Cyfrowe przetwarzanie obrazów i sygnałów
Wykład 6
AiR III

Joanna Ratajczak

KCiR (W4/K7)

Segmentacja – wydzielenie obszarów obrazu spełniających pewne kryteria jednorodności (np. kolor obszaru, poziom jasności, faktura).

Zadaniem procedur segmentacji obrazów jest jedynie podział obrazu na odpowiednie obszary a nie ich rozpoznawanie.

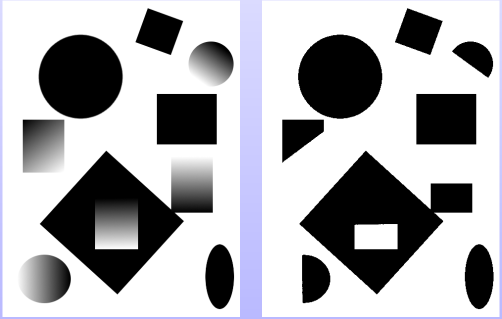
Binaryzacja – proces konwersji obrazów monochromatycznych lub kolorowych do obrazu binarnego.

- Cel: redukcja zbędnej informacji i pozostawienie informacji istotnej z punktu widzenia dalszego przetwarzania.
- Najczęściej polega na oddzieleniu obiektów od tła.
- Istotny element procesu analizy scen. Stanowi etap przygotowawczy do rozpoznawania obiektów.
- Efekt binaryzacji wpływa na ostateczny wynik złożonego procesu analizy.
- Redukcja informacji obniża złożoność algorytmów rozpoznawania → zmniejsza złożoność czasowa całego procesu analizy.
- Wykorzystywana jako wstępny etap procesu analizy dokumentów.
- Najczęściej realizowana poprzez progowanie.

Progowanie – thresholding

Binaryzacja z dolnym progiem

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) \le t \\ 1 & \text{jeżeli} \quad f(x,y) > t \end{cases}$$



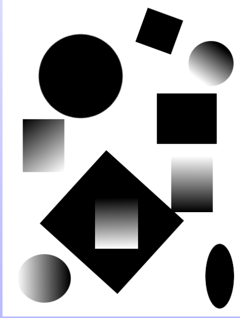
Progowanie – thresholding

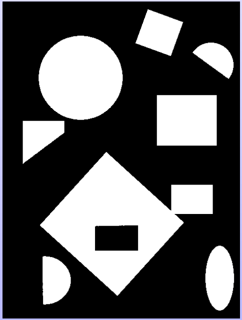
- Brana jest pod uwagę tylko intensywność piksela.
- Założenie: dwa mody histogramu reprezentują dwie rozdzielne klasy.
- Progowanie globalne vs. progowanie lokalne.
- Wybór progu:
  - automatyczny,
  - ręczny.

Progowanie – thresholding

Binaryzacja z górnym progiem

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) \geq t \\ 1 & \text{jeżeli} \quad f(x,y) < t \end{cases}$$

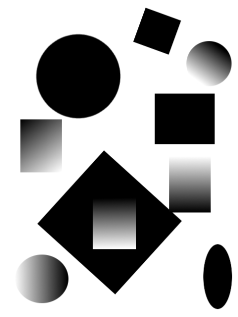


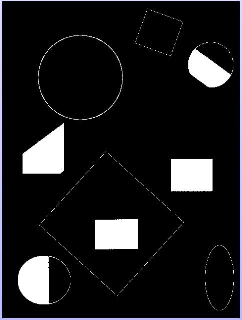


Progowanie – thresholding

Binaryzacja z podwójnym ograniczeniem

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) < u \vee f(x,y) > v \\ 1 & \text{jeżeli} \quad u \leq f(x,y) \leq v \end{cases}$$





Progowanie – thresholding

Binaryzacja warunkowa

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad u \geq f(x,y) \\ s & \text{jeżeli} \quad u < f(x,y) \leq v \\ 1 & \text{jeżeli} \quad f(x,y) > v \end{cases}$$

Binaryzacja ze zbiorem

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) \notin Z \\ 1 & \text{jeżeli} \quad f(x,y) \in Z \end{cases} ; \qquad Z \subseteq [z_1, z_k]$$

Półprogowanie (semithresholding)

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) < t \\ f(x,y) & \text{jeżeli} \quad f(x,y) \geq t \end{cases}$$

Progowanie jasności

- Próg globalny – wartość progu dobierana jest na podstawie całego obrazu.
- Próg lokalny – próg zależny od treści obrazu, tj. $f(x, y)$ oraz od pewnej cechy obrazu, np. średniej jasności w pewnym otoczeniu.
- Próg ustalany dynamicznie – wartość progu jest zależna od współrzędnych przestrzennych obrazu.

Wybór progu

- Wartość progu dobiera się najczęściej na podstawie histogramu obrazu.

Jeśli znamy wypełnienie Θ to

$$\int_{t}^{z_k} H(z)\,\mathrm{d}z = \Theta \int_{z_1}^{z_k} H(z)\,\mathrm{d}z.$$

Wybór progu – histogram bimodalny

1. Znaleźć dwa największe maksima lokalne histogramu $H(z)$ odległe od siebie co najmniej o Δ: z_i, z_j;
2. Znaleźć próg jako minimum histogramu w przedziale (z_i, z_j)

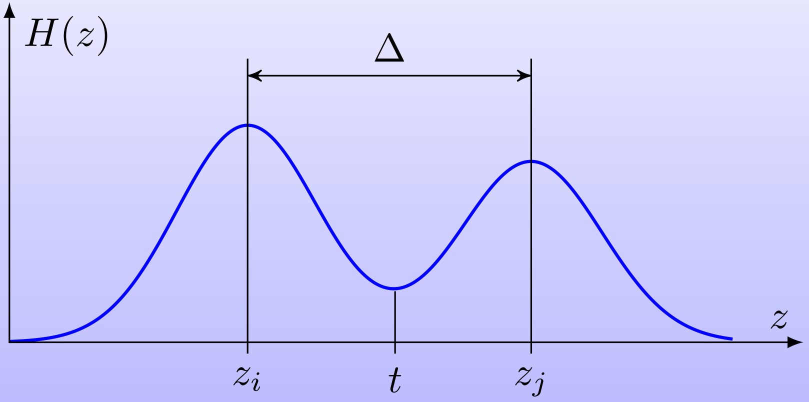
$$t : H(t) \le H(z) \qquad \text{dla} \qquad z_i < z < z_j;$$

pod warunkiem, że płaskość (flatness) histogramu

$$\frac{H(z_k)}{\min(H(z_i), H(z_j))}$$

jest mała (histogram jest istotnie bimodalny).

Wybór progu – histogram bimodalny



Progowanie optymalne

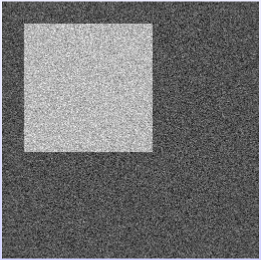
Niech histogram bimodalny będzie kombinacją wypukłą dwóch unimodalnych rozkładów jasności punktów obiektu $p(z)$ i tła $q(z)$

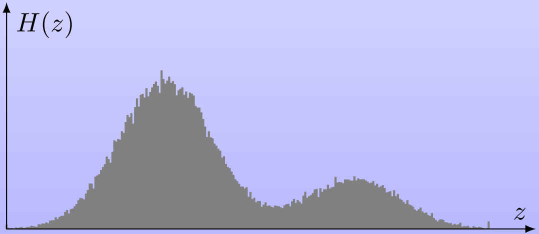
$$H(z) = \Theta p(z) + (1 - \Theta)q(z),$$

gdzie Θ jest częścią pola obrazu zajmowaną przez obiekt. Rozkłady jasności odpowiednio dla obiektu (jaśniejszego) i tła (ciemniejszego) są równe

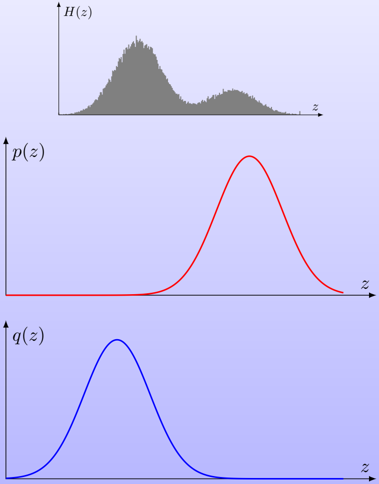
$$p(z) = N(m, s), \qquad q(z) = N(n, u), \qquad m > n.$$

Progowanie optymalne





Progowanie optymalne



Progowanie optymalne

$\epsilon_1(t) = \int_{-\infty}^{t} p(z)\,\mathrm{d}z$ – błędne zaliczenie punktów obiektu do tła

$\epsilon_2(t) = \int_{t}^{\infty} q(z)\,\mathrm{d}z$ – błędne zaliczenie punktów tła do obiektu

Minimalizujemy łączne prawdopodobieństwo błędnej klasyfikacji

$$\epsilon(t) = \Theta\epsilon_1(t) + (1-\Theta)\epsilon_2(t)$$

$$\frac{\mathrm{d}\epsilon(t)}{\mathrm{d}t} = 0 = \Theta p(z) - (1-\Theta)q(z)$$

$$\frac{\Theta}{s\sqrt{2\pi}} e^{-\frac{(t-m)^2}{2s^2}} = \frac{1-\Theta}{u\sqrt{2\pi}} e^{-\frac{(t-n)^2}{2u^2}}$$

Progowanie optymalne

$$\frac{\Theta}{s\sqrt{2\pi}}e^{-\frac{(t-m)^2}{2s^2}} = \frac{1-\Theta}{u\sqrt{2\pi}}e^{-\frac{(t-n)^2}{2u^2}} \qquad \Big/ \ln()$$

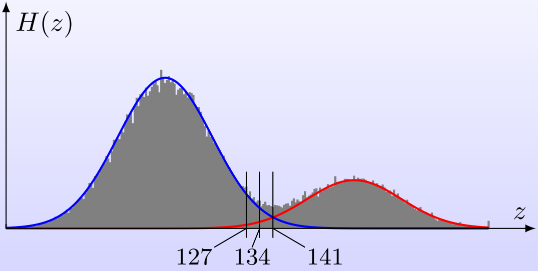
$$\ln(\Theta) - \ln(s) - \frac{(t-m)^2}{2s^2} = \ln(1-\Theta) - \ln(u) - \frac{(t-n)^2}{2u^2}$$

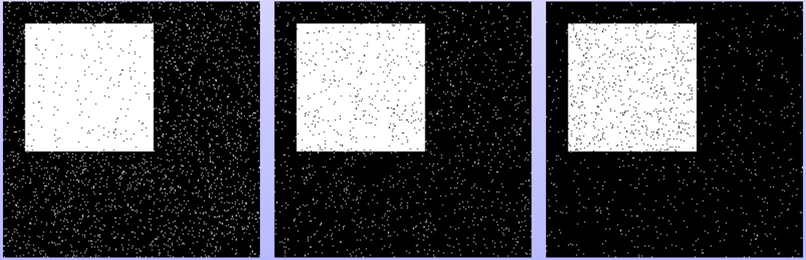
$$u^2(t-m)^2 - s^2(t-n)^2 = 2s^2u^2\ln\frac{u\Theta}{s(1-\Theta)}$$

Gdy wariancje $p(z)$ i $q(z)$ są równe ($s = u$) to

$$t = \frac{m+n}{2} + \frac{s^2}{n-m}\ln\frac{\Theta}{1-\Theta}$$

Dla $\Theta = 0.5$ optymalna wartość progu jest równa średniej arytmetycznej ze średnich jasności tła i obiektu.


$H(z)$
z
127
134
141



Progowanie w sensie Bayesa

Cel: Znaleźć podzbiór Z przedziału $[z_1, z_k]$ taki, że punkty obrazu o jasnościach z tego podzbioru zaliczamy do obiektu.

$P(o, z)$ – łączne prawdopodobieństwo tego, że punkt należy do obiektu i ma jasność z;
$P(b, z)$ – łączne prawdopodobieństwo tego, że punkt należy do tła i ma jasność z;
$P(o)$ – prawdopodobieństwo *a priori* należenia punktu do obiektu ($P(o) = \Theta$);
$P(b)$ – prawdopodobieństwo *a priori* należenia punktu do tła ($P(b) = 1 - \Theta$);
$P(z|o)$ – rozkład (warunkowy) jasności dla obiektu;
$P(z|b)$ – rozkład (warunkowy) jasności dla tła;

$$P(o, z) = P(o)P(z|o), \qquad P(b, z) = P(b)P(z|b)$$

Progowanie w sensie Bayesa

$P(o|z)$ – prawdopodobieństwo *a posteriori* należenia punktu o jasności z do obiektu;
$P(b|z)$ – prawdopodobieństwo *a posteriori* należenia punktu o jasności z do tła;
Histogram $H_n(z)$ – prawdopodobieństwo, że punkt ma jasność z.

$$P(o,z) = H_n(z)P(o|z), \qquad P(b,z) = H_n(z)P(b|z)$$

skąd otrzymujemy

$$P(o|z) = \frac{P(o)P(z|o)}{H_n(z)}, \qquad P(b|z) = \frac{P(b)P(z|b)}{H_n(z)}$$

$$Z = \{z \colon P(o|z) > P(b|z)\}$$

$$Z = \{z \colon P(o)P(z|o) > P(b)P(z|b)\}$$

Progowanie iteracyjne (Ridler, Calvard)

- Nie wymaga apriorycznej wiedzy na temat rozmiaru i położenia obiektów na scenie.
- Algorytm
  1. Oszacuj początkową wartość progu, np. przyjmij, że tło stanowią cztery narożne piksele obrazu. Pozostałe piksele stanowią piksele obiektu. Przy takim założeniu wyznacz średnią jasność tła μ_b oraz średnią jasność obiektu μ_o. Za wartość progową przyjmij

$$t^{(0)} = \frac{\mu_o + \mu_b}{2}$$

Progowanie iteracyjne (Ridler, Calvard)

2. Wyznacz wartość średnią dla klasy tła i klasy obiektu, przy czy, podziału klasy dokonaj dla progu wyznaczonego w poprzedniej iteracji algorytmu

$$\mu_o^{(n)} = \frac{\sum_{\text{obiekty}} f(x,y)}{\#\text{pikseli obiektów}}$$

$$\mu_b^{(n)} = \frac{\sum_{\text{tło}} f(x,y)}{\#\text{pikseli tła}}$$

3. Wyznacz nową wartość progu

$$t^{(n+1)} = \frac{\mu_o^{(n)} + \mu_b^{(n)}}{2}$$

4. Gdy $t^{(n+1)} = t^{(n)}$ KONIEC, w przeciwnym wypadku przejdź do punktu 2.

Progowanie podwójne (histerezowe)

- Założenie: jasny obiekt na ciemnym tle.
- Stosuje się w przypadkach, gdy klasy pikseli nie są wyraźnie odseparowane, np. granice obiektu są rozmazane, niewyraźne.
- Dwa progi:
  - radykalny t_r (punkty jaśniejsze na pewno należą do obiektu),
  - liberalny t_l (punkty należą do obiektu, o ile w ich otoczeniu jest któryś z powyższych)

$$z_1 < t_l < t_r < z_k$$

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad (\forall(\xi,\psi) \in S_{x,y})(f(\xi,\psi) < t_r) \\ 1 & \text{jeżeli} \quad (f(x,y) \geq t_l) \wedge (\exists(\xi,\psi) \in S_{x,y})(f(\xi,\psi) \geq t_r) \end{cases}$$

Progowanie lokalne (White, Rohrer)

- Wartość progu ustalana jest niezależnie dla każdego piksela obrazu na podstawie wartości średniej w otoczeniu piksela

$$g(x,y) = \begin{cases} 0 & \text{jeżeli} \quad f(x,y) < \frac{\mu(i,j)}{k} \\ 1 & \text{jeżeli} \quad f(x,y) \geq \frac{\mu(i,j)}{k} \end{cases}, \quad k > 1$$

- Założenie: piksele obiektu przyjmują wartości dużo mniejsze niż piksele tła.
- Rozmiar okna powinien być dobrany zależnie od obiektów zainteresowania na obrazie.
- Problematyczny dobór parametru k dla szerokiej klasy obrazów.

Progowanie lokalne (Bernsen)

- Metoda w wielu przypadkach dość efektywna.
- Wartość progu ustalana jest niezależnie dla każdego piksela w zależności od wartości minimalnej i maksymalnej w ustalonym otoczeniu piksela

$$t(x,y) = \frac{f_{min} + f_{max}}{2}.$$

- W przypadku słabszego kontrastu w obrębie okna niż założono, przynależność do klasy ustala się na podstawie wybranego eksperymentalnie progu globalnego.
- Rozmiar okna powinien być dobrany zależnie od obiektów zainteresowania na obrazie.

Korekcja nierówności oświetlenia

$$f(x,y) = e(x,y)g(x,y)$$

Obraz sceny testowej o stałej (znanej) funkcji odbicia $g_c(x,y) = v$

$$f_c(x,y) = e(x,y)g_c(x,y) = ve(x,y)$$

$$e(x,y) = \frac{1}{v}f_c(x,y)$$

Korekcja przy pobieraniu każdego obrazu

$$g(x,y) = \frac{f(x,y)}{e(x,y)} = v\frac{f(x,y)}{f_c(x,y)}$$

Progowanie zmienne

$$g_t(x,y) = \begin{cases} 1 & \text{jeżeli} \quad g(x,y) \geq t \qquad \text{tzn. } f(x,y) \geq \frac{t}{v} f_c(x,y) = t\, e(x,y) \\ 0 & \text{jeżeli} \quad g(x,y) < t \qquad \text{tzn. } f(x,y) < \frac{t}{v} f_c(x,y) \end{cases}$$

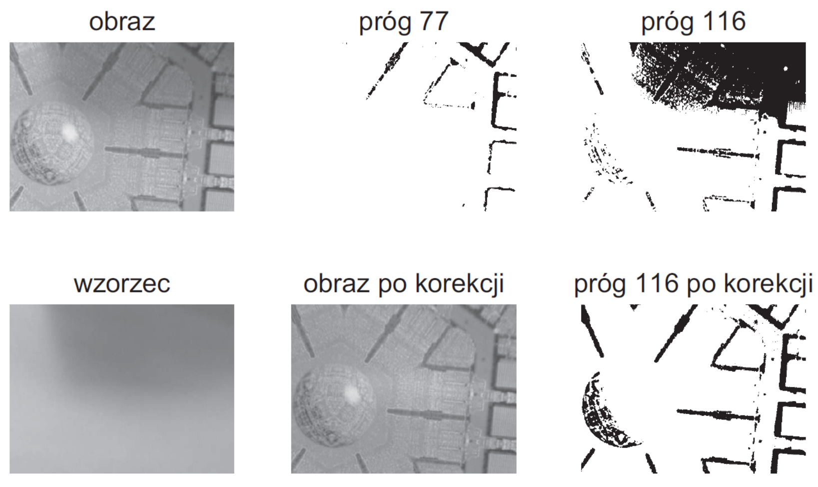
Stałe progowanie obrazu skorygowanego można zastąpić zmiennym progowaniem obrazu oryginalnego

$$g_t(x,y) = \begin{cases} 1 & \text{jeżeli} \quad f(x,y) \geq t(x,y) \\ 0 & \text{jeżeli} \quad f(x,y) < t(x,y) \end{cases},$$

gdzie

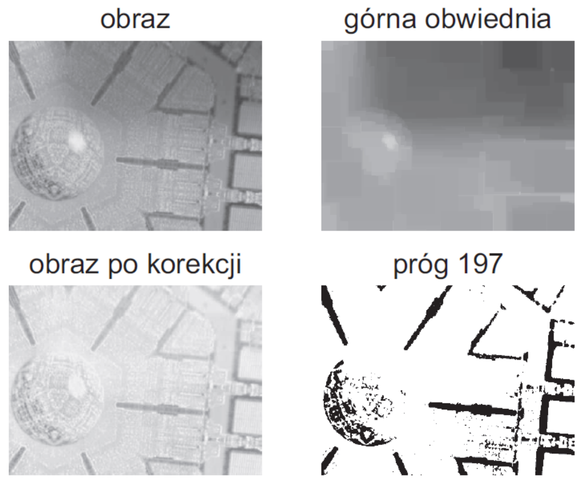
$$t(x,y) = \frac{t}{v} f_c(x,y).$$

Progowanie zmienne



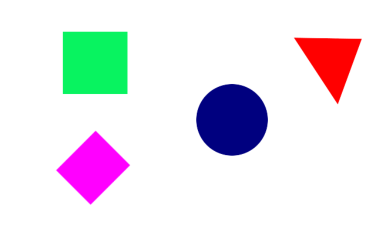
Korekcja oświetlenia bez wzorca

górna obwiednia obrazu: $upenv = Imin(Imax)$



Indeksowanie sylwetek (labelling)

Obraz wejściowy jest binarny i zawiera obszary o wartości 1 na tle o wartości 0 (wynik segmentacji obrazów lub progowania). **Cel:** przypisanie do wszystkich pikseli obiektów identyfikatorów wskazujących, do którego obiektu który piksel można przypisać.



Obszary spójne na obrazie wyjściowym mają różne wartości (od 1 do N, gdzie N jest liczbą obszarów).

Indeksowanie sylwetek (labelling)

1. Propagacja indeksu sylwetki zgodnie z przyjętymi regułami sąsiedztwa punktów w trakcie jednokrotnego przeglądania obrazu (liniami).

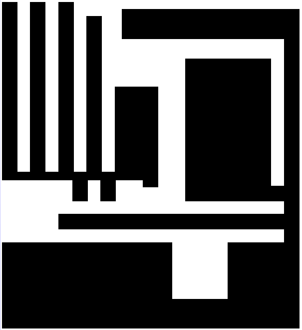
A	B	C
D	X	

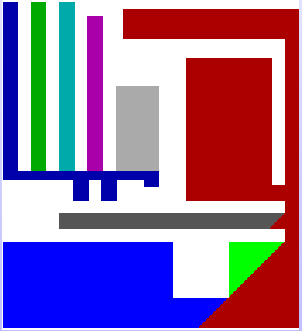
0	0	0
0	L+1	

0	0	L1
0	L1	

L1	0	L3
L1	L1	

Indeksowanie sylwetek (labelling)





R. Tadeusiewicz, P. Korohoda, Komputerowa analiza i przetwarzanie obrazów, Wyd. Fundacji Postępu Telekomunikacji, 1997

Indeksowanie sylwetek (labelling)

2. Przechodnie domknięcie relacji utworzonej w tablicy sklejeń odzwierciedlającej sąsiadowanie obszarów, którym przydzielono pierwotnie różne indeksy.

1	2	3	4	5	$\cdots$	254	255
0	0	0	0	0	$\cdots$	0	0

1	2	3	4	5	$\cdots$	254	255
1	2	3	0	0	$\cdots$	0	0

1	2	3	4	5	$\cdots$	254	255
1	2	1	4	0	$\cdots$	0	0

Indeksowanie sylwetek (labelling)

3. Konwersja otrzymanego obrazu przez tablicę sklejeń w celu połączenia podobszarów spójnych.