

Modyfikacje algorytmu GraphCut w interaktywnej segmentacji obrazów

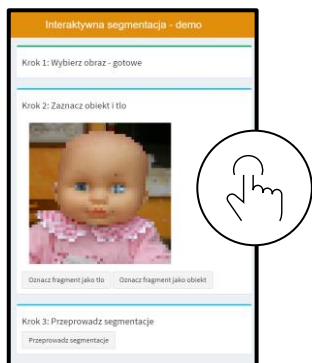
Marta Gajewska

Interaktywna segmentacja

Interaktywna segmentacja

Proces segmentacji z perspektywy użytkownika:

- wybór obrazu,
- zaznaczenie fragmentów obiektu i tła,
- obliczenia,
- otrzymanie segmentacji.

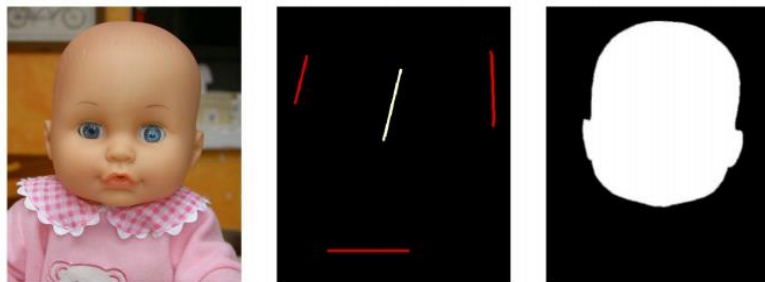


Segmentacja uzyskana w wyniku tego procesu powinna spełniać następujące **kryteria**:

- możliwe najlepsze dopasowanie do „oczekiwanej” segmentacji,
- w pełni uwzględnione informacje od użytkownika,
- brak zbędnych nieciągłości.

Implementacja

Algorytm przetestowano na **zbiorze danych** przygotowanych przez z Visual Geometry Group w Department of Engineering Science, University of Oxford. Zawiera on **151 obrazów**, wraz z zaznaczonymi przez fragmentami obiektów i tła oraz oczekiwanymi segmentacjami. Przykładowy zestaw przedstawiono poniżej.



Opisywany w tej pracy algorytm GraphCut zaimplementowano **w języku R**. Wykorzystano do tego następujące **pakiety**:

- imager (wczytywanie obrazów),
- dplyr (przetwarzanie danych),
- mixtools (algorytm EM),
- igraph (grafowa reprezentacja obrazu),
- maxflow (algorytm max-flow, znalezienie minimalnego podziału grafu),
- R Shiny (zbudowanie aplikacji),
- i inne.

Algorytm Graph Cut w interaktywnej segmentacji obrazów

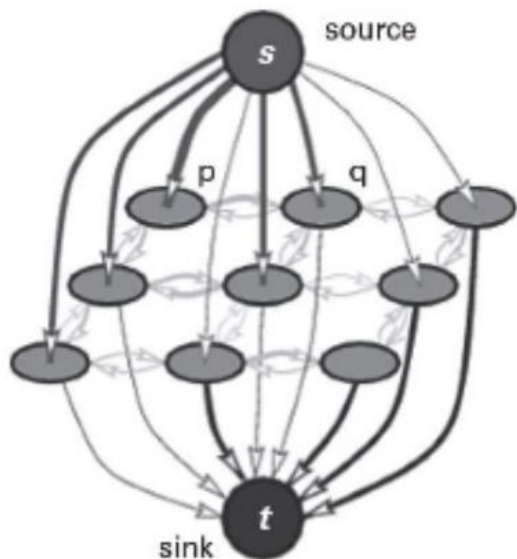
Przedstawienie obrazu w formie grafu

Wierzchołki:

- każdy z pikseli,
- dwa wierzchołki specjalne: źródło i ujście

Krawędzie:

- źródło – piksele,
- piksele – piksele,
- piksele – ujście



Określenie pojemności krawędzi

Krawędzie piksel – piksel: podobieństwo między pikselami:

$$c(v_j \rightarrow v_i) = \lambda_1 + \lambda_2 * \exp(-\beta \|f_{color}(v_i) - f_{color}(v_j)\|^2)$$

Krawędzie źródło – piksel i piksel – ujście: gęstość estymatora rozkładu kolorów w punkcie odpowiadającym kolorowi danego piksela:

$$c(s \rightarrow v_i) = h_o(f_{color}(v_i))$$

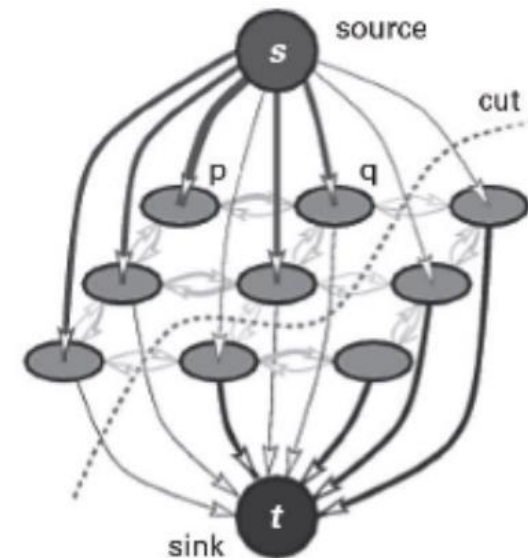
Modyfikowanie pojemności krawędzi to obszar pozwalający na najwięcej modyfikacji algorytmu

Znalezienie najlepszego podziału grafu

Szukamy przecięcia grafu, które wiąże się z najmniejszym kosztem (koszt = suma pojemności przeciętych krawędzi)

$$(\hat{s}_o, \hat{s}_t) = \arg \min_{\text{poprawne podziały}(\hat{S}_o, \hat{S}_t)} k((\hat{S}_o, \hat{S}_t)).$$

Dzięki twierdzeniu Forda – Fulkersona można sprowadzić ten problem do zagadnienia znalezienia maksymalnego przepływu.



Metody estymowania rozkładów kolorów – rodzaje rozkładów

Rozkłady kolorów obiektu i tła są estymowane na podstawie kolorów pikseli zaznaczonych przez użytkownika i przypisanych do jednej z klas. Do ich modelowania możemy używać różne rodzaje rozkładów. Standardowa metoda, to użycie rozkładu normalnego (1). W omawianej pracy zaproponowano dwie inne możliwości (2 i 3).

1 Pojedynczy rozkład normalny

Używane rozkłady

$$X_o \sim \mathcal{N}(\mu_o, \Sigma_o^2))$$

$$X_t \sim \mathcal{N}(\mu_t, \Sigma_t^2))$$

Estymacja parametrów

- Używamy estymatorów nieobciążonych, obliczonych według następujących formuł:

$$\hat{\mu}_o = \frac{1}{\|Z_o\|} \sum_{v \in Z_o} f_{color}(v)$$

$$\hat{\Sigma}_o = \frac{1}{\|Z_o\| - 1} \sum_{v \in Z_o} (f_{color}(v) - \hat{\mu}_o)(f_{color}(v) - \hat{\mu}_o)^T$$

2 Mieszanka rozkładów normalnych (3 rozkłady - obiekt / 5 rozkładów – tło)

Używane rozkłady

$$X_o \sim \sum_{i=1}^{k_o} \phi_{i,o} \mathcal{N}(\mu_{i,o}, \Sigma_{i,o})$$

$$X_t \sim \sum_{i=1}^{k_t} \phi_{i,t} \mathcal{N}(\mu_{i,t}, \Sigma_{i,t})$$

Estymacja parametrów

- $k_o = 3$ i $k_t = 5$ – wyznaczone wspólnie jako optymalny zestaw dla tego zbioru na podstawie eksperymentów
- pozostałe parametry – algorytm EM

3 Mieszanka rozkładów normalnych z dynamicznym doбором parametrów

Używane rozkłady

$$X_o \sim \sum_{i=1}^{k_o} \phi_{i,o} \mathcal{N}(\mu_{i,o}, \Sigma_{i,o})$$

$$X_t \sim \sum_{i=1}^{k_t} \phi_{i,t} \mathcal{N}(\mu_{i,t}, \Sigma_{i,t})$$

Estymacja parametrów

- k_o i k_t – wyznaczone niezależnie dla każdego z obrazów, dopasowano rozkłady dla każdego $k = 1, \dots, 15$ i wybrano optymalne na podstawie kryterium BIC
- pozostałe parametry – algorytm EM

Algorytm EM – iteracyjny algorytm służący do znajdowania parametrów (lokalnie) największej wiarygodności. Używany często w sytuacjach, gdy trudno jest bezpośrednio maksymalizować funkcję log-wiarygodności.

Metody estymowania rozkładów kolorów – przykładowe wyniki

Jak obrazuje poniższy przykład, metoda 1 ma duże problemy w przykładach, gdzie segmenty składają się z fragmentów o różnorodnej strukturze. Metoda 2 i 3 radzą sobie lepiej, i choć czasami zdarzają się duże różnice w osiągniętych wynikach (tak jak w tym przykładzie), to zazwyczaj są zwracane przez nie segmentacje są dość podobne.

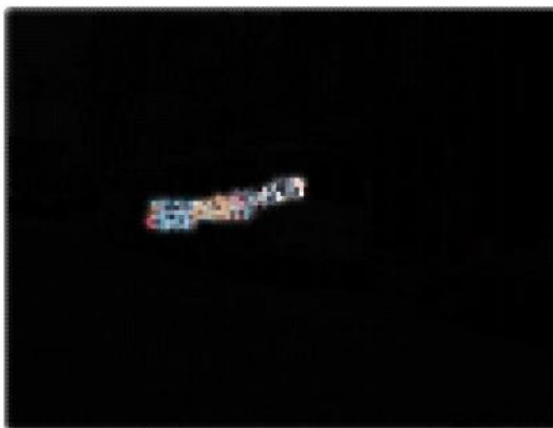
Przykładowy obraz:
bus_2007_005262

Obraz i zaznaczone obszary



1

Pojedynczy rozkład normalny



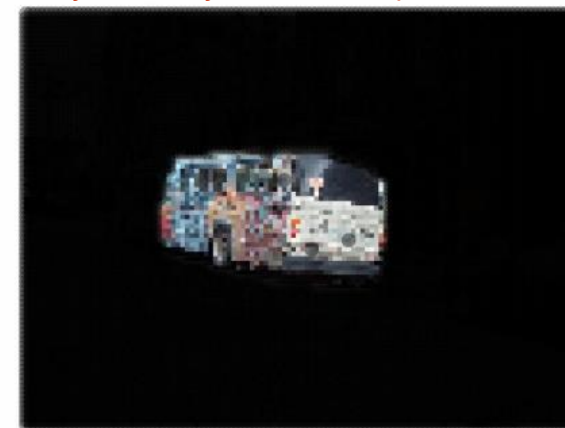
2

Mieszanka rozkładów normalnych
(3 rozkłady - obiekt / 5 rozkładów - tło)



3

Mieszanka rozkładów normalnych
z dynamicznym doбором parametrów



Metody estymowania rozkładów kolorów – porównanie jakości segmentacji

Wszystkie 3 modyfikacje algorytmu zostały przetestowane na wspomnianym wcześniej zbiorze obrazów. Dla każdej z otrzymanych segmentacji obliczono 5 wybranych metryk jakości. Poniższe tabele przedstawiają średnie wartości wybranych metryk otrzymane dzięki 3 wersjom algorytmu segmentacyjnego.

1	Pojedynczy rozkład normalny		2	Mieszanka rozkładów normalnych (3 rozkłady - obiekt / 5 rozkładów - tło)		3	Mieszanka rozkładów normalnych z dynamicznym doбором parametrów	
	dokładność	0.82		dokładność	0.88		dokładność	0.87
	zbal. dokładność	0.72		zbal. dokładność	0.78		zbal. dokładność	0.78
	precyzja	0.92	\leq	precyzja	0.93	\approx	precyzja	0.93
	wrażliwość	0.86		wrażliwość	0.92		wrażliwość	0.91
	specyficzność	0.58		specyficzność	0.63		specyficzność	0.64

Test Wilcoxona dla metody 1 i 2

Metryka	Różnica	P-wartość
dokładność	+0.06	<0.001
zbal. dokładność	+0.06	<0.001
precyzja	+0.01	0.1
wrażliwość	+0.06	0.03
specyficzność	+0.05	0.06

Wyniki uzyskane przy użyciu klasycznej estymacji – metodą 1 – są wyraźnie gorsze niż wyniki otrzymane dla metod używających mieszanek rozkładów normalnych. W celu sprawdzenia istotności statystycznej różnic pomiędzy metodą 1 i 2 przeprowadzono test Wilcoxona, którego wyniki prezentuje poniższa tabela. Metoda ustalania parametrów w mieszkankach rozkładów (globalnie, przy metodzie 2, oraz dynamicznie, przy metodzie 3) nie powoduje znaczących różnic w wynikach.

Wnioski

1 Pojedynczy rozkład normalny

- Najszybsza z metod
- Gorsza jakość segmentacji

2 Mieszanka rozkładów normalnych (3 rozkłady - obiekt / 5 rozkładów - tło)

- Średnia prędkość
- Dobra jakość segmentacji
- Wymaga zbioru testowego – dobra dla obrazów ustalonego typu (np. obszary nowotworowe na zdjęciach rentgenowskich)


3 Mieszanka rozkładów normalnych z dynamicznym doбором parametrów

- Najwolniejsza z metod
- Dobra jakość segmentacji
- Ustalenie parametrów nie wymaga zbioru testowego – dobra dla obrazów nieznanego typu

Przykład

Krok 1: Wybierz obraz - gotowe


Krok 2: Zaznacz obiekt i tło



Krok 3: Przeprowadź segmentację

Krok 1: Wybierz obraz - gotowe

Krok 2: Zaznacz obiekt i tło



Poczekaj...

Segmentacja w toku...

Krok 3: Przeprowadź segmentację



Wynik segmentacji:

