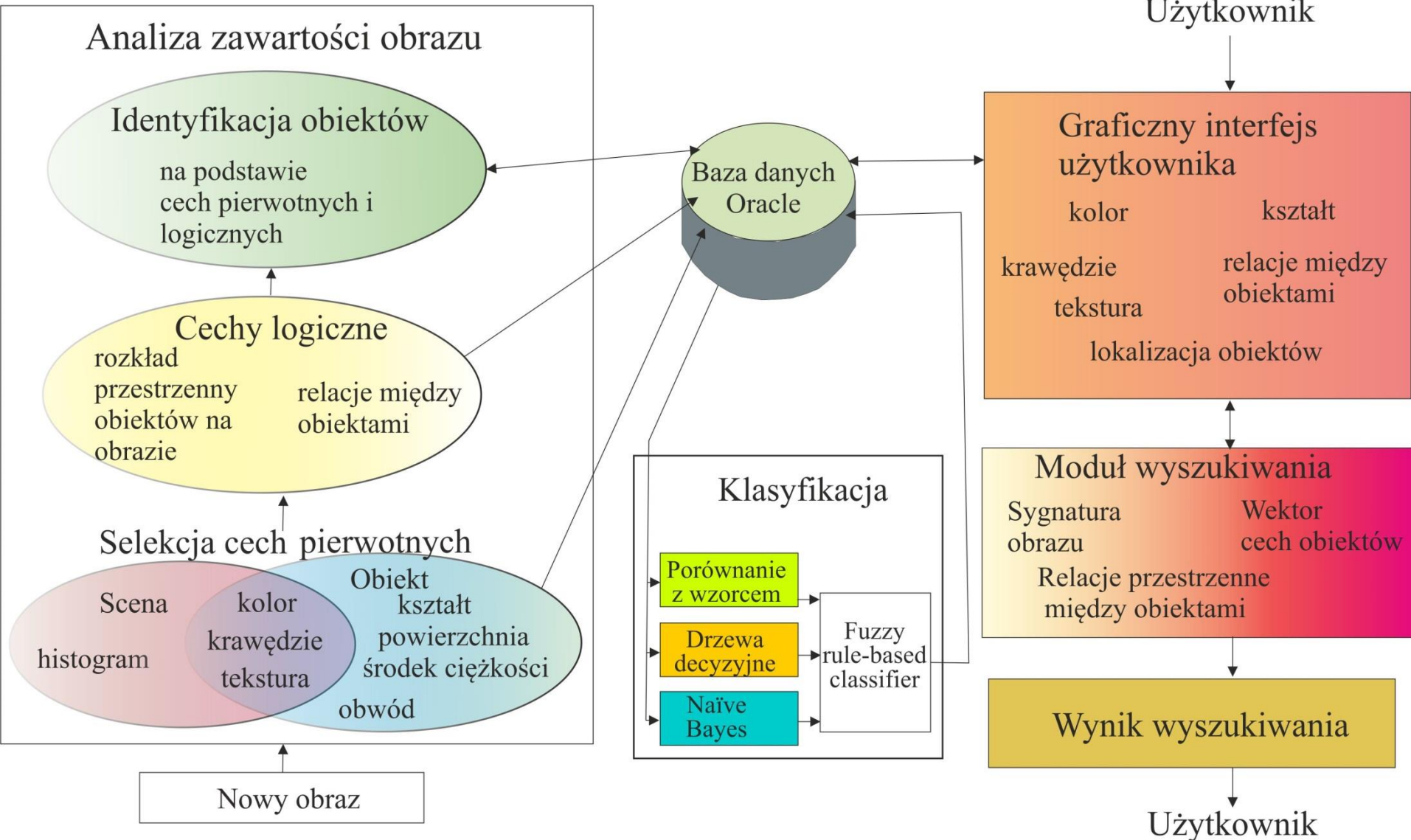
Struktura bazy danych - Oracle



Schemat ogólny obrazowej bazy danych

Przetwarzanie wsadowe

Przetwarzanie bezpośrednie



Biblioteka wzorców

Biblioteka wzorców zawiera informację o wzorcowych wektorach cech obiektu graficznego, zakresie wartości i wagach przypisanych tym wartościom. Dla każdego obiektu mamy więc wzorcowy wektor cech P_k z wagami

$$\mu_{P_k}(f_i) \in [0,1]$$

Każdy obiekt jest przyporządkowywany do określonej klasy z biblioteki wzorców poprzez szukanie najmniejszej odległości pomiędzy wektorami O i P_k

$$d(F_O, P_k) = \left[\sum_{i=1}^r \mu_{P_k}(f_i) |F_O(f_i) - P_k(f_i)|^m \right]^{1/m}$$

Na tej podstawie każdy obiekt otrzymuje etykietę L_k

Classifiers

- Similarity to patterns

$$d(O, P_k) = \left[\sum_{i=1}^r \xi_{P_k}(f_i) |O(f_i) - P_k(f_i)|^m \right]^{1/m}$$

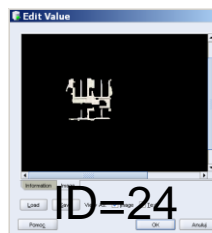
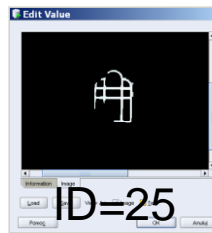
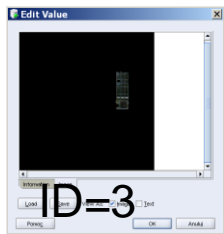
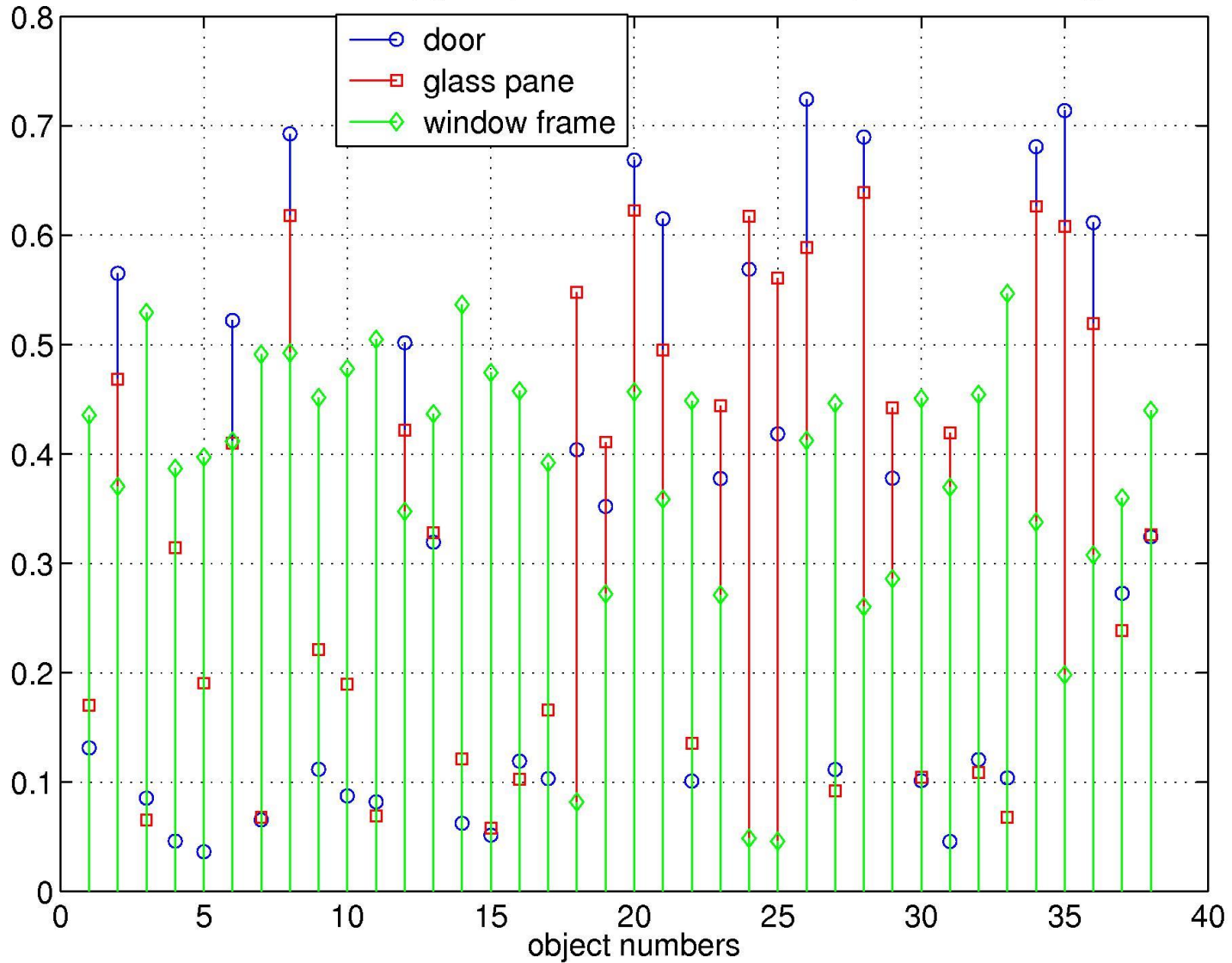
where weights satisfy $\xi_{P_k}(f_i) \in [0,1]$ and weights for Zernike's complex features satisfy

$$\xi_{CV} = \sqrt{\frac{\delta_{\text{Re}}^2 + \delta_{\text{Im}}^2}{\bar{x}_{\text{Re}}^2 + \bar{x}_{\text{Im}}^2}}$$

- Decision tree. We use the hierarchic method. A more general division is created by k-means clustering of the whole data set into four clusters. The most numerous classes of each cluster (metaclass) are assigned to four decision trees, which results in 8 classes for each one.
- Naïve Bayes classifier

Klasyfikacja obiektów

distances d for door, glass pane and window frame patterns with weights



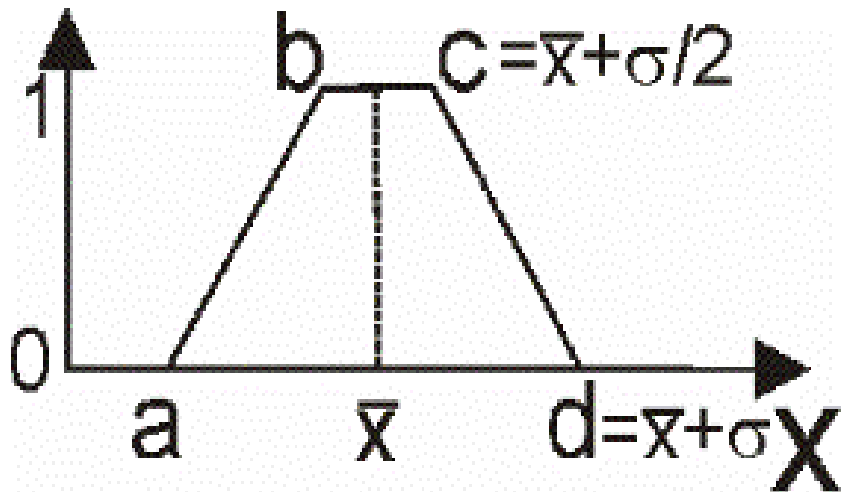
Fuzzy set

According to Zadeh, a fuzzy set F in U is uniquely specified by its membership function $\mu_F: U \rightarrow [0,1]$. Thus, the fuzzy set is described as follows:

$$F = \{(u, \mu_F(u)) | u \in U\}$$

For our purpose, we use a trapezoidal membership function which is defined by four parameters a, b, c, d :

$$\text{trap mf}(u; a, b, c, d) = \begin{cases} 0, & u < a \\ (u - a)/(b - a), & a \leq u \leq b \\ 1, & b \leq u \leq c \\ (d - u)/(d - c), & c \leq u \leq d \\ 0, & d < u \end{cases}$$



Fuzzy rule-based classifiers 1

Let us consider an M -class classification problem in an n -dimensional normalized hyper-cube $[0,1]^n$. For this problem, we use fuzzy rules of the following type:

Rule R_q : If x_1 is A_{q1} and ... and x_n is A_{qn} then Class C_q with CF_q ,

where R_q is the label of the q^{th} fuzzy rule, $\mathbf{x} = (x_1, \dots, x_n)$ is an n -dimensional feature vector, A_{qi} is an antecedent fuzzy set ($i = 1, \dots, n$), C_q is a class label, CF_q is a real number in the unit interval $[0,1]$ which represents a rule weight. The rule weight can be specified in a heuristic way. We use the n -dimensional vector $\mathbf{A}_q = (A_{q1}, \dots, A_{qn})$ to represent the antecedent part of the fuzzy rule R_q in the above formula in a concise manner.

Fuzzy rule-based classifiers 2

A set of fuzzy rules S of the type shown in the rule R_q forms a fuzzy rule-based classifier. When an n -dimensional vector $\mathbf{x}_p = (x_{p1}, \dots, x_{pn})$ is presented to S , first the *compatibility grade* of \mathbf{x}_p with the antecedent part \mathbf{A}_q of each fuzzy rule R_q in S is calculated as the product operator:

$$\mu_{A_q}(\mathbf{x}_p) = \mu_{A_{q1}}(x_{p1}) \times \dots \times \mu_{A_{qn}}(x_{pn}) \text{ for } R_q \in S,$$

where $\mu_{A_{qi}}(\cdot)$ shows the membership function of A_{qi} .

Then a single winner rule $R_{w(\mathbf{x}_p)}$ is identified for \mathbf{x}_p as

follows:
$$w(\mathbf{x}_p) = \arg \max_q \{CF_q \times \mu_{A_q}(\mathbf{x}_p) \mid R_q \in S\}$$

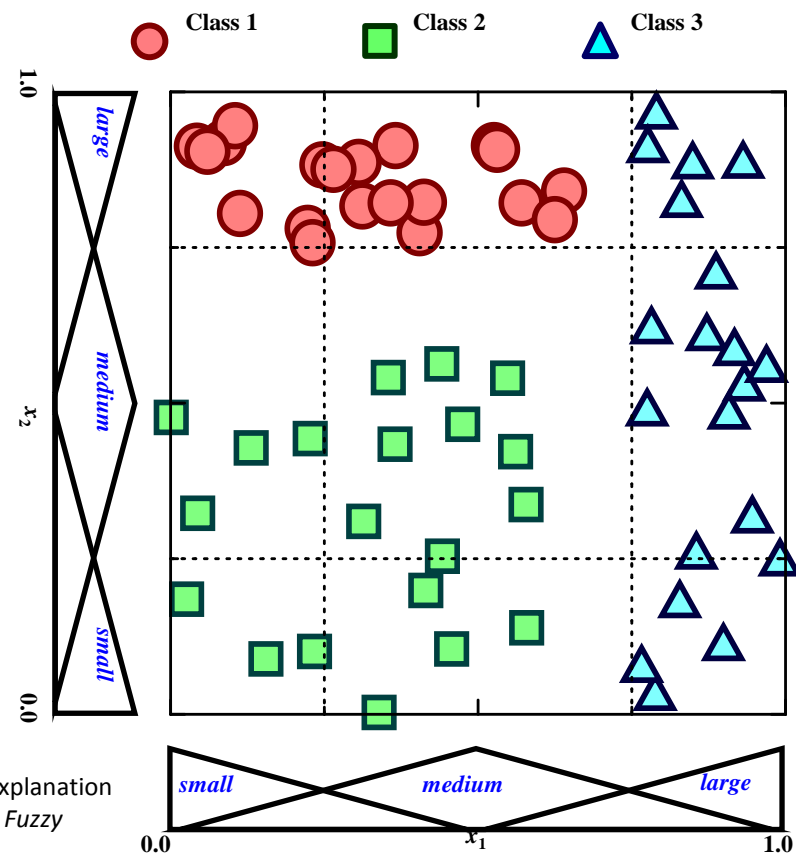
where $w(\mathbf{x}_p)$ denotes the rule index of the winner rule for \mathbf{x}_p .

The vector \mathbf{x}_p is classified by the single winner rule $R_{w(\mathbf{x}_p)}$ belonging to the respective class. We use the single winner-based fuzzy reasoning method for pattern classification.

Fuzzy rule-based classifier proposed by Ishibuchi

An ideal theoretical example of a simple three-class, two-dimensional pattern classification problem with 20 patterns from each class is considered by H. Ishibuchi and Y. Nojima.

In this example three linguistic values (*small*, *medium* and *large*) were used as antecedent fuzzy sets for each of the two attributes and 3×3 fuzzy rules were generated.



Fuzzy rule-based classifier with 9 fuzzy rules

S_1 :

R_1 : If x_1 is *small* and x_2 is *small* then Class2 with 1.0,

R_2 : If x_1 is *small* and x_2 is *medium* then Class2 with 1.0,

R_3 : If x_1 is *small* and x_2 is *large* then Class1 with 1.0,

R_4 : If x_1 is *medium* and x_2 is *small* then Class2 with 1.0,

R_5 : If x_1 is *medium* and x_2 is *medium* then Class2 with 1.0,

R_6 : If x_1 is *medium* and x_2 is *large* then Class1 with 1.0,

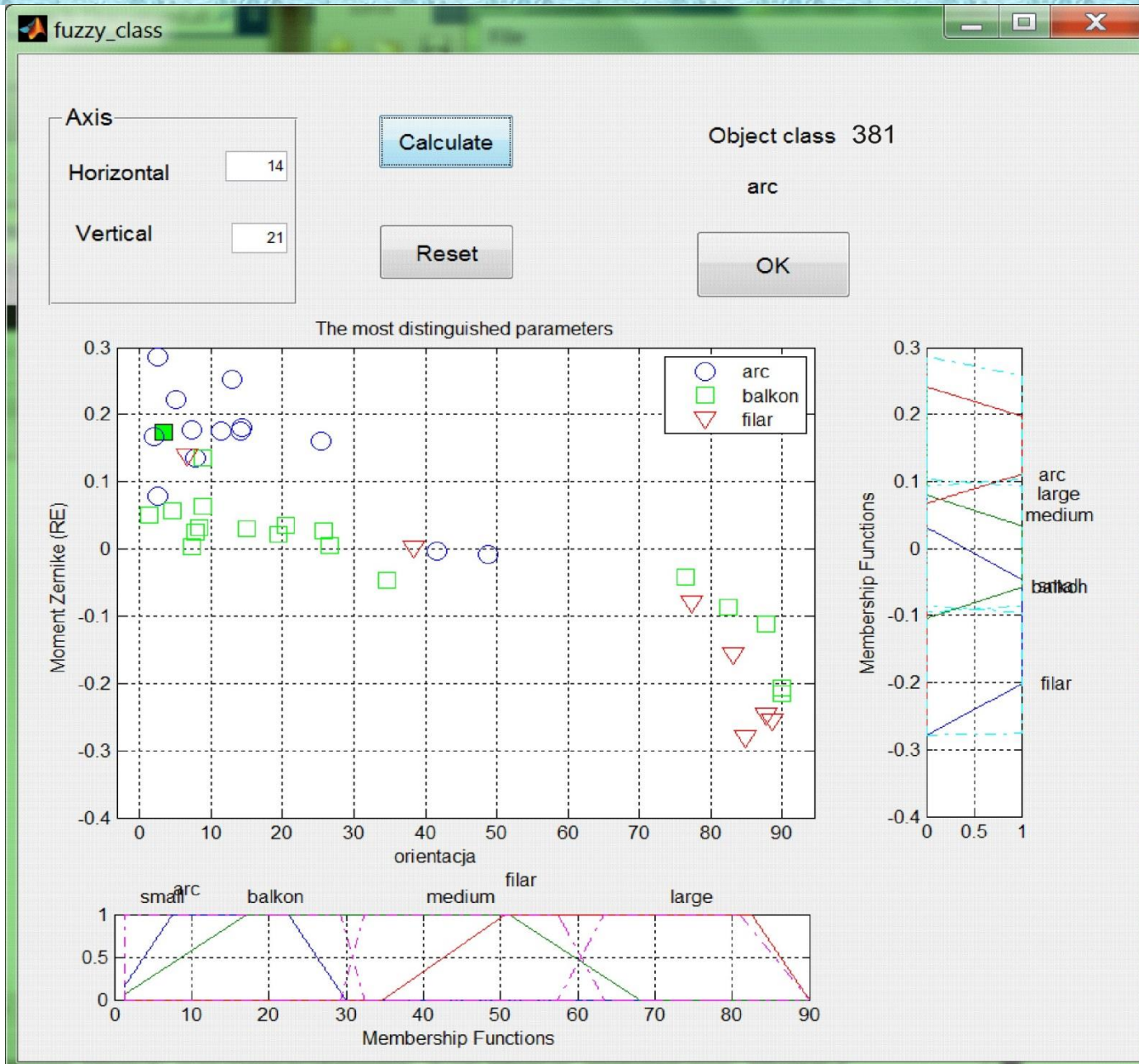
R_7 : If x_1 is *large* and x_2 is *small* then Class3 with 1.0,

R_8 : If x_1 is *large* and x_2 is *medium* then Class3 with 1.0,

R_9 : If x_1 is *large* and x_2 is *large* then Class3 with 1.0.

R_3	R_6	R_9
R_2	R_5	R_8
R_1	R_4	R_7

Three-class problem for two features



Fuzzy rule-based classifier with 9 fuzzy rules

For our fuzzy rule-based classifier we have classified data from a training subset according to the fuzzy rule-based classifier S_2 with 9 fuzzy rules constructed automatically as follows:

S_2 : fuzzy rule-based classifier with 9 fuzzy rules

R_1 : If x_1 is *small* and x_2 is *small* then non-defined with 1.0,

R_2 : If x_1 is *small* and x_2 is *medium* then balkon with 1.0,

R_3 : If x_1 is *small* and x_2 is *large* then arc with 1.0,

R_4 : If x_1 is *medium* and x_2 is *small* then non-defined with 1.0,

R_5 : If x_1 is *medium* and x_2 is *medium* then balkon with 1.0,

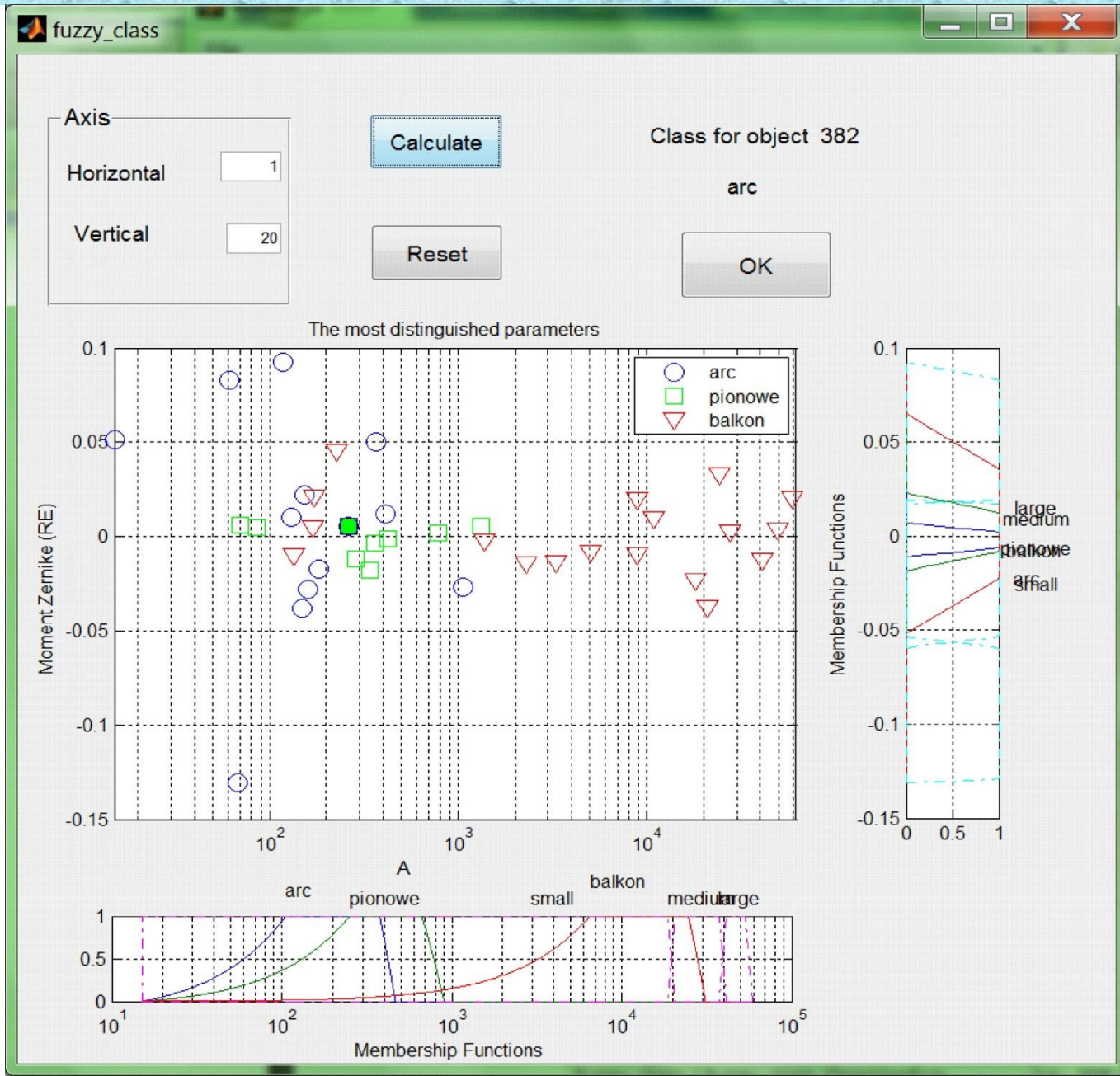
R_6 : If x_1 is *medium* and x_2 is *large* then non-defined with 1.0,

R_7 : If x_1 is *large* and x_2 is *small* then filar with 1.0,

R_8 : If x_1 is *large* and x_2 is *medium* then non-defined with 1.0,

R_9 : If x_1 is *large* and x_2 is *large* then non-defined with 1.0.

Three-class problem for two features



Fuzzy rule-based classifier with 9 fuzzy rules

For our fuzzy rule-based classifier we have classified data from a training subset according to the fuzzy rule-based classifier S_3 with 9 fuzzy rules constructed automatically as follows:

S_3 : fuzzy rule-based classifier with 9 fuzzy rules

R_1 : If x_1 is *small* and x_2 is *small* then non-defined with 1.0,

R_2 : If x_1 is *small* and x_2 is *medium* then arc with 1.0,

R_3 : If x_1 is *small* and x_2 is *large* then arc with 1.0,

R_4 : If x_1 is *medium* and x_2 is *small* then non-defined with 1.0,

R_5 : If x_1 is *medium* and x_2 is *medium* then vertical lines with 1.0,

R_6 : If x_1 is *medium* and x_2 is *large* then non-defined with 1.0,

R_7 : If x_1 is *large* and x_2 is *small* then non-defined with 1.0,

R_8 : If x_1 is *large* and x_2 is *medium* then balkon with 1.0,

R_9 : If x_1 is *large* and x_2 is *large* then non-defined with 1.0.

Axis

Horizontal

Vertical

Calculate

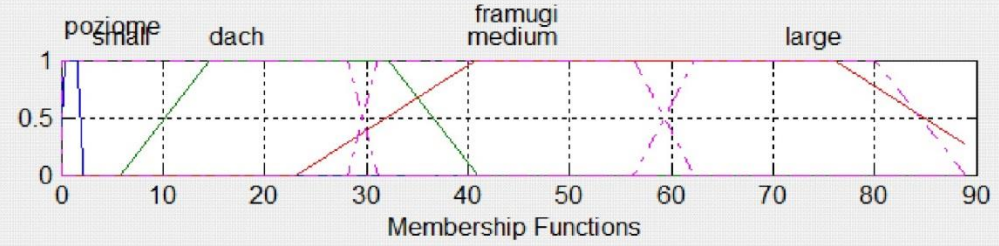
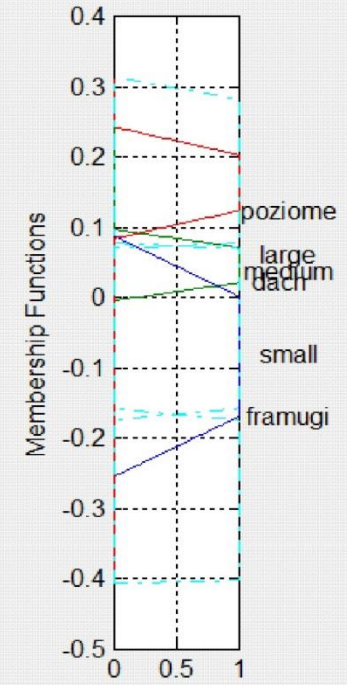
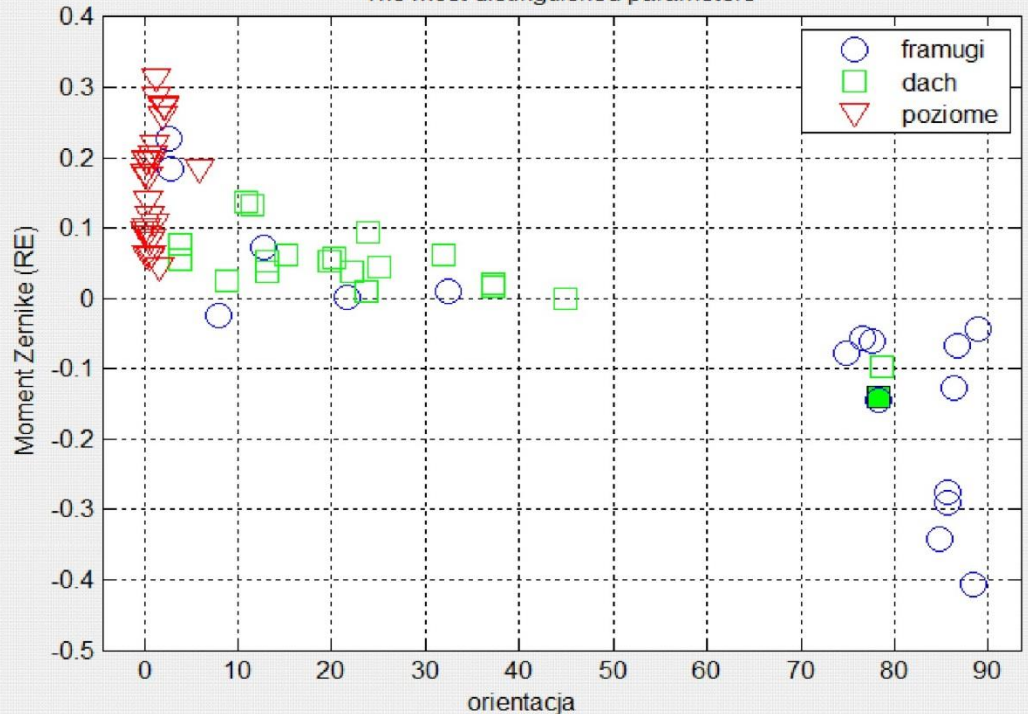
Class for object 383

framugi

Reset

OK

The most distinguished parameters



Fuzzy rule-based classifier with 9 fuzzy rules

S_4 : fuzzy rule-based classifier with nine fuzzy rules

- R_1 : If x_1 is *small* and x_2 is *small* then non-defined with 1.0,
- R_2 : If x_1 is *small* and x_2 is *medium* then roof with 1.0,
- R_3 : If x_1 is *small* and x_2 is *large* then horizontal lines with 1.0,
- R_4 : If x_1 is *medium* and x_2 is *small* then non-defined with 1.0,
- R_5 : If x_1 is *medium* and x_2 is *medium* then roof with 1.0,
- R_6 : If x_1 is *medium* and x_2 is *large* then non-defined with 1.0,
- R_7 : If x_1 is *large* and x_2 is *small* then frame with 1.0,
- R_8 : If x_1 is *large* and x_2 is *medium* then frame with 1.0,
- R_9 : If x_1 is *large* and x_2 is *large* then non-defined with 1.0.

Classification precision

Precision	Similarity to pattern	Decision trees	Naïve Bayes	FRBC
Total (for 30 classes)	21.5%	68.6%	53.8%	88%
Window-pane	16.1%	72%	31.3%	89.7%
Window	46.7%	61%	82.2%	57.6%
Brick wall	9%	45.5%	32%	90.9%
Arc	63.6%	68.2%	65%	58%
Roof edge	8.4%	86.7%	61.4%	93.9%

Opis rozkładu przestrzennego obiektów graficznych na obrazie z użyciem analizy składowych głównych 1

W bazie danych mamy M klas obiektów oznaczonych etykietami L_1, L_2, \dots, L_M . Możemy zatem określić sygnaturę obrazu I_j jako następujący wektor:

$$\text{Signature } (I_j) = [\text{nobc}_{i1}, \text{nobc}_{i2}, \dots, \text{nobc}_{iM}]$$

gdzie: nobc_{ik} oznacza liczbę obiektów o_{ij} o klasie L_k w danym obrazie I_j .

Poza tą informacją rozpatrujemy też wzajemne położenie obiektów w obrazie w oparciu o analizę składowych głównych (PCA).

Opis rozkładu przestrzennego obiektów graficznych na obrazie z użyciem analizy składowych głównych 2

Porównujemy dwa obrazy składające się ze zbioru

obiektów $I_i = \{O_{i1}, O_{i2}, \dots, O_{in}\}$

opisanych przez środki ciężkości $C_{ij} = (x_{ij}, y_{ij})$ i klasę L_{ij}

Formalnie obraz I składa się z n obiektów i k klas gdzie $k \leq N$, ponieważ kilka obiektów może być tej samej klasy.

Jako linię odniesienia do liczenia kąta pomiędzy obiektami wybieramy największą odległość między centroidami obiektów istniejącą na obrazie

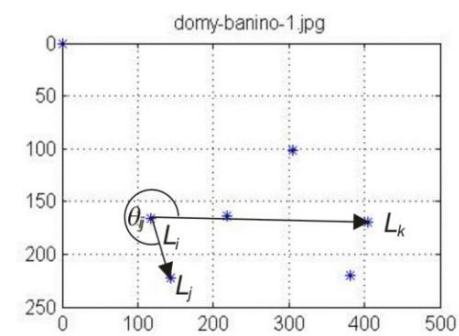
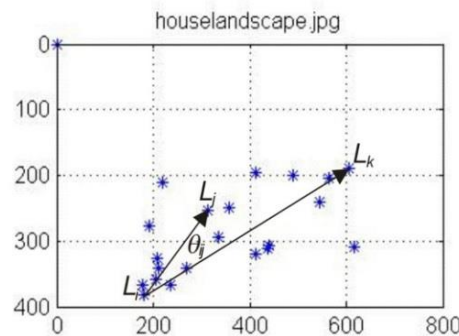
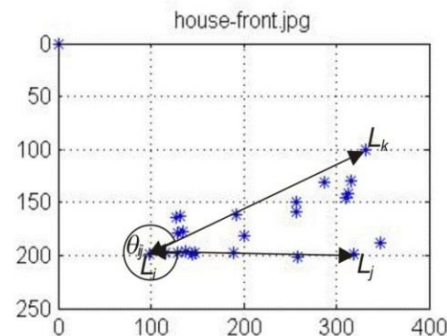
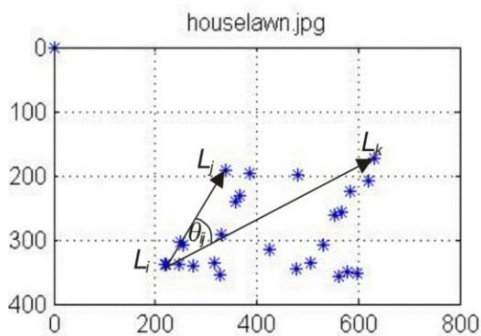
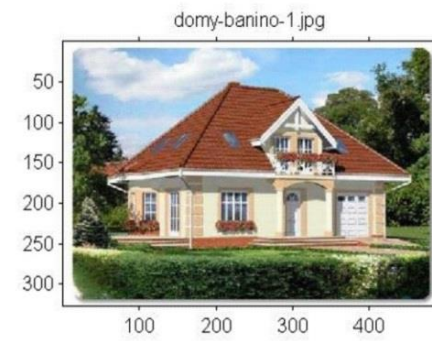
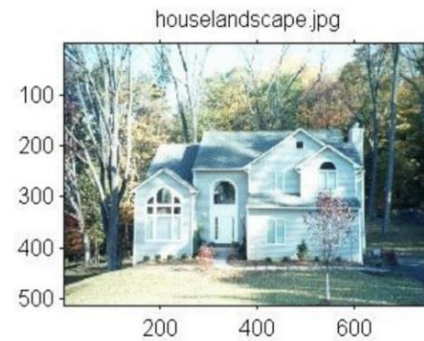
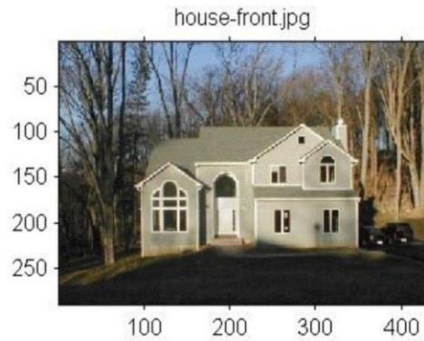
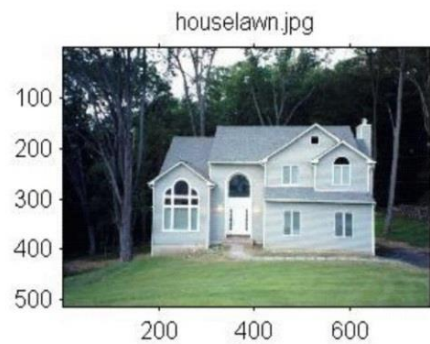
$$\text{dist}(C_p, C_q) = \max \{ \text{dist}(C_i, C_j) \mid \forall i, j \in \{1, 2, \dots, k\} \text{ and } L_i \neq L_j \}$$

Następnie od tej linii liczymy kąty θ_{ij} pomiędzy kolejnymi środkami ciężkości. W ten sposób metoda jest niezmiennicza względem obrotu.

Określenie kąta między obiektami przy wyznaczeniu macierzy do analizy składowych głównych 3

Otrzymujemy więc trójki (L_i, L_j, θ_{ij}) , gdzie kąt jest liczony względem linii odniesienia. Dostajemy $T = n(n-1)/2$ takich trójek dla obrazów o n obiektach.

Następnie poszukujemy wektorów własnych PCV_i macierzy kowariancji zgodnie z metodą analizy składowych głównych.



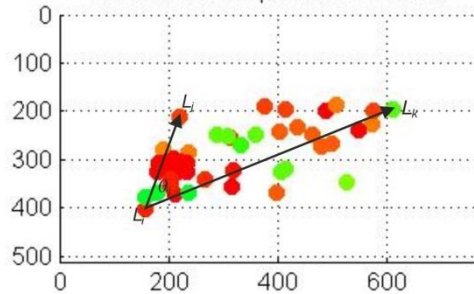
Determination of angle relative to the reference direction

housetandscape.jpg



housetandscape.jpg

object centroids
colours indicate particular classes

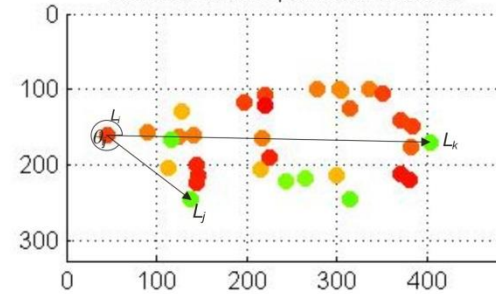


domy-banino-1.jpg

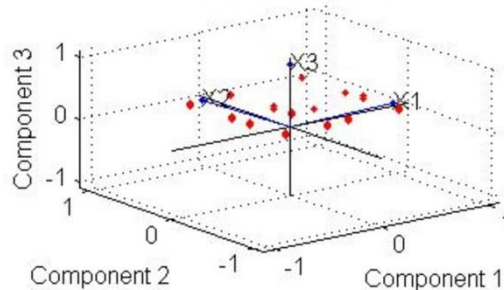


domy-banino-1.jpg

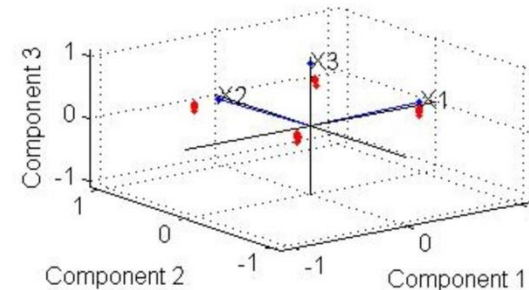
object centroids
colours indicate particular classes



three components accounting
for 98% of the variance



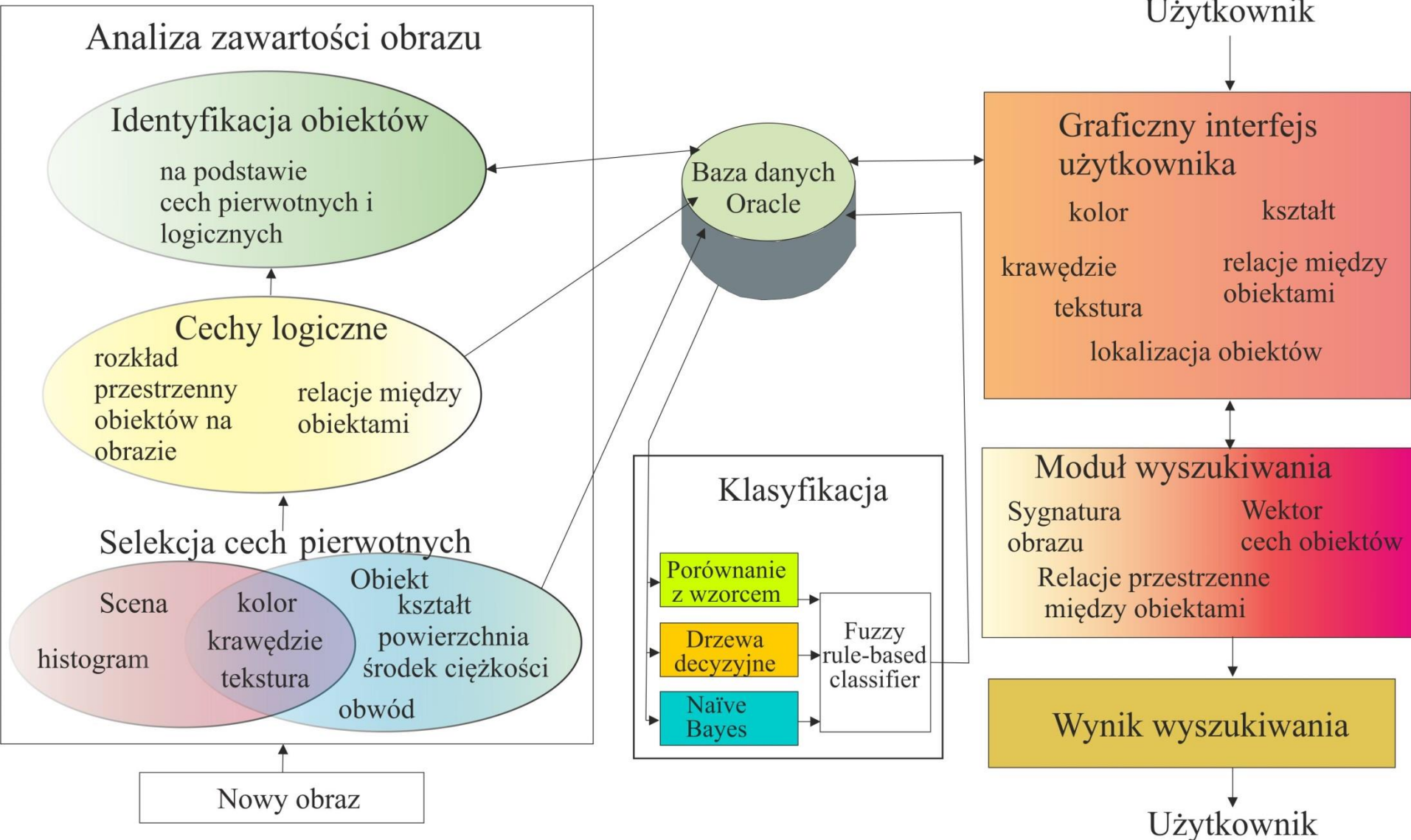
three components accounting
for 98% of the variance



Schemat ogólny obrazowej bazy danych

Przetwarzanie wsadowe

Przetwarzanie bezpośrednie



Query types

1. Query by keywords
2. Query by example
3. Query by canvas
4. Query by sketches
5. Query by spatial icons
6. Semantic query
7. Designed query for semantic retrieval.

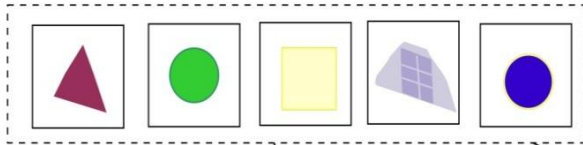
Query types

Query by keywords

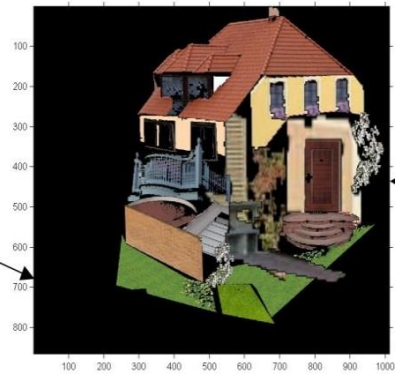
a big house



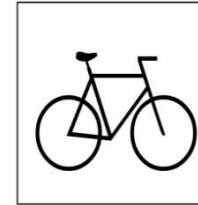
Query by example



Query by canvas



Designed query



Query by sketches



Query by spatial icons



Semantic query

Graficzne zapytanie do bazy poprzez GUI

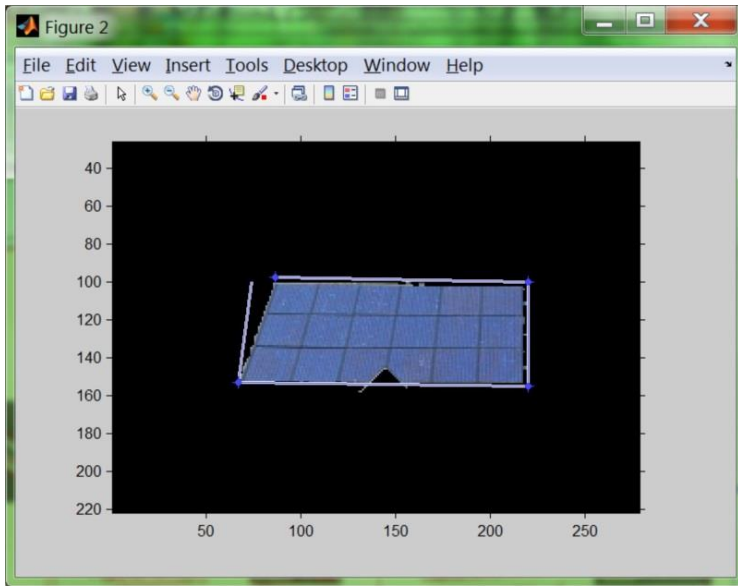
The image displays a graphical user interface (GUI) for image processing, specifically for color selection and edge detection. The interface is divided into several sections:

- Top Left:** A small window titled "Query_menu" with the text "Wybór elementu Kolor elementu". It contains a small image of a house with a coordinate grid (x: 0-800, y: 0-600) and a "Button Group" with a "wybierz" button.
- Top Center:** A window titled "Zmiana_kolor" with a "Wybrany obraz" dropdown menu and a "Zmień kolor obrazu" button. It displays the selected image (a house) with a coordinate grid.
- Top Right:** A "Paleta Kolorów" (Color Palette) with a grid of color swatches. Below the palette, the RGB values for the selected color are shown: R: 0.10588, G: 0.3098, B: 0.20784. A "More Colors ..." button is also present.
- Bottom Left:** A "Color" dialog box with a grid of color swatches, a "More Colors..." button, and "OK" and "Cancel" buttons.
- Bottom Center:** A large image of the house with a red edge detection overlay. The image has a coordinate grid (x: 0-900, y: 0-600).

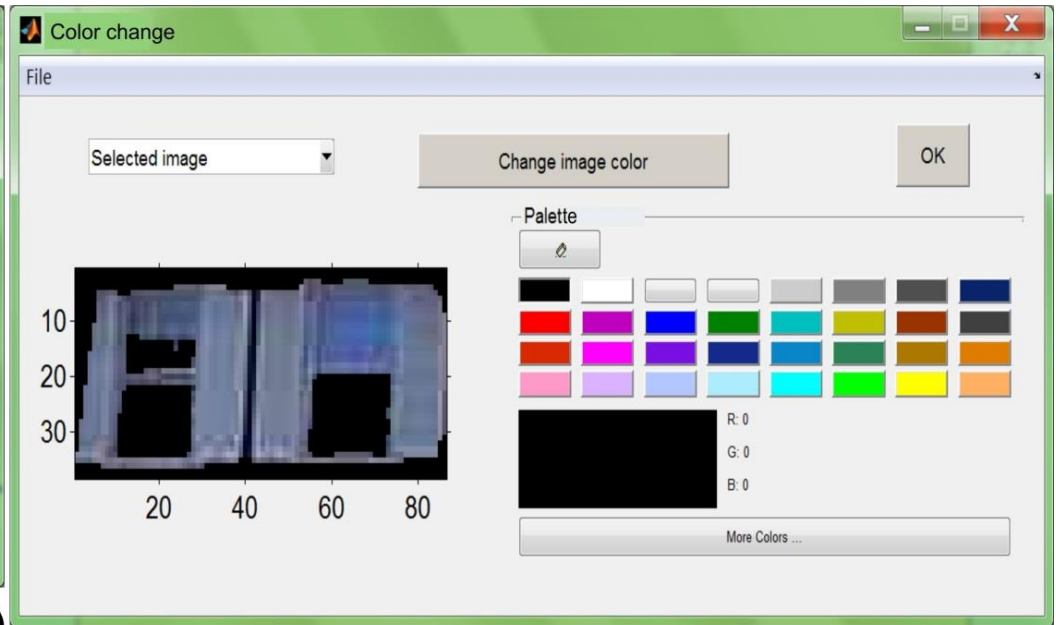
Graficzne zapytanie do bazy poprzez GUI

The image displays a MATLAB 7.8.0 (R2009a) environment with a custom GUI. The main window, titled 'Query_menu', features a central plot of a house image with red contour lines. The plot axes range from 0 to 900 on the x-axis and 0 to -600 on the y-axis. To the left of the plot is a menu titled 'Wybór elementu' with options: 'kontur domu', 'dachy', 'okna', 'balkony', 'drzwi', and 'Close'. Below the menu is a 'Push Button'. On the far left, a yellow sidebar contains five image thumbnails, each with a 'wybierz' button. The thumbnails show different parts of the house: a roof with chimneys, a red-tiled roof, a close-up of a roof, a close-up of a roof, and another red-tiled roof. At the bottom of the sidebar is an orange button labeled 'Inne'. The MATLAB interface includes a Command Window with the text 'New to MATLAB? Watch this Video' and a 'zmiana_kolor' window with a 'Zmień kolor obrazu' button and a color palette.

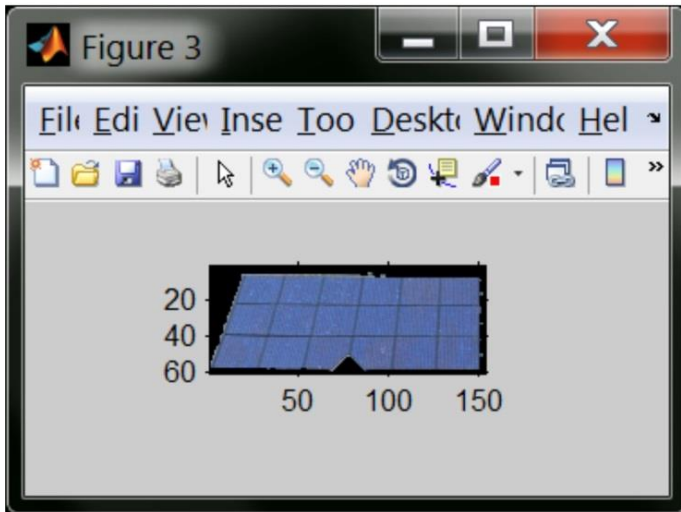
Graficzny interfejs użytkownika



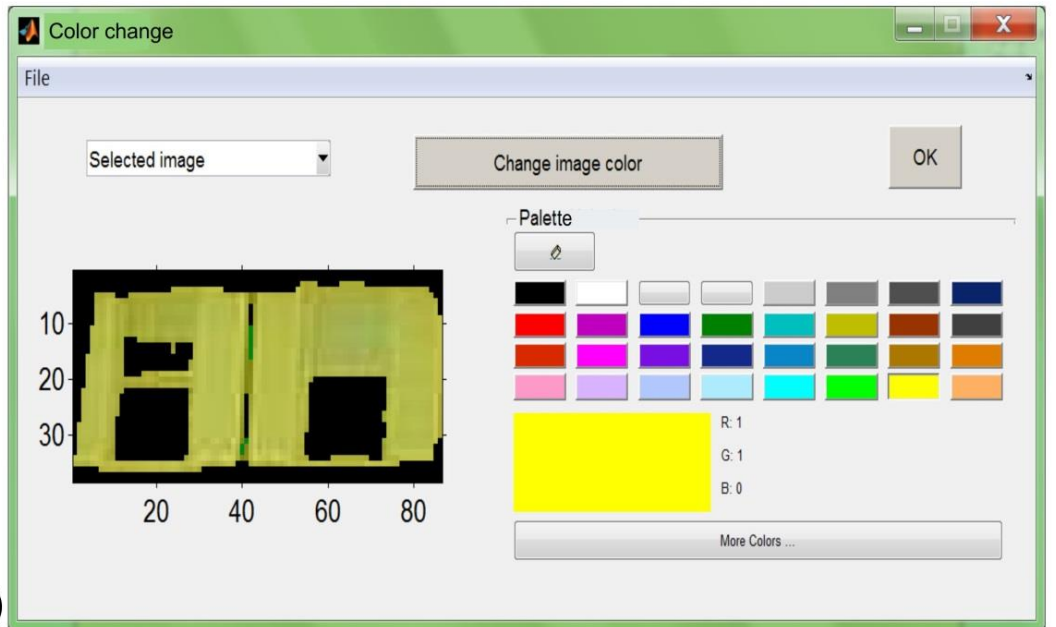
a)



c)



b)

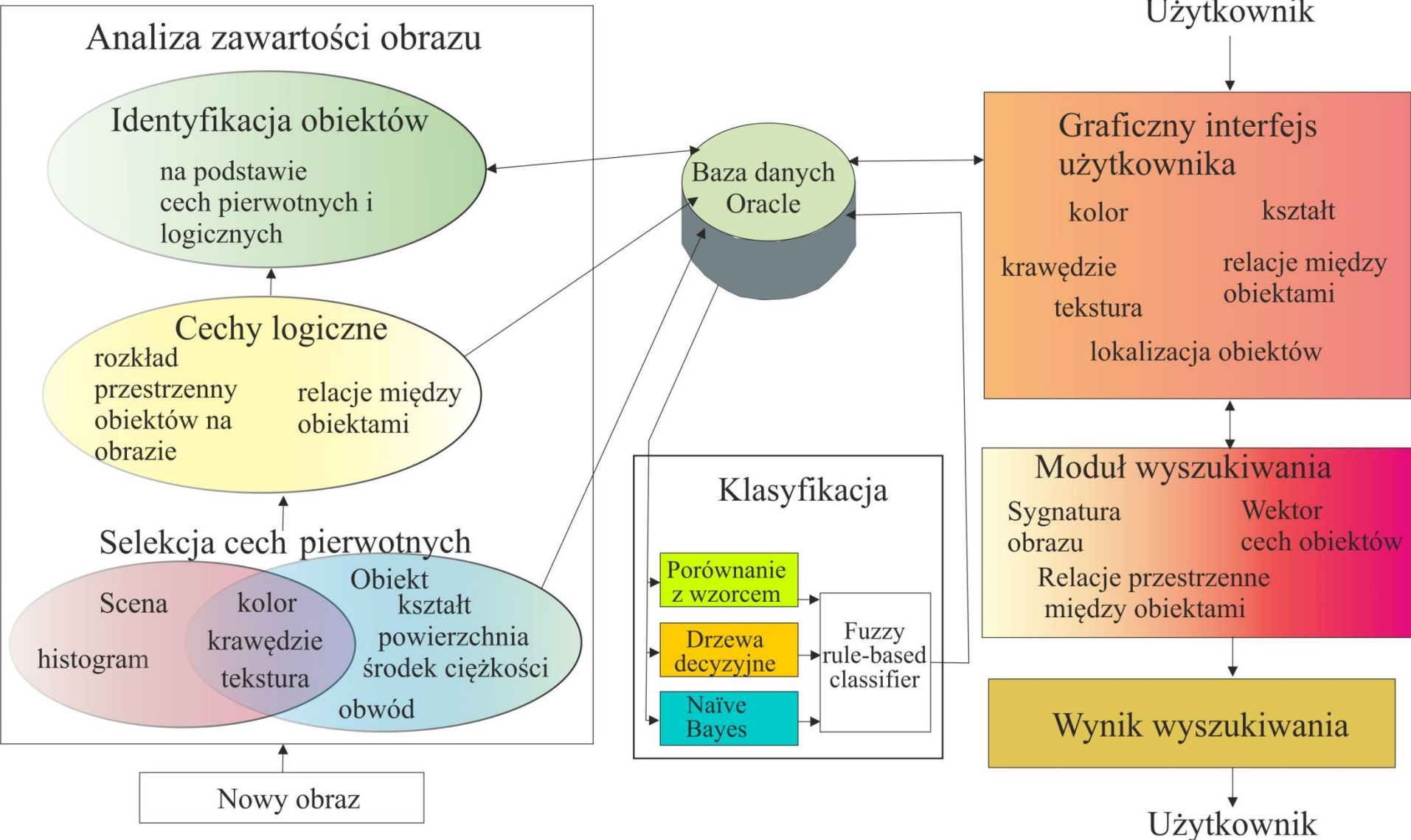


d)

Schemat ogólny obrazowej bazy danych

Przetwarzanie wsadowe

Przetwarzanie bezpośrednie



Silnik wyszukiwania I etap

Obraz z bazy możemy oznaczyć jako $I^b = \{o_1^b, o_2^b, \dots, o_m^b\}$.
Niech zapytaniem będzie obraz I^q taki, że $I^q = \{o_1^q, o_2^q, \dots, o_n^q\}$.
Aby odpowiedzieć na to zapytanie porównujemy z nim każdy obraz I^b z bazy w następujący sposób.

Najpierw określamy podobieństwo sim_{sgn} pomiędzy I^q i I^b licząc odległość między sygnaturami obrazów:

$$\text{sim}_{\text{sgn}}(I_q, I_b) = \sum_i (\text{nob}_{qi} - \text{nob}_{bi})$$

Wtedy odległość tych wektorów można traktować jako zmodyfikowaną odległość Hamminga, zakładając, że $\text{sim}_{\text{sgn}} \geq 0$ i $\max_i (\text{nob}_{qi} - \text{nob}_{bi}) \leq \text{tr}$, gdzie tr jest różnicą obiektów danej klasy, którą dopuszczamy pomiędzy wektorami. Jeśli to podobieństwo jest mniejsze od założonego progu wtedy obraz I^b jest odrzucany z dalszych porównań.

Silnik wyszukiwania II etap

Jeśli to podobieństwo jest większe, to przechodzimy do następnego etapu i znajdujemy podobieństwo sim_{PCV} między rozkładem przestrzennym obiektów w obrazach I^q i I^b licząc odległość euklidesową między wektorami PCV:

$$\text{sim}_{\text{PCV}}(I^q, I^b) = 1 - \sqrt{\sum_{i=1}^3 (\text{PCV}_i^b - \text{PCV}_i^q)^2}$$

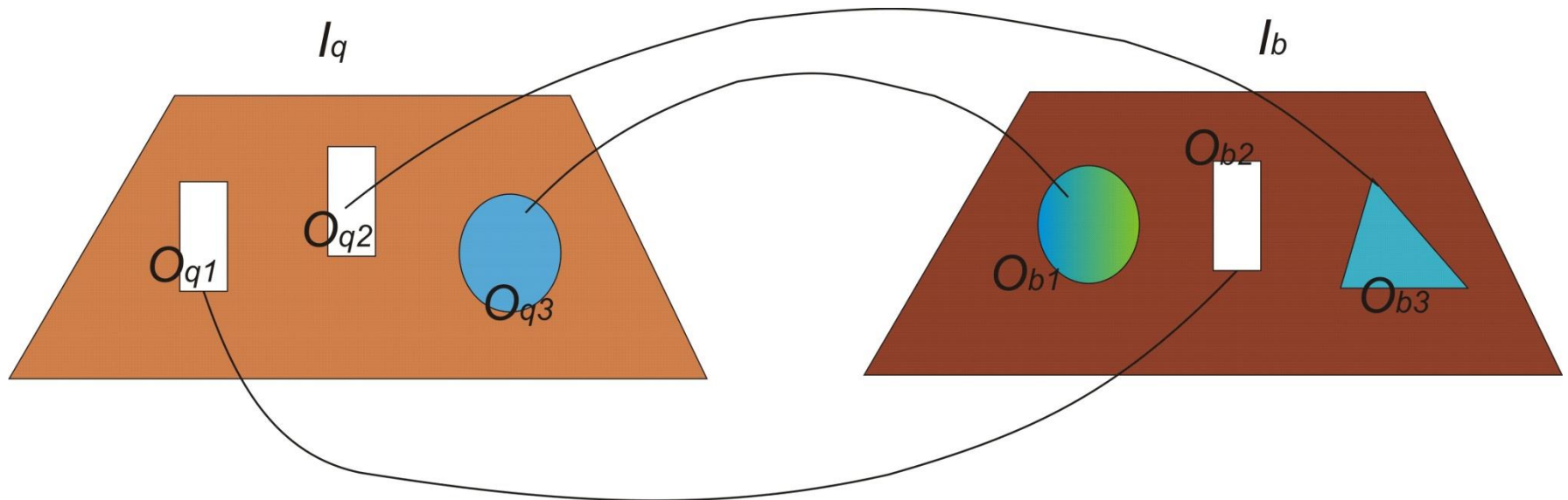
Jeśli z kolei to podobieństwo okaże się mniejsze od wartości progowej obraz I^b jest odrzucany. W przeciwnym razie, przechodzimy do następnego kroku, w którym liczymy podobieństwo poszczególnych obiektów wchodzących w skład obu obrazów I^q i I^b . Dla każdego obiektu o_i^q wchodzącego w skład obrazu I^q , znajdujemy najbardziej podobny obiekt w tej samej klasie na obrazie I^b .

Dopasowanie parami z eliminacją

Dla każdego obiektu $o_j^b \in L_i^q$ można obliczyć $\text{sim}_{\text{ob}}(o_i^q, o_j^b)$ między obiektami:

$$\text{sim}_{\text{ob}}(o_i^q, o_j^b) = 1 - \sqrt{\sum_l (F^q o_{il} - F^b o_{jl})^2}$$

gdzie: l jest numerem cechy w wektorze cech F_o poszczególnych obiektów. Może zajść jednak sytuacja przedstawiona na rysunku:



Silnik wyszukiwania III etap

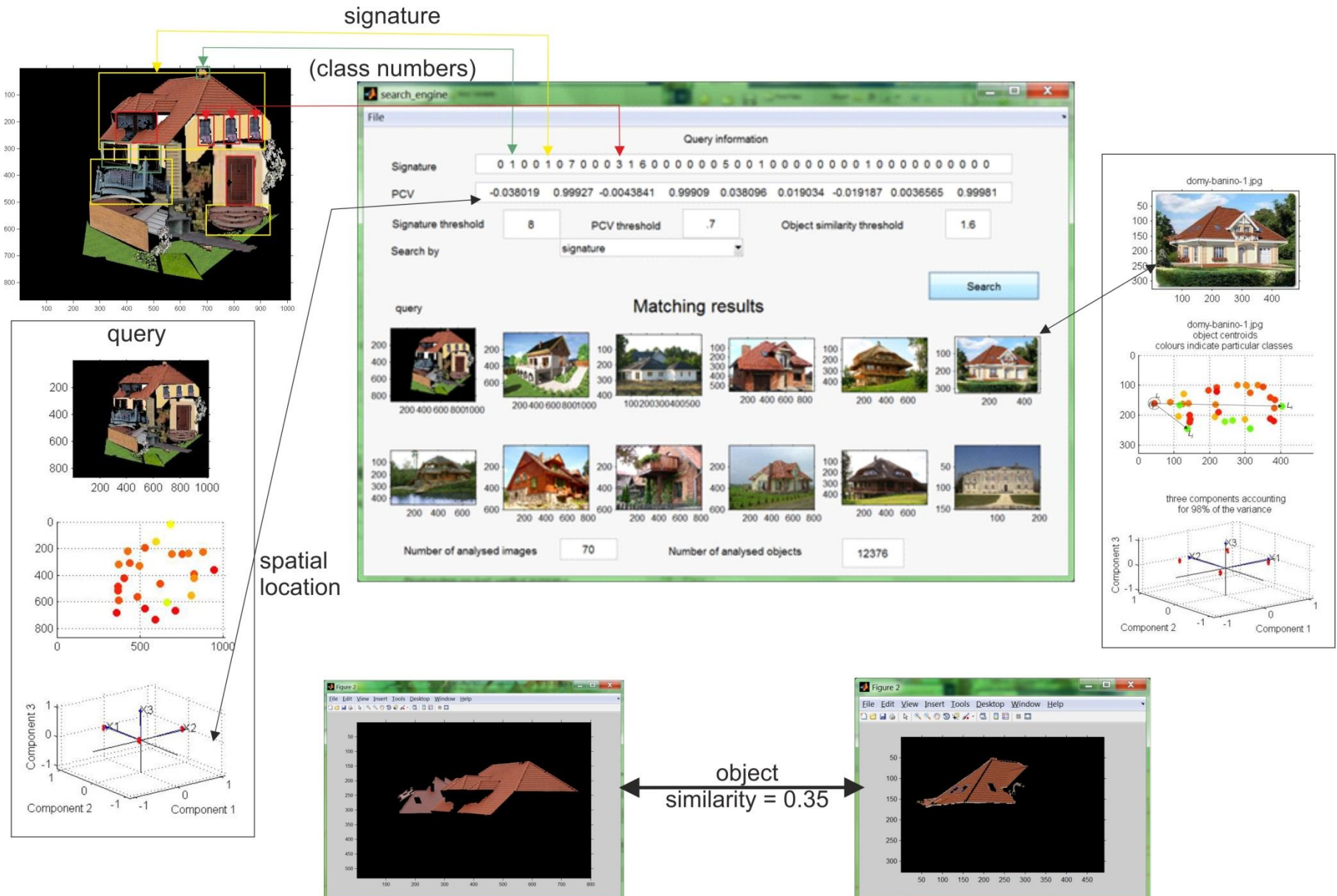
Z powyższego algorytmu dostajemy wektor podobieństw pomiędzy zapytaniem I^q i obrazem I^b :

$$\text{sim}(I^q, I^b) = \begin{bmatrix} \text{sim}_{\text{ob}}(o_1^q, o_1^b) \\ \vdots \\ \text{sim}_{\text{ob}}(o_n^q, o_n^b) \end{bmatrix}$$

gdzie: n jest liczbą obiektów reprezentujących obraz I^q .

Porządkujemy obrazy I^b względem zapytania I^q , na podstawie $\text{sim}_{\text{ob}}(o_i^q, o_j^b)$

Graficzny schemat działania silnika wyszukiwania



Wyniki

search_engine

File

Query information

Signature

1 2 1 1 1 0 11 2 0 6 0 0 6 0 0 0 0 1 1 1 0 0 0 0 0 2 0 0 0 0 3 1 0 0 0

PCV

0.98833 -0.15232 0.00059419 0.15232 0.98833 0.002044 -0.00089859 -0.0019296 1

Signature threshold

20

PCV threshold

4

Object similarity threshold

.9

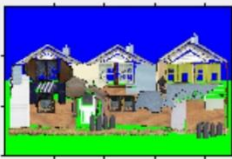











Search by

signature

Matching results

Search

query

 <p>200 400 600</p> <p>200 400 600 800</p>	<p>ssim Index is 0.1399</p>  <p>100 200 300</p> <p>200 400</p>	<p>ssim Index is 0.0012</p>  <p>100 200 300 400 500</p> <p>200 400 600</p>	<p>ssim Index is 0.2009</p>  <p>100 200 300 400</p> <p>100200300400500</p>	<p>ssim Index is 0.2221</p>  <p>200 400 600</p> <p>200 400 600 800 1000</p>	<p>ssim Index is 0.0642</p>  <p>100 200 300</p> <p>200 400</p>
<p>ssim Index is 0.1443</p>  <p>100 200 300</p> <p>100200300400500</p>	<p>ssim Index is 0.1505</p>  <p>100 200 300 400</p> <p>200 400 600</p>	<p>ssim Index is -0.0378</p>  <p>100 200 300 400</p> <p>200 400 600</p>	<p>ssim Index is 0.0833</p>  <p>100 200 300 400</p> <p>200 400 600</p>	<p>ssim Index is 0.0749</p>  <p>100 200 300 400 500</p> <p>200 400 600</p>	<p>ssim Index is 0.1745</p>  <p>50 100 150</p> <p>50 100 150 200 250</p>

Number of analysed images

70

Number of analysed objects

20664

Wyniki

search_engine

File

Query information

Signature

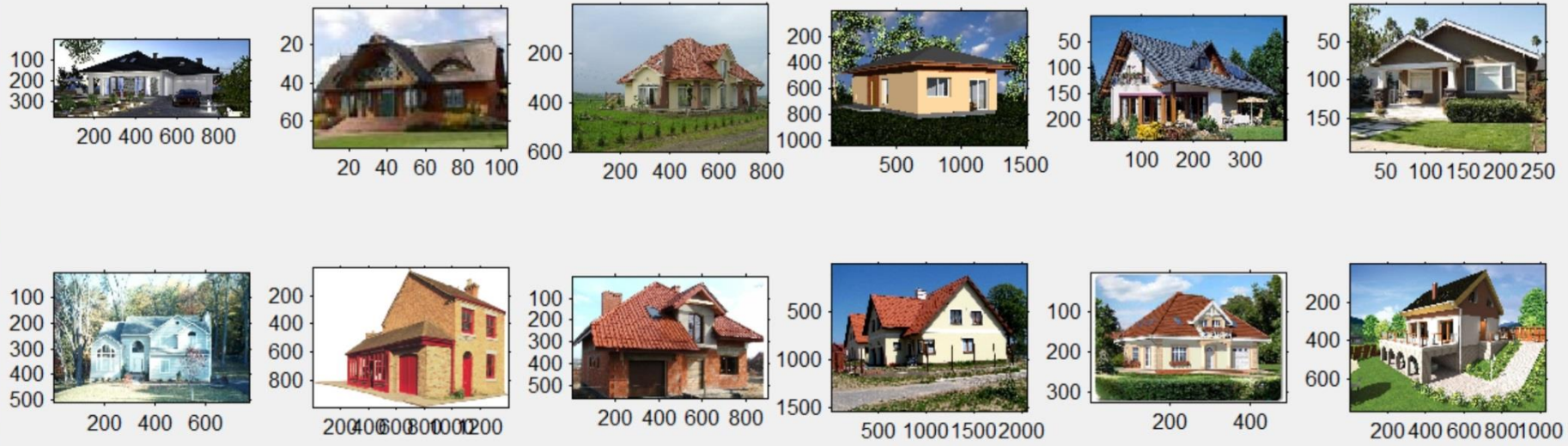
PCV

Signature threshold PCV threshold Object similarity threshold

Search by

query

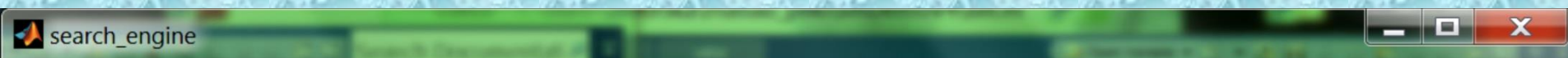
Matching results



Number of analysed images

Number of analysed objects

Wyniki



File

Query information

Signature

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

PCV

0.08686 0.99612 0.013823 0.99601 -0.086548 -0.021784 0.020504 -0.01566 0.99967

Signature threshold

6

PCV threshold

3

Object similarity threshold

1

Search by

signature

Matching results

Search

query



5001000150020002500

ssim Index is 0.3898



100
200
300
400

200 400 600

ssim Index is 0.1836



500
1000
1500

500 1000 1500 2000

ssim Index is 0.2217



50
100
150

50 100 150 200 250

ssim Index is 0.0493



100
200
300
400
500

200 400 600

ssim Index is 0.2230



200
400
600
800

200 400 600

ssim Index is 0.3072



100
200
300
400
500

200 400 600 800

ssim Index is 0.2168



200
400
600

200 400 600 800

ssim Index is 0.4157



500
1000
1500

500 1000 1500 2000 2500

ssim Index is 0.2019



100
200
300

200 400

ssim Index is 0.1019



200
400
600
800
1000

500 1000 1500

ssim Index is 0.3247



50
100
150

50 100 150 200 250

Number of analysed images

70

Number of analysed objects

6930

Wyniki

search_engine

File

Query information

Signature: 0 1 0 0 1 0 7 0 0 0 3 1 6 0 0 0 0 0 0 5 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0

PCV: -0.038019 0.99927 -0.0043841 0.99909 0.038096 0.019034 -0.019187 0.0036565 0.99981

Signature threshold: 8 PCV threshold: .7 Object similarity threshold: 1.6

Search by: signature

Search

query

Matching results

Number of analysed images: 70 Number of analysed objects: 12376

Wyniki

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

flesh and blood - Idioms b... x

Transactions on Image Processi... x

Szukaj w Google x +

https://www.google.pl/search?tbs=sbi:AMhZZiv3ti36Riff6kF6j-gDPftNoq;

Szukaj

Często odwiedzane Pasek zakładek Ang Szkoła Praca fr Info konfer



Rozmiar obrazu:
688 × 625

Nie znaleziono innych rozmiarów tego obrazu.

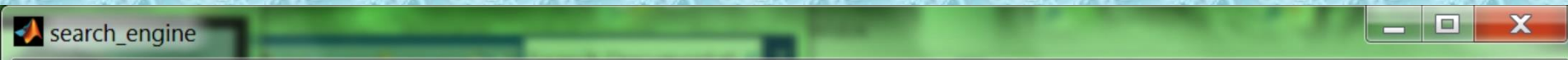
Wskazówka: wprowadź opisowe słowo w polu wyszukiwania.

[Obrazy podobne wizualnie](#)

[Zgłoś grafiki](#)



Wyniki



File

Query information

Signature

4 1 0 0 1 0 10 0 0 0 3 1 6 0 0 0 0 0 0 4 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0

PCV

-0.020557 0.99975 -0.0088261 0.99978 0.02052 -0.004195 0.0040129 0.0089104 0.99995

Signature threshold

10

PCV threshold

4

Object similarity threshold

1

Search by

signature

Matching results

Search

query



200
400
600
800
200 400 600

corr2D = -0.0311



100
200
300
400
500
200 400 600 800

corr2D = -0.0628



200
400
600
800
200 400 600 800

corr2D = -0.0167



200
400
600
800
200 400 600 800

corr2D = -0.0127



100
200
300
400
500
200 400 600

corr2D = 0.1125



500
1000
1500
2000
500 1000 1500 2000

corr2D = 0.0566



200
400
600
800
1000
200 400 600 800 1000

corr2D = 0.0915



100
200
300
400
200 400 600

corr2D = 0.0359



200
400
600
800
200 400 600 800

corr2D = -0.0039



100
200
300
400
100 200 300 400 500

corr2D = 0.1250



100
200
300
400
100 200 300 400 500

corr2D = 0.0237



100
200
300
200 400

Number of analysed images

69

Number of analysed objects

11968

Wyniki

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

Match Up Results x IEEE Signal Processing Soc... x IEEE Signal Processing Soc... x W Scale-invariant feature... x Szukaj w Google x

https://www.google.pl/search?tbs=sbi:AMhZZish2tT1z- Szukaj

Często odwiedzane Pasek zakładek Ang Szkoła Praca fr info konfer



Rozmiar obrazu:
795 × 856

Nie znaleziono innych rozmiarów tego obrazu.

Wskazówka: wprowadź opisowe słowo w polu wyszukiwania.

Obrazy podobne wizualnie

Zgłoś grafiki



Proponowane zagadnienia

- Implementacja miękkich metod wyszukiwania
- Przetestowanie różnych miar podobieństwa
- Większe uwzględnienie semantyki
 - Dołożenie ontologii
- Opracowanie obiektywnych metod oceny semantycznego podobieństwa obrazów
 - (the universal similarity index SSIM nie mierzy tego co potrzeba)
- Połączenie systemu z internetem

Literatura

- Russ J. C.: *The Image Processing Handbook*, wyd. 2, CRC Press, 1994.
- *Artificial intelligence for Maximizing Content Based Image Retrieval*, red. Zongmin Ma, Information Science Reference, NY, 2009.
- Deb S.: *Multimedia Systems and Content-Based Image retrieval*, Idea Group Publishing, Melbourne, 2004.
- Jaworska T., *Object extraction as a basic process for content-based image retrieval (CBIR) system*, Opto-Electronics Review, Vol. 15, Nr. 4, str. 184-195, 2007.
- Jaworska T., *A Search-Engine Concept Based on Multi-Feature Vectors and Spatial Relationship*, LECTURE NOTES IN ARTIFICIAL INTELLIGENCE nr 7022, str. 137–148, 2011.
- Jaworska T., *Query techniques for CBIR*, W: ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTERING nr 400, Flexible Query Answering Systems 2015, red. T. Andreassen, H. Christiansen, J. Kacprzyk, H. Larsen, G. Pasi, O. Pivert, G. De Tre, M. A. Vila, A. Yazici, S. Zadrozny, str. 403-416, Springer, 2015.