

9 — Segmentacja obrazów

9.1 Cel

- zapoznanie z metodami segmentacji obrazów:
 - segmentacją przez rozrost,
 - segmentacją przez podział i scalanie,

9.2 Wstęp

W ramach dotychczas wykonanych ćwiczeń poznaliśmy segmentację z wykorzystaniem binaryzacji (progowania) – tj. na podstawie jasności (koloru) poszczególnych pikseli. Wykonaliśmy dwa warianty metody: globalny i lokalny oraz przetestowaliśmy różne podejścia do automatycznego wyznaczania progu bianryzacji: Otsu, Yen, Sauvola.

Ponadto poznaliśmy możliwość segmentacji na podstawie krawędzi z wykorzystaniem transformaty Hough'a.

W tym ćwiczeniu poznamy dwie inne metody podziału obrazu na fragmenty:

- segmentację przez rozrost obszaru (ang. *region growing*),
- segmentację przez podział i scalanie (ang. *split and merge*).

9.3 Podstawy

Niech R oznacza obszar równy całemu analizowanemu obrazowi. Segmentację możemy uznać za proces podziału R na n podobszarów R_1, R_2, \dots, R_n takich że:

- a. $\bigcup_{i=1}^n R_i = R$
- b. R_i – składa się z połączonych ze sobą pikseli,
- c. $R_i \cap R_j = \emptyset$ dla wszystkich i i $j, i \neq j$,
- d. $Q(R_i) = \text{TRUE}$ dla $i = 1, 2, \dots, n$
- e. $Q(R_i \cup R_j) = \text{FALSE}$ dla każdych sąsiednich R_i i R_j .

gdzie: symbole \cup i \cap oznaczają odpowiednio sumę i iloczyn zbiorów, a Q jest pewnym predykatem.

Punkt a. oznacza, że segmentacja musi być kompletna tj. każdy piksel powinien zostać przyporządkowany do jakiegoś zbioru. Punkt b. oznacza, że piksele w ramach jednego podobszaru muszą być ze sobą połączone (na zasadzie sąsiedztwa 4 lub 8 punktowego). Punkt c. oznacza, że dowolne różne podobszary muszą być rozłączne. Punkt d. oznacza, że wszystkie piksele będące w ramach jednego podobszaru muszą spełniać pewną własność. Przykładowo może to być ten sam lub podobny odcień szarości. Punkt e. oznacza, że dwa sąsiednie podobszary muszą być różne w sensie predykatu Q (inaczej powinny zostać uznane za ten sam podobszar).

9.4 Segmentacja przez rozrost obszaru

Pomysł jest następujący. Wybieramy (jak? – o tym później) piksele startowe (ang. *seed*) i od nich zaczynamy segmentację. Odbywa się ona na zasadzie sprawdzania czy sąsiednie piksele (sąsiedztwo 4 lub 8 punktowe) są podobne do centralnego pod względem jakieś cechy (predykatu

Segmentacja obrazów

54

Q). Jeśli tak to oznaczane są jako należące do tej samej klasy co piksel centralny. Ponadto stają się one kolejnymi punktami startowymi metody. Zatem procedura ma charakter rekurencyjny.

Wybór punktów startowych może być podyktowany charakterem problemu (przykładowo wiemy gdzie na pewno zaczynają się obiekty). W ogólnym przypadku trzeba założyć, że pikselem startowym może być każdy piksel, co oczywiście wpływa na złożoność metody.

Kolejnym problemem jest wybór kryterium stopu tj. kiedy nasza procedura rekurencyjna ma się zakończyć. Dla danego podobszaru będzie to moment, kiedy nie istnieją już piksele, które można do niego dołączyć.

Warto w tym miejscu zwrócić uwagę, że stosowanie "sztywnego" warunku – np. różnica jasności pomiędzy pikselem centralnym, a analizowanym jest mniejsza niż 5 – może często dać niepożądane wyniki, gdyż nie uwzględnia pewnych globalnych właściwości. Przykładowo, może się okazać, że jeśli na obrazie występuje niewielki gradient to za należące do tego samego obszaru uznane zostaną piksele o zupełnie różnych jasnościach. Możliwa jest też sytuacja odwrotna. Duże zróżnicowanie wartości na obrazie spowoduje zbyt duże "poszarpanie" wykrytych obszarów.

Jednym z możliwych rozwiązań jest uzależnienie kryterium podobieństwa (predykatu Q) od własności obrazu np. średniej jasności w obrębie danego obszaru. Można również dodać inne kryteria np. kształt podobszaru itp.

Uwaga. Pojęcie segmentacja przez rozrost to pewna koncepcja podejścia do segmentacji, a nie konkretna metoda. Na etapie projektowania algorytmu należy skupić się na konstrukcji kryterium podobieństwa (tj. co i jak ma być ze sobą porównywane) oraz ewentualnym uzupełnianiu metody o dodatkowe kryteria.

Zadanie: zaprojektować system segmentacji wybranej struktury na obrazie MRI (np. stawu kolanowego). Punkt startowy wyznaczany będzie "ręcznie" (poprzez kliknięcie na obrazie).

1. Otwórz program Matlab. Ustal ścieżkę *Current Directory* na swój własny katalog roboczy.
Utwórz nowy m-plik (*New Script*) lub (*New->Script*). Na początku wykonaj polecenia
`close all; clearvars; clc;`
2. Wczytaj obraz *knee.png* – obraz MRI stawu kolanowego. Wyświetl go. Założmy, że chcemy dokonać segmentacji górnej kości. Przyjęliśmy, że punkt startowy metody wyznaczany będzie w sposób ręczny. Do pobrania położenia kurSORA myszy na ekranie służy funkcja *ginput*. Jej argumentem jest liczba punktów do pobrania. My potrzebujemy 1. Uwaga 1. Pobrane współrzędne należy zaokrąglić (*floor* lub *round*). Uwaga 2. Dla potrzeb testów dobrze jest wpisać punkt startowy na "sztywno". Pozwoli to uniknąć dość irytującego klikania przy każdym uruchomieniu skryptu. Uwaga 3. Proszę nie zapomnieć o konwersji obrazu na typ *double*.
3. Metodę zaimplementujemy z wykorzystaniem stosu. Uwaga. Podany poniżej opis jest tylko jedną z możliwych realizacji (niekoniecznie najlepszą). Na początek tworzymy dwie macierze o rozmiarach takich jak analizowany obraz. W jednej będziemy zapisywać odwiedzone lokalizacje (*visited*), a w drugiej rezultaty segmentacji (*segmented*). Tworzymy też stos tj. tablicę o np. 10000 wierszy i dwóch kolumnach (współrzędne) – *stack*. Do stosu dodajemy wskaźnik na element na górze (*iStack*). Obie macierze i stos tworzymy wypełnione zerami (funkcja *zeros*).
4. W pierwszym kroku metody na stos odkładamy współrzędne wybranego przez użytkownika piksela. Oznaczamy go również jako odwiedzony (macierz *visited*) i zaliczony do obiektu (macierz *segmented*).
5. Pozostałe działanie odbywać się będzie w pętli *while*, której warunkiem stopu jest obecność elementów na stosie. W iteracji należy pobrać współrzędne piksela ze stosu. Następnie sprawdzamy, czy dla tego piksela można określić kontekst o rozmiarze 3×3 tj. czy ma wszystkich sąsiadów. Uwaga. Przyjmujemy tutaj uproszczenie – nie segmentujemy brzegu obrazka (ramki o szerokości 1 piksela).

9.4 Segmentacja przez rozrost obszaru

55

6. W kolejnym kroku rozpisyjemy pętlę po otoczeniu 3×3 (`x2 for`). Wewnątrz obliczamy odległość pomiędzy pikselem centralnym, a każdym z kontekstu. Przyjmijmy, że będzie to moduł z różnicy jasności. Jeśli wartość modułu będzie mniejsza od zdefiniowanego progu (proszę przyjąć jako początkową wartość 4) oraz rozpatrywany piksel nie był wcześniej odwiedzany to oznaczamy go jako należący do obiektu oraz jego współrzędne "odkładamy" na stosie. Uwaga. Pierwsza część warunku logicznego to nasz predykat Q. Za piksele podobne uznajemy takie, których różnica w jasności jest mniejsza niż zadany próg. Niezależnie od wyniku testu oznaczamy piksel jako odwiedzony (żeby wielokrotnie nie analizować tych samych lokalizacji).
7. Poza pętlą `while` proszę wyświetlić rezultat segmentacji. Czy wyniki są poprawne? Proszę poeksperymentować z wartością progu.
8. Przykład ukazuje wspomniany wcześniej problem z "globalnym" podejściem do predykatu Q. Jeśli próg będzie mały, to wyznaczymy jedynie niewielki fragment kości. Natomiast zwiększenie progu skutkuje segmentacją nadmiarową. Mówiąc kolokwialnie, na obrazie znajdzie się "ścieżka" po której możliwe jest przejście od obszaru jasnego do ciemnego nie "łamiąc" progu odległości pomiędzy sąsiednimi pikselami.
9. Aby zaradzić powyższemu problemowi, można za kryterium podobieństwa przyjąć, nie różnicę jasności względem piksela centralnego, a od globalnie wyznaczonego i aktualizowanego progu. W najprostszym przypadku może to być średnia jasność w wyznaczonym obszarze. W celu implementacji mechanizmu wystarczy dodać dwie zmienne: średnią (mV) oraz licznik pikseli uznanych za należące do obiektu (nS). Przy każdym zdjęciu ze stosu licznik jest zwiększany o 1. Aktualizacja średniej następuje na podstawie równania:

$$mV_n = \frac{mV_{nS-1}(nS - 1) + I}{nS} \quad (9.1)$$

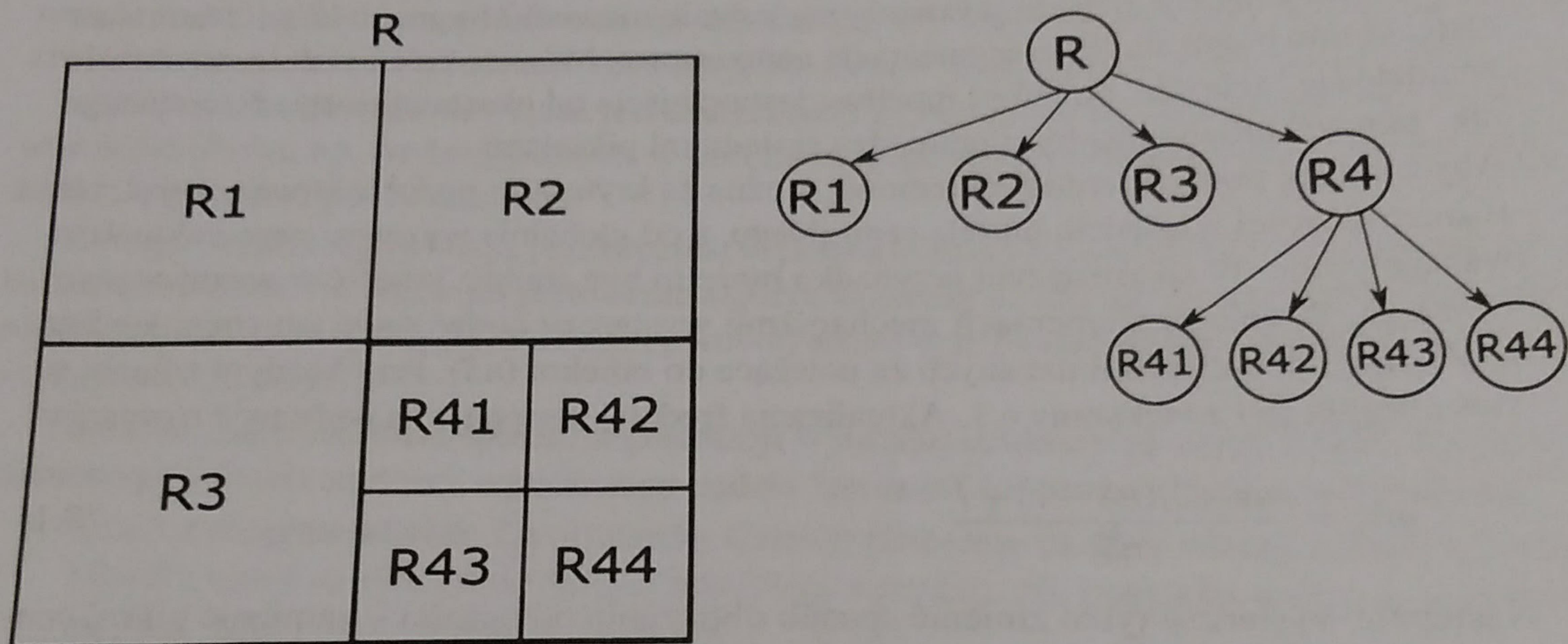
Następnie wystarczy tylko zmienić sposób obliczania odległości – zamienić piksel centralny na wartość średnią. Proszę spróbować jak działa metodą z taką modyfikacją. Proszę się liczyć z koniecznością zwiększenia progu (nawet dość znaczną).

10. Poprawić działanie metody może również dodanie filtracji uśredniającej np. filtrem Gaussa. Realizacja: `imfilter` i `fspecial`.
11. Możliwa jest również wizualizacja działania algorytmu "na bieżąco". W tym celu wynik należy wyświetlić na obrazie w pętli `while` np. `Icopy(segmented>0) = 255;` i wykorzystać polecenie `drawnow`. `Icopy` to kopia obrazu wejściowego. Oczywiście należy się liczyć ze znacznym spowolnieniem działania.

9.5 Segmentacja przez podział i łączenie

Opisaną powyżej procedurę segmentacji przez rozrost można określić jako podejście z dołu do góry, od szczegółu do ogółu (ang. *bottom-up*) – zaczynamy od pojedynczego piksela i dochodzimy do całego obrazu lub jego fragmentu. Segmentację można również zacząć od całego obrazka i dokonywać jego podziału na fragmenty, które ew. mogą zostać później połączone. Takie podejście można określić jako od góry do dołu, od ogółu do szczegółu (ang. *top-down*).

Niech R oznacza cały obraz, a Q wybrany predykat. Sprawdzamy, czy zachodzi zależność $Q(R) = \text{TRUE}$. Zwykle tak nie jest, gdyż to oznaczałoby obecność na obrazie tylko jednego obszaru o "spójnych" właściwościach, czyli segmentacja w takim przypadku byłaby zbędna. Jeśli $Q(R) = \text{FALSE}$ dzielimy obszar R na podobszary. Zwykle stosuje się tutaj podział na cztery jednakowe kwadratowe podobszary. Zilustrowano to na rysunku 9.1.



Rysunek 9.1: Przykład Podziału obrazu na kwadratowe podobszary. Źródło: opracowanie własne.

Podział ten można również opisać w formie drzewa, gdzie każdy wierzchołek może mieć dokładnie 4 następcy. W takim ujęciu korzeń to cały obraz. Dla nowo powstałych podobszarów R_1, R_2, R_3, R_4 sprawdzany jest predykat Q . Jeśli nie jest on spełniony to następuje dalszy podział – w przykładzie na $R_{41}, R_{42}, R_{43}, R_{44}$. Procedurę kontynuujemy do momentu kiedy wszystkie podobszary będą spójne. Uwaga. Zwykle określa się minimalny rozmiar podobszaru (większy niż 1 pikselm np. blok o rozmiarze 8×8).

W wyniku procedury otrzymujemy podział obrazka, przy czym często się może zdarzyć, że leżące obok siebie podobszary będą spełniać predykat tj. $Q(R_i \cup R_j) = \text{TRUE}$ dla sąsiadujących R_i i R_j . Jest to sprzeczne z podpunktem e. w warunkach segmentacji. Zjawisko to eliminowane jest w procedurze łączenia. Dla każdego z podobszarów sprawdza się jego sąsiadów i jeśli spełniony jest warunek $Q(R_i \cup R_j) = \text{TRUE}$ to następuje połączenie. Procedura kontynuowana jest do momentu, kiedy niemożliwe jest dalsze łączenie.

Dla metody możliwe są również uproszczenia. Przykładowo osłabia się warunek na łączenie z $Q(R_i \cup R_j) = \text{TRUE}$ na $Q(R_i) = \text{TRUE} \text{ i } Q(R_j) = \text{TRUE}$, czyli warunki muszą być spełnione dla każdego podobszaru osobno, ale już nie dla połączenia.

Uwaga. Podobnie jak dla segmentacji przez rozrost, segmentacja przez podział i łączenie to też tylko pewna koncepcja (pewne ramy) algorytmu. Stworzenie konkretnego rozwiązania jest zadaniem dla osoby realizującej dany system wizyjny i wymaga analizy konkretnego problemu.

Uwaga. Podany poniżej sposób implementacji nie jest ani jednym z możliwych, ani optymalnym obliczeniowo. Ma on za zadanie tylko zilustrować koncepcję tej metody segmentacji.

1. Utwórz nowy m-plik (*New Script*) lub (*New->Script*). Na początku wykonaj polecenia
`close all; clearvars; clc;`
2. Wczytaj obraz *umbrella.png*. Wyświetl go. Założmy, że chcemy dokonać segmentacji poszczególnych fragmentów kolorowego parasola. Nasz algorytm opierać się będzie na podejściu podziału i łączenia. Jako kryterium podziału zastosujemy "jednorodność" danego obszaru, którą można opisać poprzez odchylenie standardowe. Przy scalaniu będziemy brać pod uwagę średni ocień koloru tj. łączyć podobszary o zbliżonym odcieniu.
3. W pierwszym etapie należy dokonać konwersji do przestrzeni barw HSV (funkcja `rgb2hsb`). Następnie wybrać tylko składową H (tj. odcień). Proszę nie zapomnieć o konwersji obrazka na typ `double`.
4. Procedurę podziału wygodnie jest zrealizować w formie rekurencji. Utwórz funkcję np. `split`, która jako argumenty przyjmować będzie obraz oraz cztery współrzędne analizowanego podobszaru. W pierwszym kroku należy obliczyć średnią i odchylenie standardowe dla rozważanego fragmentu obrazu (funkcje `mean` i `std`). Uwaga. Wcześniej obraz należy zamienić na wektor np. `I (:)`.
5. Następnie sprawdzamy czy odchylenie jest mniejsze niż zadany przez nas próg (np. 0.05) oraz czy nie osiągnęliśmy limitu podziału (np. bok kwadratu 8 pikseli). Oba potrzebne progi możemy zrealizować za pomocą zmiennych globalnych dostępnych w Matlabie. Deklaruje się je jako global np. `global sTh`. Powyższa linijka powinna się znaleźć w każdym pliku w którym używana jest zmienna. Potrzeby rozmiar kwadratu obliczany na podstawie współrzędnych (przekazanych do funkcji jako argumenty).
6. Jeśli podobszar nie jest jednorodny i nie osiągnęliśmy minimalnego rozmiaru podobszaru to dokonujemy podziału na cztery części. Wyznaczamy rozmiar aktualnego podobszaru (na podstawie jego współrzędnych). Z ich wykorzystaniem otrzymujemy cztery identyczne podobszary (I_1, I_2, I_3, I_4). Następnie dla każdego z nich wywołujemy funkcję `split` – rekurencja. Najtrudniejsze jest odpowiednie podanie współrzędnych. Mają to być **rzeczywiste** współrzędne podobszaru we współrzędnych globalnych (obrazu w pełnej rozdzielczości). Podpowiedź. Trzeba "odpowiednio" wykorzystać współrzędne podobszaru przed podziałem (tj. argumenty funkcji) oraz rozmiar podobszaru. Należy zwrócić uwagę na to, aby nie wystąpił błąd przesunięcia o 1 tj. współrzędna była większa/mniejsza o 1 od rzeczywistej.
7. Jeśli podobszar jest jednorodny lub nie możemy już dalej prowadzić podziału to:
 - zapisujemy indeks danego podobszaru: `segRes (yS:yE, xS:xE) = index;`. `segRes` to globalna macierz o rozmiarze takim jak obraz wejściowy, a `yS, yE, xS, xE` to współrzędne podobszaru. W macierzy `segRes` zapisywane będą jednorodne obszary. Zmienna `index` jest globalnym licznikiem podobszarów. Powinna zostać zainicjowana wartością 1 przed pierwszym wywołaniem funkcji `split`. Następnie, każdorazowo po przypisaniu należy ją inkrementować.
 - zapisujemy średnią podobszaru. Wykorzystujemy macierz globalną `MRes` oraz kod podobny do opisanego powyżej. Wartości te wykorzystamy przy etapie scalania.

Uwaga. Podany sposób zapisu wyników podziału jest dość nieefektywny. Następuje powielenie tej samej informacji. Lepszym pomysłem byłoby wykorzystanie podejścia opartego o grafy, jednak jest ono trudniejsze do "szymbkiej" implementacji.

8. Drugi etap to łączenie. Idea jest nieco zbliżona do segmentacji przez rozrost. Wybieramy dany podobszar i analizujemy sąsiednie podobszary. Jeśli są one podobne tołączamy je do aktualnie rozpatrywanego. Za kryterium podobieństwa przyjmujemy niewielką różnicę w uśrednionym odcieniu barwy (składowa H).
9. Procedurę realizujemy wewnątrz pętli `while`. Warunkiem jej stopu jest przekroczenie przez licznik (np. `i`) wartości `index`, co oznacza, że przeanalizowane zostały wszystkie

znalezione w pierwszym etapie podobszary. W pierwszym kroku "wycinamy" maskę pikseli o rozpatrywanym indeksie tj. $IB = segRes == i$; . Następnie sprawdzamy, czy maska zawiera elementy niezerowe tj. czy nie jest "pusta". Taki przypadek może zajść jeśli pewien podobszar został dołączony do innego i zmienione zostały jego indeksy. Dla "pustej" maski inkrementujemy licznik i i przechodzimy do następnej iteracji – continue.

10. Dla rozpatrywanego podobszaru znajdujemy współrzędną lewego górnego rogu. Można to zrobić z wykorzystaniem składni: $[yF, xF] = find(IB, 1, 'first');$ – znajduwanie współrzędnych pierwszego niezerowego elementu.
11. Następnie należy znaleźć sąsiadów rozpatrywanego obszaru. Można wykorzystać dylatację maski IB (przypomnienie – `imdilate`) z elementem strukturalnym w postaci kwadratu o rozmiarze 3×3 `strel('square', 3)`. Następnie od maski po dylatacji odejmujemy maskę oryginalną. Otrzymujemy "ramkę", którą wykorzystujemy do "wycięcia" (mnożenie punktowe) fragmentu z macierzy $segRes$. Z tego fragmentu wybieramy elementy niezerowe (funkcja `nonzeros`) i unikalne (tj. eliminacja duplikatów) – `unique`.
12. Mając wektor z indeksami sąsiadów realizujemy pętlę w której dla każdego sąsiada:
 - wycinamy opowiadającą mu maską i znajdujemy jej lewy górny róg – w sposób analogiczny do opisanego powyżej,
 - sprawdzamy, czy moduł z różnicy pomiędzy średnimi odcieniami barwy w dwóch analizowanych podobszarach jest mniejszy od progu (np. $5/255$). Jeśli tak to łączymy obszary tj. sąsiadowi przypisujemy indeks rozpatrywanego $segRes(IBS) = i$, gdzie IBS – maska podobszaru sąsiedniego.

Do poprawnego działania konieczna jest jeszcze flaga informująca czy nastąpiło połączenie. Jej wartość ustawiamy na 0 przed pętlą po sąsiednich podobszarach. Jeśli wystąpiło połączenie to wartość ustawiamy na jeden.

Po pętli po sąsiadach sprawdzamy flagę. Jeśli ma ona wartość 0 tj. nie nastąpiło połączenie obszarów to licznik i jest inkrementowany.

13. Uwaga. Obliczenia mogą chwilę trwać. Można wykonać wizualizację działania (wyświetlenie wyników wewnątrz pętli), ale trzeba się liczyć ze spowolnieniem działania aplikacji.
14. Na koniec zastosujemy jeszcze dwie proste metody filtracji wyników. Po pierwsze wyeliminujemy obszary o rozmiarze mniejszym niż zadany (np. 100 pikseli). W tym celu wyznaczamy wektor unikalnych indeksów: $U = unique(segRes)$; . Następnie implementujemy pętlę `for` po tych indeksach. Wewnątrz wycinamy maskę dla rozpatrywanego indeksu i obliczamy jej pole (`sum`). Jeśli jest ono mniejsze niż próg to cały podobszar "zerujemy" – $segRes(C) = 0$; , gdzie C – maska podobszaru.
15. W drugim kroku przeprowadzamy przeindeksowanie indeksów na pierwsze N liczb całkowitych. Ponownie wyznaczamy unikalne indeksy. Następnie w pętli `for` wycinamy maskę o indeksie $U(i)$, a do wyniku przypisujemy iterator pętli $segRes(C) = i$. Do wizualizacji segmentacji możemy wykorzystać funkcję `label2rgb`.
16. Proszę poeksperymentować z różnymi parametrami. Wyniki proszę zaprezentować przedziem.