

Segmentacja obrazu

Krzysztof Krawiec

Przetwarzanie i Rozpoznawanie Obrazów

April 20, 2016

Definicja: Identyfikowanie¹ obszarów² obrazu, których wygląd jest jednolity³ dla obserwatora.

Komentarze:

- 1 Jednoznaczne ustalenie, do którego obszaru należy dany punkt obrazu,
- 2 Lub innych elementów (składowych pierwotnych), np. krawędzi,
- 3 Jednolitość może być definiowana na różne sposoby (poziom jasności, charakterystyka kolorystyczna, tekstura, etc.)

Obszar = zbiór punktów obrazu R pozostających w relacji spójności, tj. takich że $\forall p_1, p_2 \in R$ istnieje ścieżka p łącząca p_1 z p_2 taka że $p \in R$.



<https://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

Segmentacja to podział obszaru (obrazu) R na rozłączne obszary $R_1, R_2, \dots, R_N$ takie, że:

- ❶ $R_1, R_2, \dots, R_N$ tworzą *partycję* (podział) zbioru R , tj.
 - ❶ $\bigcup_{i=1}^N R_i = R$,
 - ❷ Obszary są parami rozłączne: $i \neq j \implies R_i \cap R_j = \emptyset$
- ❷ Obszary R_i są spójne,
- ❸ Każdy obszar powinien spełniać pewien test jednolitości (predykat logiczny P),
- ❹ Dla żadnej pary przyległych obszarów, ich suma nie może spełniać P .

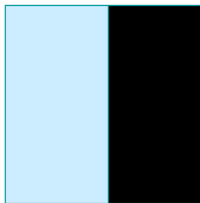
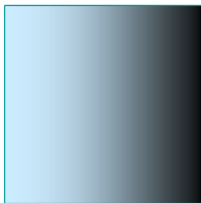
Uwaga:

- Wiele algorytmów definiuje partycję implícite, najczęściej za pomocą wartości pikseli.
- Wiele prostych algorytmów segmentacji nie definiuje jawnie P .

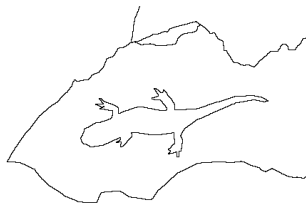
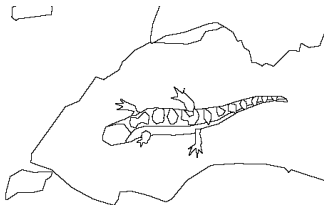
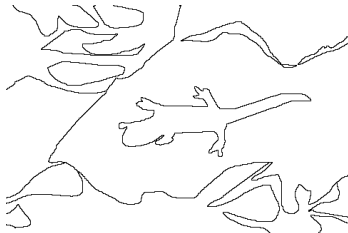
- Przejście od ilościowej do jakościowej informacji (reprezentacji obrazu)
- Segmentacja jest często zależna od zastosowania (specjalizowana)
 - Konieczność uwzględnienia wiedzy
 - zdroworoządkowej (ang. common-sense knowledge),
 - dziedzinowej (ang. domain)
 - Algorytmy można podzielić na:
 - uniwersalne,
 - dedykowane (np. analiza 3D scen rzeczywistych).
- W ogólności problem źle sformułowany (ill-posed) z matematycznego punktu widzenia.

W rzeczywistych obrazach często występują regiony, które, choć z semantycznego punktu widzenia nie powinny ulegać podziałowi na 'segmenty', ulegają mu ze względu na zmienną charakterystykę, wynikającą np. z:

- gradientu oświetlenia,
- zniekształceń perspektywicznych,



Wyniki segmentacji ręcznej (trzy niezależne osoby)



<https://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

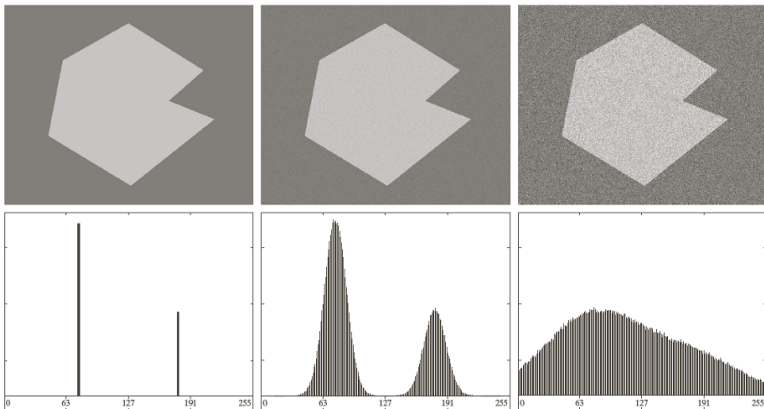
- przez progowanie,
- przez analizę skupień,
- przez wykrywanie krawędzi,
- metody obszarowe, tj. przez rozrost i/lub podział obszarów,
- inne, np.
 - przez dopasowywanie wzorca (template matching)
 - wododziałowa
 - przez podział grafu
 - algorytm mean-shift

Idea: binaryzacja obrazu na podstawie histogramu.
(stąd alternatywna nazwa: *histogram-based segmentation*)

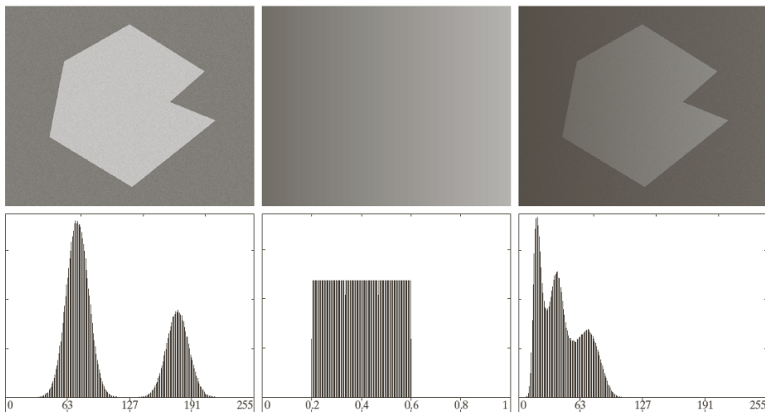
$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \leq T \end{cases}$$

Problemy:

- dobór progu(-ów) T ,
- ignorowanie relacji sąsiedztwa (przestrzennego charakteru obrazu),
- problematyczne w obrazach kolorowych (wielomodalnych).



Obrazy i histogramy z szumem o odchyleniu standardowym: 0, 10, 50.



Obraz w prawej kolumnie to iloczyn obrazów z lewej i środkowej kolumny.

Algorytm ISODATA, zasadniczo równoważny algorytmowi k-means:

T_0 = średnia jasność w całym obrazie

$k = 0$

Do

przeprowadź progowanie z progiem T_k

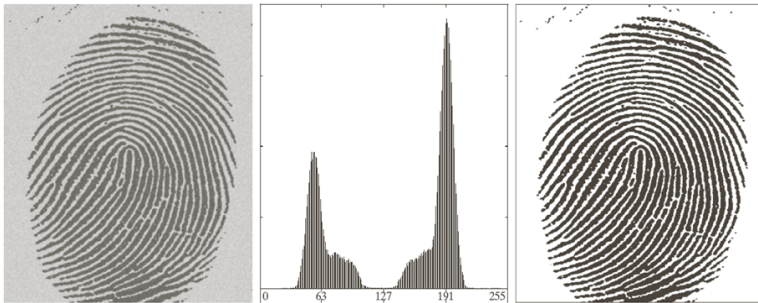
L_B = średnia jasność pikseli o jasnościach $< T_k$

L_F = średnia jasność pikseli o jasnościach $\geq T_k$

$k++$

$T_k = (L_B + L_F)/2$

While $T_k \neq T_{k+1}$



Idea: minimalizacja wariancji wewnątrzklasowej¹:

$$\sigma_{intra}^2(t) = P_1 \sigma_1^2 + P_2 \sigma_2^2$$

gdzie:

σ_i – średnia jasność w klasie [jasności] i

P_i – prawdopodobieństwo przynależności piksela do klasy i , tj.

$$P_1 = \Pr(f(x, y) < t), \quad P_2 = \Pr(f(x, y) \geq t),$$

Optymalny próg (wyszukiwany przeszukiwaniem pełnym):

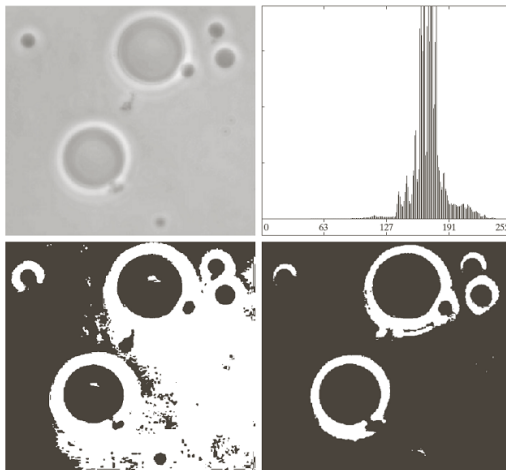
$$\arg \max_k \sigma_{intra}^2(t)$$

Otrzymana wielkość, po normalizacji (σ^2 - wariancja jasności w całym obrazie):

$$\frac{\sigma_{intra}^2(t)}{\sigma^2}$$

może być interpretowana jako miara separowalności.

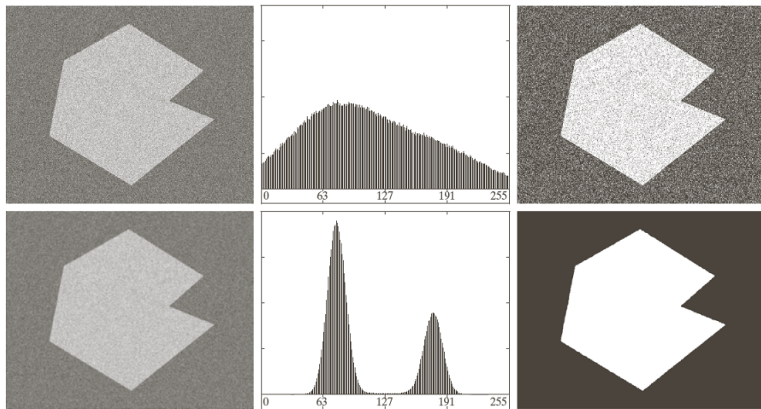
¹Równoważna maksymalizacji wariancji międzyklasowej.



Dół lewa: wynik segmentacji algorytmem Isodata;

Dół prawa: wynik segmentacji algorytmem Otsu.

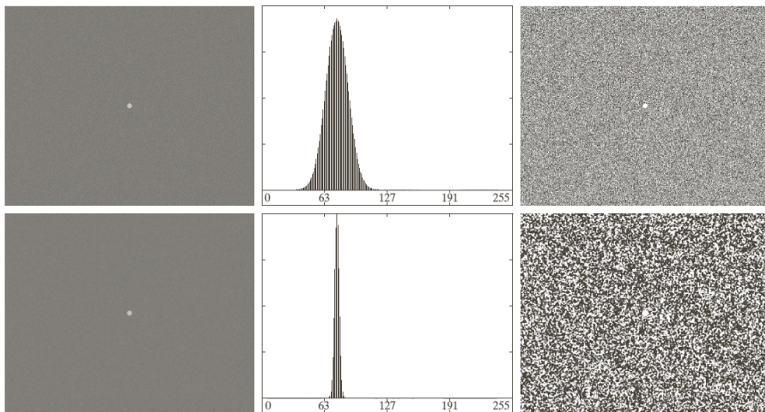
Przykład: Metoda Otsu w obecności szumu



Góra: segmentacja obrazu oryginalnego.

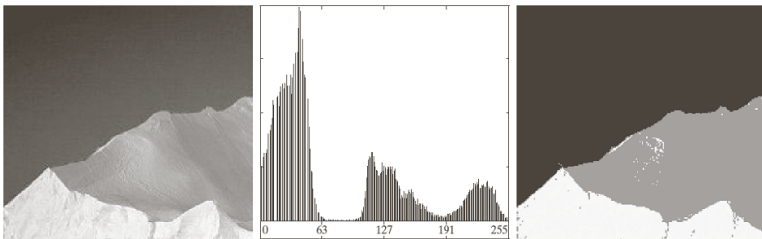
Dół: segmentacja obrazu po uprzednim przefiltrowaniu prostym filtrem uśredniającym 5×5 .

Przykład: Metoda Otsu dla małych obiektów



Góra: segmentacja obrazu oryginalnego. Dół: segmentacja obrazu po uprzednim przefiltrowaniu prostym filtrem uśredniającym 5×5 .

Ideę maksymalizacji odległości międzyskupieniowej da się łatwo uogólnić na wiele progów.

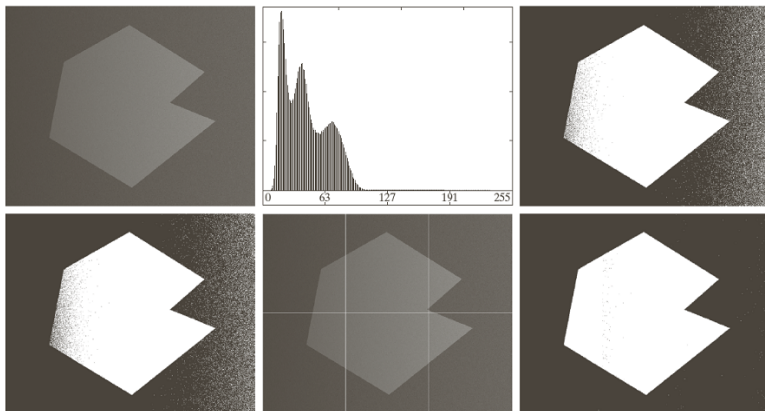


Idea: zamiast z całego obrazu, próg wyznaczany jest na podstawie fragmentu obrazu (lub, w skrajnym przypadku, [dużego] lokalnego otoczenia badanego punktu).

Metody wyznaczania lokalnego progu:

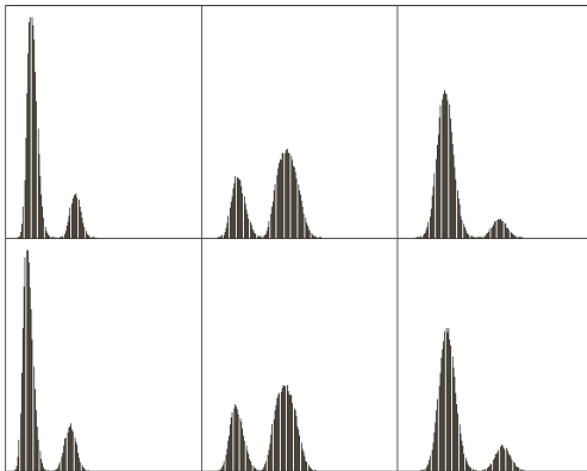
- Średnia,
- Mediana,
- Midrange
- Otsu

Progowanie z podziałem obrazu – przykład

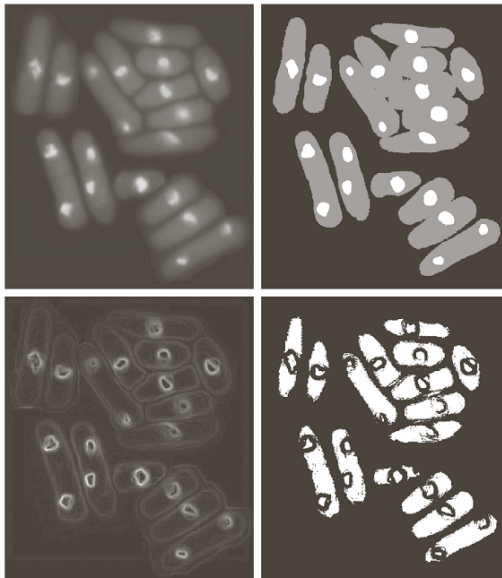


PG: Isodata; LD: Otsu; PD: Otsu na podobrazach

Histogramy podobrazów:



Idea: próg dla danego punktu wyliczany na podstawie jego lokalnego otoczenia
(otoczenie $n \times n$)



Próg wyliczany z fragmentu wiersza pikseli (otoczenie $n \times 1$)

and ninety six between Stockley
of Knox. And State of Tennessee
Andrew Jackson of the County
of Aboke said of the other part
said Stockley Donelson for A
of the Sum of two thousand
and paid the receipt where
hath And by these presents
self alien enkooff And Confe
Jackson his heirs And a
certain traits or parcels of La
sand airer/ony thousand are

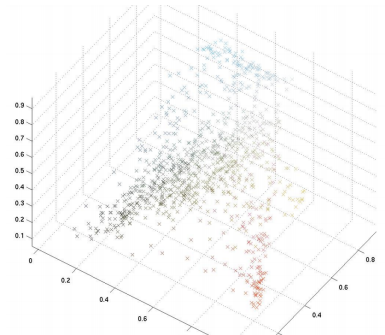
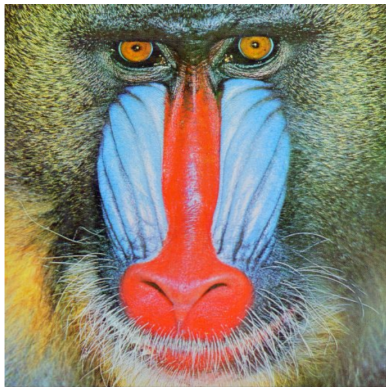
and ninety six between Stockley
of Knox. And State of Tennessee
Andrew Jackson of the County
of Aboke said of the other part
said Stockley Donelson for A
of the Sum of two thousand
and paid the receipt where
hath And by these presents
self alien enkooff And Confe
Jackson his heirs And a
certain traits or parcels of La
sand airer/ony thousand are

and ninety six between Stockley
of Knox. And State of Tennessee
Andrew Jackson of the County
of Aboke said of the other part
said Stockley Donelson for A
of the Sum of two thousand
and paid the receipt where
hath And by these presents
self alien enkooff And Confe
Jackson his heirs And a
certain traits or parcels of La
sand airer/ony thousand are

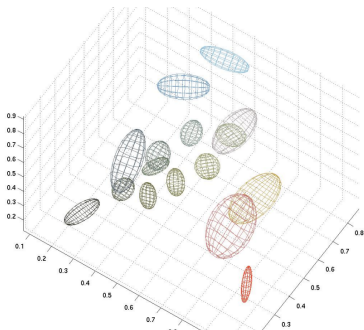
Uogólnienie segmentacji przez progowanie: punkty obrazu grupowane są na podstawie ich wielowymiarowej charakterystyki (progowanie – przypadek jednowymiarowy).

Typowa implementacja:

- klasyczna analiza skupień (algorytm Lloyd'a)
- algorytm EM (*expectation maximization*) – modelowanie wielowymiarowego rozkładu cech za pomocą sumy ważonej funkcji Gaussowskich (mixture of Gaussians)



(za: Silvio Savarese)



(za: Silvio Savarese)

Idea: zamiast wykrywać obszary, spróbujmy wykrywać granice pomiędzy nimi.

Najczęściej stosowane:

- splot (np. filtry Sobela, Robertsa, etc.),
- gradient, Laplasjan,
- podejścia morfologiczne.

Segmentacja przez wykrywanie krawędzi

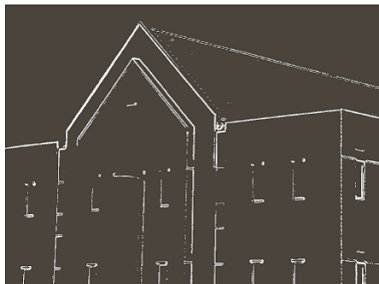
Zalety:

- dość szybkie,
- w porównaniu z segmentacją przez progowanie: brak problemów z ustalaniem wartości progu.

Wady:

- konieczność prześledzenia krawędzi i analizy topologicznej (to już nie jest tak szybkie),
- podatne na szumy,
- problemy z wykrywaniem „rozmytych” granic pomiędzy obszarami,
- niezbędny element: operator wykrywania krawędzi (por. wcześniejsze wykłady).

Segmentacja przez wykrywanie krawędzi - przykład



(obraz wejściowy, moduł gradientu, wynik algorytmu Canny)

Segmentacja przez wykrywanie krawędzi - przykład



Wymaga sprecyzowania dwóch komponentów:

- *Strategii przeszukiwania* przestrzeni wszystkich możliwych partycji
Na przykład: startując od małych elementów obrazu, sukcesywnie łącz sąsiadujące elementy obrazu w coraz większe, o ile spełniają przyjęte test jednolitości.
 - rozrost (łącznie),
 - podział,
 - mieszane.
- *Testu/predykatu jednolitości* $P(R)$: zwraca wartość tak/nie zależności od tego czy obszar R można uznać za reprezentację tego samego (pojedynczego) obiektu w obrazie.
 - Zazwyczaj zależy od funkcji charakterystycznej f (jasność, barwa, tekstura),
 - ... ale też niekiedy od innych charakterystyk, np. wielkości R .

Przykład testu jednolitości

Maksymalne odchylenie jasności względem średniego poziomu jasności w obszarze.

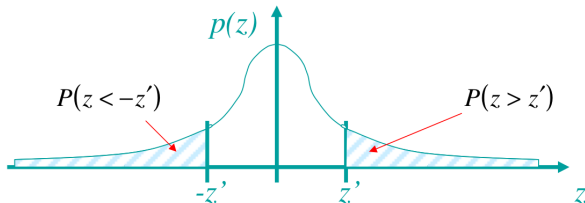
- Średni poziom jasności: $\bar{f}(R) = \frac{1}{|R|} \sum_{p \in R} f(p)$
- Test jednolitości: $P(R) \iff \max_{p \in R} |f(p) - \bar{f}(R)| \leq T$

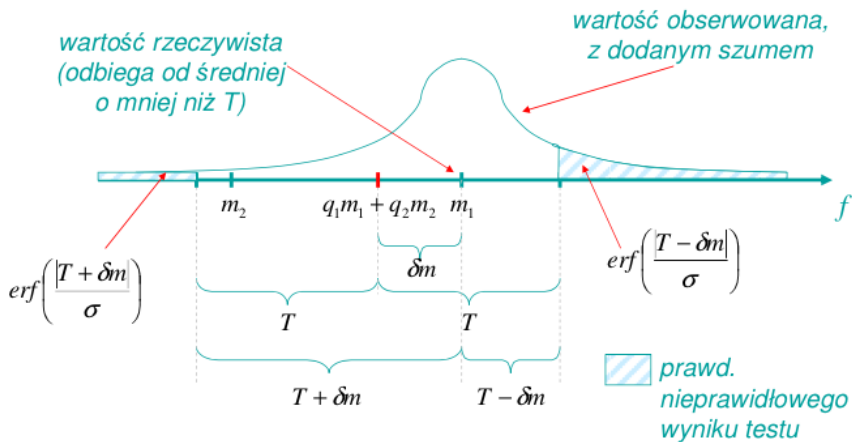
Niestety taki test jednolitości charakteryzuje się wysokim ryzykiem popełnienia błędów

1. uznania obszaru jednolitego za niejednolity (błąd I rodzaju),
2. uznania obszaru niejednolitego za jednolity (błąd II rodzaju).

Przyczyna: dla błędów I rodzaju:

- w obecności szumu (np. $\mathcal{N}(0, \sigma)$), prawdopodobieństwo błędu dla pojedynczego punktu jest małe $\Pr_p = \Pr[|f(p) + \mathcal{N}(0, \sigma) - \bar{f}(R)| > T]$,
- ale dla wszystkich punktów w R wynosi ono $1 - (1 - \Pr_p)^{|R|}$, i jest $\gg 0$





- Oparte na bardziej 'odpornych' statystykach jasności, np. wariancji: $\sigma_R \leq T$
- Oparte na bardziej wyrafinowanych cechach obrazu, np. histogram, macierz sąsiedztwa.
- Różnicowe: $P(R_1, R_2)$ zamiast $P(R)$, tj. może porównywać charakterystyki R_1 i R_2 z $R_1 \cup R_2$
 - Np. zmiana wariancji jasności.
- Biorące pod uwagę cechy topologiczne, np. długość granicy pomiędzy obszarami.

Idea: startując z 'elementarnych', małych obszarów, łącz je w coraz większe obszary do momentu gdy wynikowe obszary przestaną spełniać test jednolitości.

L - lista regionów

Inicjalizuj(L)

do

for wszystkie pary regionów przyległych $(R_1, R_2) \in L \times L$

$R \leftarrow R_1 \cup R_2$

if TestJednolitości(R) then

$L.$ Usun(R_1)

$L.$ Usun(R_2)

$L.$ Dodaj(R)

endif

endfor

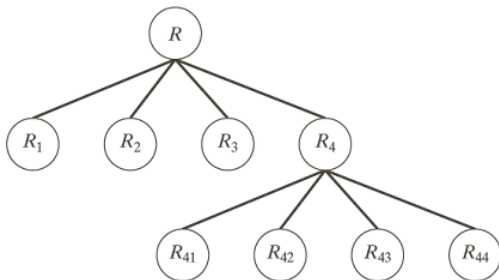
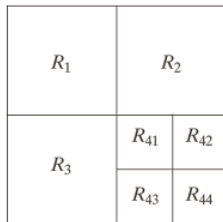
while not(WarunekStopu(L))

Inicjalizacja L : małe fragmenty obrazu (np. obszary 2×2)

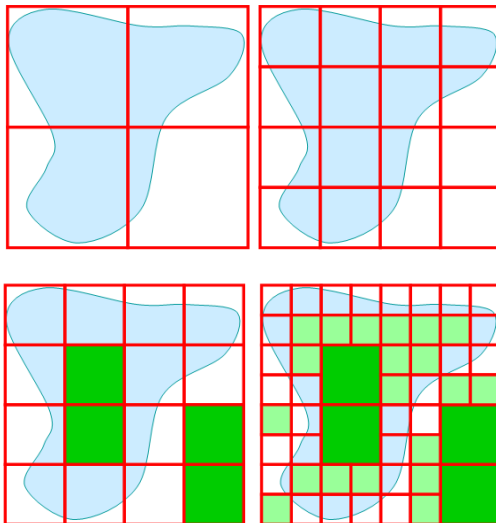
- Zastąpienie testu jednolitości **kryterium** jednolitości i wykonywanie operacji łączenia w kolejności wyznaczonej tym kryterium (podobne regiony będą łączone wcześniej)
 - czyli: przejście na algorytm zachłanny
- Zaleta: większa wierność obrazu wynikowego w stosunku do oryginalnego,
- Wada: większy koszt obliczeniowy

Idea: startując z całego obrazu, sukcesywnie dziel go na coraz mniejsze podobrazy do momentu gdy wynikowe podobrazy spełnią test jednolitości.

```
L - lista obszarów
L.Dodaj(jeden obszar składający się z całego obrazu)
do
  for wszystkie regiony  $R \in$ 
    L takie że not(TestJednolitości( $R$ ))
      Dokonaj podziału  $R$  na podobszary  $R_1$  i  $R_2$ 
        L.Dodaj( $R_1$ )
        L.Dodaj( $R_2$ )
        L.Usuń( $R$ )
      endif
    endfor
  while not(WarunekStopu(L))
```

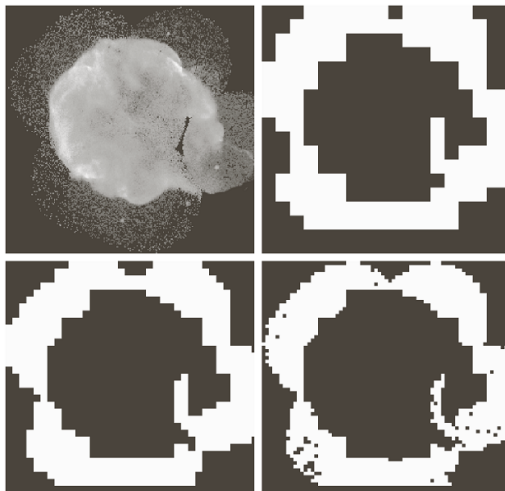
- Całkowita liczba węzłów drzewa: $\sum_{k=0}^n 4^k = \frac{4}{3}4^n$, gdzie n - liczba poziomów (głębokość) drzewa + 1.
- Zauważmy że pełne drzewo czwórkowe (dla całego obrazu) ma głębokość $\log_4(\text{liczbaPikseli})$
- Zatem: węzłów drzewa jest tylko o ok. 33% więcej niż punktów obrazu.



Wymagane dalsze połączenie węzłów drzewa „w poprzek” jego struktury, zgodnie z sąsiedztwem kwadratów w obrazie.

Segmentacja obrazu supernowej Cygnus Loop (pasmo Rentgenowskie)
zorientowana na wydzielenie obszarów o znacznej wariancji.

- Nietypowy test jednolitości: $\sigma > a \wedge \bar{f} < b$, gdzie a i b to parametry.
- Obraz oryginalny i wynik segmentacji z ograniczeniem na rozmiar regionu 32x32, 16x16, i 8x8 pikseli.



Zakłada że dany jest pewien zbiór punktów startowych (ziaren, *seeds*), od których zaczyna się rozrost obszarów.

Dany punkt przyłączany jest do ziarna s jeżeli:

- spełnia test jednolitości,
- istnieje ścieżka punktów spełniających test jednolitości, łącząca punkt p z s

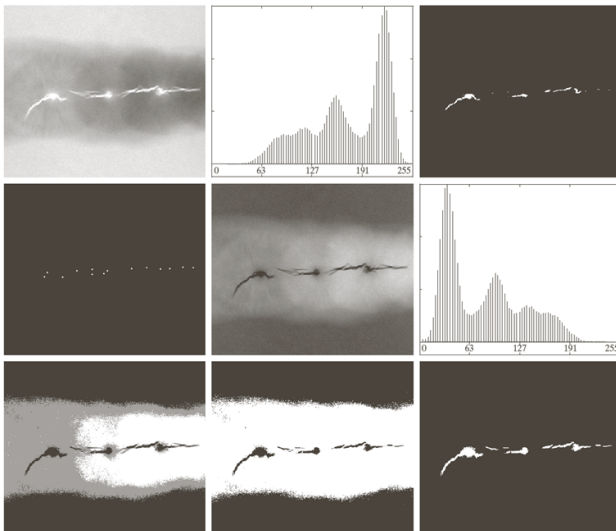
Legenda do następnego slajdu:

a	b	c
d	e	f
g	h	i

FIGURE 10.51 (a) X-ray image of a defective weld. (b) Histogram. (c) Initial seed image. (d) Final seed image (the points were enlarged for clarity). (e) Absolute value of the difference between (a) and (c). (f) Histogram of (e). (g) Difference image thresholded using dual thresholds. (h) Difference image thresholded with the smallest of the dual thresholds. (i) Segmentation result obtained by region growing. (Original image courtesy of X-TEK Systems, Ltd.)

Wykorzystany test jednolitości: $|f(p) - f(s)| \leq T$

Algorytm z 'ziarnem' – przykład



Motywacje: wady wcześniej wymienionych algorytmów, tj.:

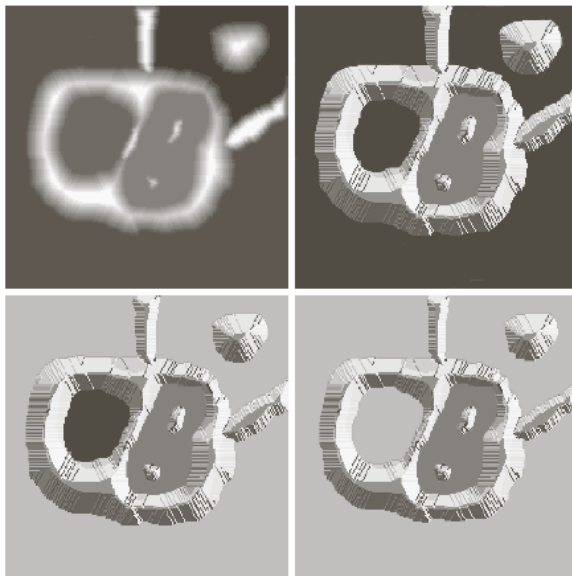
- algorytmy łączenia, startując z pojedynczych pikseli, silnie zależą od właściwości tych pikseli
 - ponadto, dla pojedynczych pikseli nie da się policzyć żadnych statystyk opisowych,
- z drugiej strony, algorytmy podziału mogą nie 'odkryć' pewnych detali obrazu.

Rozwiązanie:

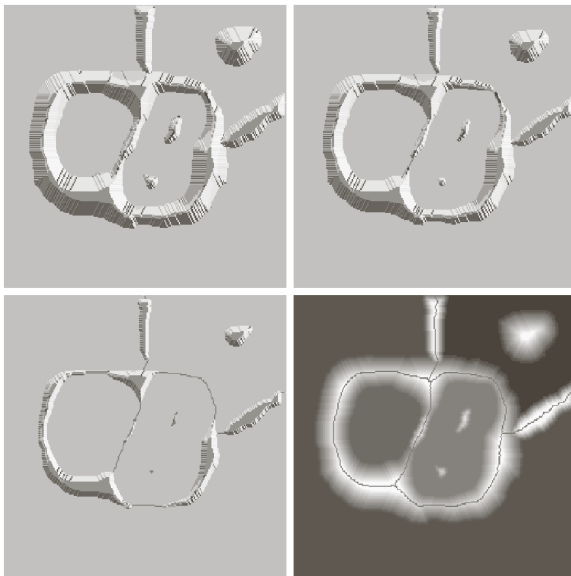
- Algorytmy hybrydowe, tj. podziału i łączenia.
- Idea: start z pośredniego poziomu drzewa (np. bloki pikseli 8x8).

Ang. *watershed segmentation*

- Klasa algorytmów segmentacji stosująca analogię map reliefowych:
 - obraz wejściowy interpretowany jako powierzchnia topograficzna (w 3 wymiarach)
 - obszar odpowiada *zlewisku*
- Wszystkie punkty obrazu można podzielić na 3 kategorie:
 - punkty reprezentujące lokalne minima,
 - punkty leżące na 'zboczach'
 - punkty leżące na 'graniach'
- Punkty w których spotykają się rozlewiska wyznaczają granice pomiędzy obszarami.
- Punkt przynależy do tego zlewiska (obszaru) do którego spłynęłaby kropla wody upuszczona na obraz gradientowy.

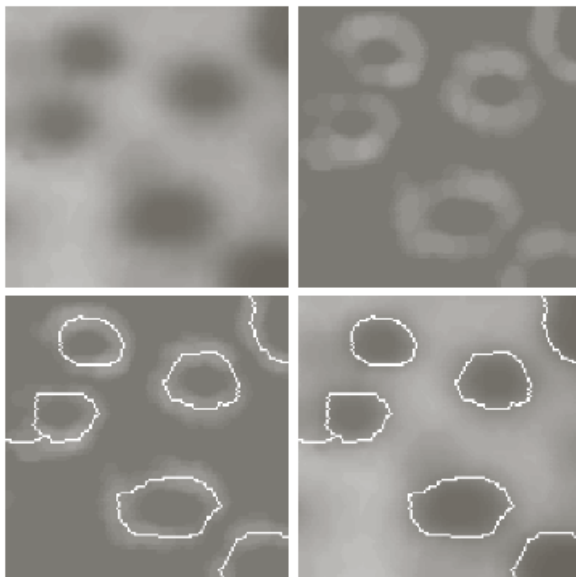


Obraz oryginalny, interpretacja topograficzna (rendering 3D, z cieniem), oraz dwa pierwsze etapy zalewania zlewisk (począwszy od najniżej położonych dolin

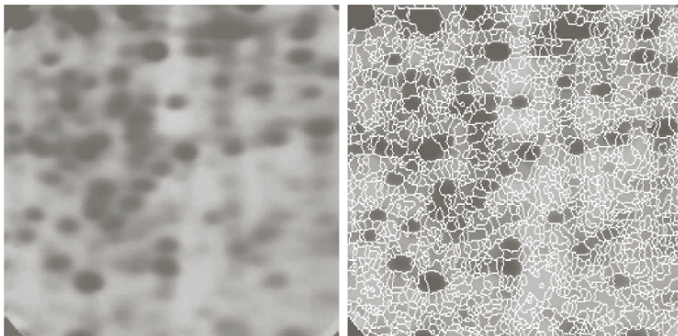


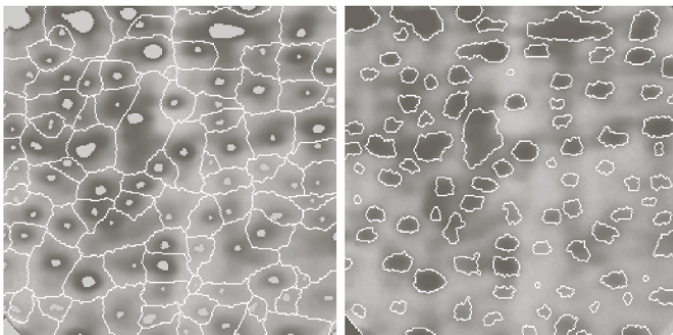
Woda wypełniające poszczególne zlewiska zaczyna spotykać się na graniach;
budujemy 'tamy' aby nie dopuścić do połączenia zlewisk.

- Istotne: Obraz wejściowy dla algorytmu segmentacji wododziałowej jest zazwyczaj obrazem gradientowym otrzymanym z oryginalnego obrazu.
- Możliwe realizacje praktyczne:
 - „Rozlewanie” (flood fill) startujące z lokalnych minimów gradientu i rejestrowaniem punktów gdzie gdzie spotykają się zlewiska
 - Wykorzystanie odległości od minimów lokalnych
- (dają różne wyniki)



- Przed obliczeniami warto policzyć histogram obrazu g :
 - z histogramu natychmiast dostajemy wartości min i max
 - być może nie wszystkie jasności występują w obrazie; jeśli tak jest to niewystępujące poziomy jasności warto pominąć (tj. zwiększać n o odpowiednio większe przyrosty).
- Częsty problem: oversegmentation. Przyczyna: duża liczba minimów lokalnych.
 - Możliwe rozwiązanie: wykorzystanie znaczników (markers)
 - Znaczniki definiują nowe punkty startowe dla procesu budowania zlewisk (czyli $T[\min+1]$).





Znaczniki: lokalne minima w *oryginalnym* obrazie przepuszczonym przez filtr dolnoprzepustowy. Wynik w prawym obrazie uzyskany przez zastosowanie algorytmu wododziałowego do każdego z regionów z lewego obrazu z osobna.

- Segmentacja wododziałowa da się wprost zastosować dla obrazów z gradacją stopni szarości.
- Co z obrazami binarnymi?
 - Dla 'płaskich' obszarów, zlewiska w segmentacji wododziałowej przyrastają bardzo szybko.
- Czy istnieje możliwość takiego 'uciąglenia' obrazu binarnego aby segmentacja wododziałowa dała sensowny wynik?
 - Odpowiedź: Distance transform.

Distance transform (transformacja odległościowa): (wymaga obrazu binarnego)

Dla każdego piksela należącego do obiektu:

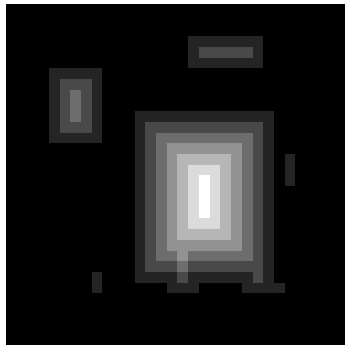
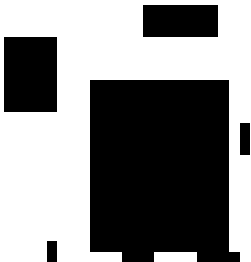
- Znajdź najbliższego sąsiada należącego do tła
- Oblicz odległość względem tego sąsiada
- Przypisz pikselowi wynikowemu wartość tej odległości.

Stopnie szarości w obrazie wynikowym wizualizują powyższą odległość.

	1	1	1	1	1	
	1	1	1	1	1	
	1	1	1	1	1	
	1	1	1	1	1	
	1	1	1	1	1	

	1	1	1	1	1	
	1	2	2	2	1	
	1	2	3	2	1	
	1	2	2	2	1	
	1	1	1	1	1	

Obraz oryginalny i transformacja odległościowa przy metryce miejskiej (dla powyższego obrazu efekt dla metryki Euklidesowej byłby taki sam).



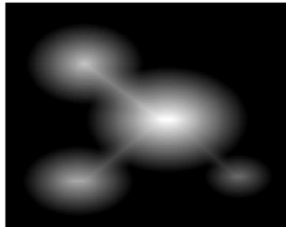
Uwagi:

- Powiązanie ze szkieletem obiektu (zobacz rozdział Opis obrazów i obiektów).
- Realizacja techniczna: obraz wynikowy musi być typu rzeczywistego (float, double).
- Szybka implementacja z wykorzystaniem erozji.
 - Kształt elementu strukturalnego determinuje metrykę transformaty.
 - Możliwe usprawnienia?

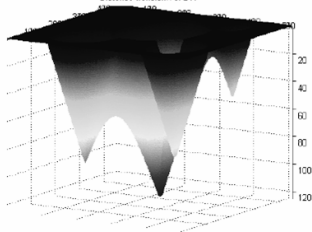
BW



Distance transform of BW



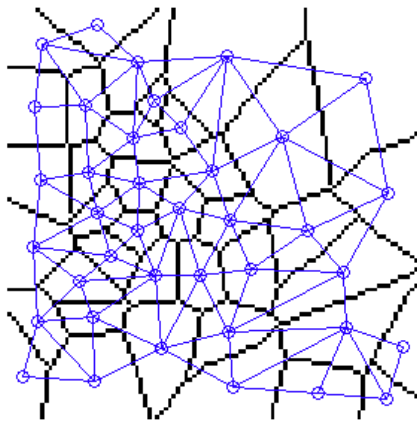
Distance transform of BW



Watershed transform of BW



- Mapa pokolorowanych obszarów
- Graf przyległości obszarów (*region adjacency graph*, RAG), graf nieskierowany taki że:
 - węzły = obszary
 - krawędzie = relacja sąsiedztwa pomiędzy obszarami.



Graf czy multigraf?

- Planarny, a zatem w konsekwencji:
 - Spełnia wzór Eulera: $vertices - edges + f_{acets} = 2$
 - Da się pokolorować 4 kolorami
- Jeżeli piksele granic pomiędzy obszarami wraz z relacją ich sąsiedztwa potraktować jako graf, to RAG jest grafem dualnym względem takiego grafu
- Mała liczba krawędzi a duża liczba węzłów: dogodne reprezentowanie w postaci listy,
- Cenna informacja dla etapu wnioskowania z obrazu; łatwo zdefiniować wiele informatywnych cech RAG, np.:
 - v, e, f
 - Średni stopień wierzchołka
- ... i/lub pojedynczych obszarów/węzłów w RAG:
 - Stopień wierzchołka (np. przypadek szczególny: stopień=1)
 - Średnia odległość regionu przyległego

Analiza wyników działania algorytmów segmentacji 'na oko' nie jest dobrym pomysłem. Potrzeba obiektywnych miar.

Pytania:

- Jak odnieść wynik działania algorytmu segmentacji do rzeczywistych regionów w obrazie?
- Jak porównać wyniki działania dwóch algorytmów segmentacji?

Jedna z często wykorzystywanych miar: Dice statistics:

$$S(R_A, R_B) = \frac{2|R_A \cap R_B|}{|R_A \cup R_B|}$$

gdzie R_A i R_B to

- Obszary znalezione przez dwa algorytmy segmentacji A i B, lub
 - Obszar zdefiniowane przez eksperta (*ground truth*, A) i obszar znaleziony przez algorytm B
- Bardzo zbliżony do Jaccard index: $|R_A \cap R_B| / |R_A \cup R_B|$
- Można uogólnić na wiele obszarów

- Wiele algorytmów segmentacji da się uogólnić na przypadek obrazów wielokanałowych.
- Wymagana modyfikacja: redefinicja kryterium (-ów) jednolitości obszarów.
- W ogólności po segmentacji obrazów wielokanałowych można często spodziewać się lepszych wyników (niż dla obrazów monochromatycznych), ponieważ dysponujemy bogatszą informacją o poszczególnych pikselach.





[D. Comaniciu and P. Meer]