

Robert Świta
Zbigniew Suszyński
Katedra Systemów Multimedialnych
i Sztucznej Inteligencji
Wydział Elektroniki i Informatyki
Politechnika Koszalińska

Inicjalizacja segmentacji k-means uwzględniająca rozkład gęstości pikseli

Słowa kluczowe: FA, KKZ, k-means, kmeans++, segmentacja

Wstęp

Segmentacja jest zasadniczym etapem w rozpoznawaniu obrazów. Jednocześnie jest ona problemem NP-trudnym, w rozwiązaniu którego często posługujemy się metodami numerycznymi, takimi jak algorytm k-means. Decydujący wpływ na błąd segmentacji k-means ma początkowy wybór środków segmentów (Pena et al., 1999 [1]). Standardowym podejściem jest tu jednak zupełnie losowy ich wybór i przyjęcie statystycznie najlepszego rozwiązania. Heurystyczne metody inicjalizacji algorytmu k-means często maksymalizują odległości wzajemne środków segmentów, nie uwzględniają natomiast gęstości rozkładu wartości pikseli ze względu na dużą złożoność obliczeniową. W niniejszym artykule przedstawiono metodę inicjalizacji, która uwzględnia gęstość poprzez oszacowanie jej wartości dla każdego piksela.

1. Segmentacja k-means

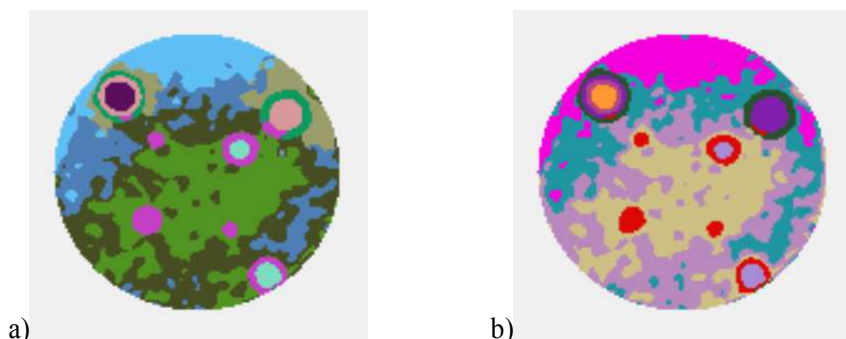
Celem segmentacji jest zmiana podziału zbioru n -elementowego na zbiór o mniejszej liczbie elementów k , minimalizująca pewną funkcję kosztu będącą błędem segmentacji. Z tego punktu widzenia segmentacja jest przekształceniem podobnym do aproksymacji, ale operującym na zbiorach. Metody segmentacji najczęściej dzieli się na metody klastrowe, które podczas łączenia pikseli biorą pod uwagę wyłącznie wartość przetwarzanego piksela i metody kontekstowe, które uwzględniają również wartości pikseli sąsiednich – z otoczenia przetwarzanego piksela. Segmentacja jest problemem NP-trudnym nawet w przypadku podziału pojedynczego obrazu na dwa segmenty. W celu szybkiego znalezienia rozwiązania w postaci minimum lokalnego wykorzystuje się najczęściej metody iteracyjne. Najpopularniejszym takim algorytmem jest klastrowa segmentacja k-means (Forgy

1965 [2]; MacQueen 1967 [3]; Lloyd 1982 [4], Wu et al., 2008 [5]). Sposób podziału obrazu na segmenty metodą k-means zależy jedynie od wartości piksela, nie uwzględniając ani jego położenia, ani wartości pikseli sąsiednich, dlatego jej wynikiem są w ogólności klastry niespójne. Liczba segmentów-klastrów k , na które obraz zostanie podzielony ustalana jest z góry. W pierwszej iteracji wybieranych jest k wartości środków segmentów i piksele przydzielane są do najbliższych im segmentów. Środki klastrów są następnie przeliczane do wartości średniej ze wszystkich pikseli należących do danego klastra i procedura przydziału jest przeprowadzana ponownie. Algorytm w każdej iteracji wykonuje zatem dwie operacje: przypisania pikseli do klastrów i korekcji wartości środków klastrów. Nietypową cechą k-means wśród algorytmów segmentacji jest możliwość dokładnego określenia błędu. Algorytm k-means minimalizuje sumę kwadratów odległości wartości pikseli od przyporządkowanych im środków segmentów.

$$error = \sum_{i=1}^P \min_{j \in \{1, S\}} (x[i] - c[j])^2 \quad (1)$$

Algorytm k-means jest zawsze zbieżny, gdyż średnia arytmetyczna jest estymatą błędu, ale iteracje można również przerwać w momencie uzyskania wystarczająco małego błędu segmentacji. Jeśli funkcją kosztu byłaby suma wartości bezwzględnych odległości pikseli od przypisanych środków, w drugim kroku każdej iteracji należałoby korygować środki segmentów do wartości mediany z pikseli należących do klastra (k-median).

Algorytm k-means można również wykorzystać do klasteryzacji danych wielowymiarowych. Procedura segmentacji sekwencji obrazów ma za zadanie wyznaczyć zbiory składające się z podobnych charakterystyk wartości pikseli w funkcji numeru kadru. Segmentacja k-means sekwencji obrazów może być zastosowana w przypadku, gdy, dla kolejnych kadrów, wartości pikseli należących do tego samego klastra zmieniają się w podobny sposób. Sytuacja taka ma miejsce m.in. dla sekwencji obrazów termicznych uzyskanych techniką termowizji aktywnej podczas rejestracji odpowiedzi temperaturowej próbki na pobudzenie cieplne o określonym widmie. W każdym kadrze sekwencji uzyskujemy wówczas różne obrazy, ale grupy pikseli z obszarów o podobnych parametrach cieplnych będą przyjmować podobne wartości. Metoda k-means porównuje ze sobą wektory wartości pikseli w kolejnych kadrach, czyli charakterystyki pikseli. Po segmentacji, dla każdego klastra, otrzymujemy jego środek w postaci wektora – średniej charakterystyki klastra w funkcji kadru. Segmentacji mogą być poddane zarówno sekwencje w dziedzinie czasu jak i częstotliwości (Rys. 1).



Rys. 1. Obraz segmentacyjny sekwencji obrazów termicznych w dziedzinie czasu a) i sekwencji amplitudowej b)

Dla sekwencji obrazów definiujemy błąd bezwzględny segmentacji na S obiektów jako sumę odchyłeń wartości wszystkich P pikseli od przyporządkowanych im środków segmentów dla wszystkich K kadrów:

$$error = \sum_{i=1}^P \min_{j=1..S} \sum_{k=1}^K (x[k, i] - c[k, j])^2 = \sum_{i=1}^P \min_{j=1..S} \|x[i] - c[j]\|^2 \quad (2)$$

Warto podkreślić, że algorytm k-means łatwo zaimplementować z wykorzystaniem przetwarzania równoległego pikseli w celu przyspieszenia obliczeń.

Algorytm k-means jest szczególnym przypadkiem segmentacji FCM (Fuzzy C-Means) [6], w której zakłada się, że piksele mogą należeć do segmentów zgodnie z określoną funkcją wagową μ . Zakładając, że S oznacza liczbę segmentów a r stopień rozmycia przynależności pikseli do segmentów:

$$\mu(x[i], c[j]) = \frac{\|x[i] - c[j]\|^{-2/(r-1)}}{\sum_{k=1}^S \|x[i] - c[k]\|^{-2/(r-1)}} \quad (3)$$

W każdej iteracji środki klastrów modyfikowane są przez wyznaczenie środka ciężkości segmentów:

$$c[j] = \frac{\sum_{i=1}^P \mu(x[i], c[j]) x[i]}{\sum_{i=1}^P \mu(x[i], c[j])} \quad (4)$$

Jeśli tylko dla jednego segmentu waga jest niezerowa, algorytm degeneruje się do metody k-means. Algorytm FCM w ogólności znajduje lepsze rozwiązanie, jest jednak znacznie bardziej złożony obliczeniowo, jest trudniej zbieżny i wymaga dobrania odpowiedniego współczynnika rozmycia r .

2. Popularne metody inicjalizacji

Algorytm k-means jest zawsze zbieżny jeśli miarą odległości wektorów jest metryka euklidesowa, ale o szybkości zbieżności i dokładności rozwiązania decyduje przede wszystkim wybór początkowy środków klastrów (Pena et al., 1999 [1]). Istnieje kilka popularnych metod inicjalizacji środków klastrów:

1. Inicjalizacja Forgy'ego [2] – przyjmuje się wartości losowe, wybrane spośród danych wejściowych, przy jednorodnym rozkładzie prawdopodobieństwa.
2. KKZ (Katsavounidis et al. [7]) – na pierwszy środek można wybrać piksel o najmniejszej lub największej wartości. Kolejne przyjmują wartości pikseli położonych najdalej od znalezionych już środków. Maksymalizując odległość pomiędzy środkami gwarantuje się dobrą separację segmentów. Metoda jest bardzo szybka, ponieważ przechowując tablicę minimalnych odległości każdego piksela od znalezionych środków wybieramy na nowy środek piksel o maksymalnej odległości i korygujemy tablicę tylko dla pikseli najbliższych znalezionemu środkowi. Wadą tej metody jest częste tworzenie segmentów z niewielkiej liczby pikseli o wartościach skrajnych, tzw. outlierów.
3. k-means++ (D.Arthur, S.Vassilvitskii [8]) - przyjmuje się wartości losowe, wybrane spośród danych wejściowych, przy rozkładzie prawdopodobieństwa proporcjonalnym do odległości od już wyznaczonych środków. W tym celu wyznaczamy sumę odległości wszystkich pikseli od znalezionych środków. Losujemy pewną odległość d z zakresu od 0 do tej sumy i znajdujemy indeks piksela, dla którego suma odległości pikseli przekroczy wartość d . Znaleziony piksel dodajemy do wektora środków. Metoda łączy w sobie zatem wykorzystanie statystyki i analizę odległości od już wyznaczonych środków.

3. Inicjalizacja uwzględniająca gęstość pikseli

Powyższe inicjalizacje algorytmu k-means nie uwzględniają jednak gęstości określonej jako rozkład wartości pikseli. Wymaga ona obliczenia wzajemnych odległości pomiędzy pikselami, czyli liczby operacji rzędu $O(n^2)$. Dla problemów wielowymiarowych w celu oszacowania gęstości można posłużyć się strukturą drzewa kd dzielącego przestrzeń danych w kolejnych wymiarach (S.Redmond i C.Heneghan [9]). O dużej gęstości piksela decydują niewielkie odległości od jego sąsiadów. Możemy tę funkcję określić zatem np. poprzez sumowanie odwrotności odległości od pozostałych pikseli (Inverse Distance Weighting Function):

$$\rho[k] = \sum_{i=1}^P \frac{1}{\|x[i] - x[k]\|^2 + \varepsilon} \quad (5)$$

Czynnik ε jest dodatni i związany z kwantyzacją danych, w wyniku której niektóre odległości pomiędzy pikselami są zerowe. Wyznaczenie tak zdefiniowanej funkcji jest bardzo kosztowne obliczeniowo, ale można jej wartość poddać oszacowaniu. Załóżmy, że odległości piksela k od pozostałych są zawsze takie same, wówczas:

$$\rho[k] = \sum_{i=1}^P \frac{1}{\|x[i] - x[k]\|^2 + \varepsilon} \cong \frac{P}{d[k] + \varepsilon} \quad (6)$$

Gęstość piksela k możemy określić korzystając z rozwinięcia w szereg Taylora funkcji $(d + 1)^{-1} = 1 - d + d^2 - d^3 + \dots$, ponieważ:

$$\rho[k] = \sum_{i=1}^P \frac{1}{\|x[i] - x[k]\|^2 + \varepsilon} = \frac{1}{\varepsilon} \sum_{i=1}^P \left(\frac{\|x[i] - x[k]\|^2}{\varepsilon} + 1 \right)^{-1} \quad (7)$$

Rozwijając również prawą stronę przybliżenia (3), otrzymujemy:

$$\begin{aligned} \frac{1}{\varepsilon} \left(P - \frac{1}{\varepsilon} \sum_{i=1}^P \|x[i] - x[k]\|^2 + \frac{1}{\varepsilon^2} \sum_{i=1}^P \|x[i] - x[k]\|^4 - \dots \right) \\ \cong \frac{P}{\varepsilon} \left(1 - \frac{d[k]}{\varepsilon} + \left(\frac{d[k]}{\varepsilon} \right)^2 - \dots \right) \end{aligned} \quad (8)$$

Z porównania pierwszych składników rozwinięć wynika, że odległość d dla piksela k powinna być równa średniej ze wszystkich odległości tego piksela względem pozostałych. Można ją szybko wyznaczyć dla wszystkich pikseli na podstawie wartości średniej \bar{x} i wariancji $\text{Var}[x]$, przy założeniu jednakowego prawdopodobieństwa przyjmowanych wartości przez piksele.

$$\begin{aligned} d[k] &= \frac{1}{P} \sum_{i=1}^P \|x[i] - x[k]\|^2 = \\ &= \frac{1}{P} \sum_{i=1}^P \|x[i] - \bar{x}\|^2 + \|\bar{x} - x[k]\|^2 = \text{Var}[x] + \|\bar{x} - x[k]\|^2 \end{aligned} \quad (9)$$

Przybliżenie będzie prawdziwe jeśli ciągi nie będą się rozbiegać zbyt szybko, tzn. dla $\varepsilon \geq d[k]$. Ponieważ czynnik ε powinien być jak najmniejszy i stały dla wszystkich pikseli, przyjmiemy za jego wartość średnią ze wszystkich odległości $d[k]$

$$\varepsilon = \frac{1}{P} \sum_{i=1}^P d[k] = \frac{1}{P} \sum_{i=1}^P \left(\text{Var}[x] + \|\bar{x} - x[k]\|^2 \right) = 2 \cdot \text{Var}[x] \quad (10)$$

Gęstość piksela możemy zatem oszacować jako:

$$\rho[k] \cong \frac{P}{3 \cdot \text{Var}[x] + \|\bar{x} - x[k]\|^2} \quad (11)$$

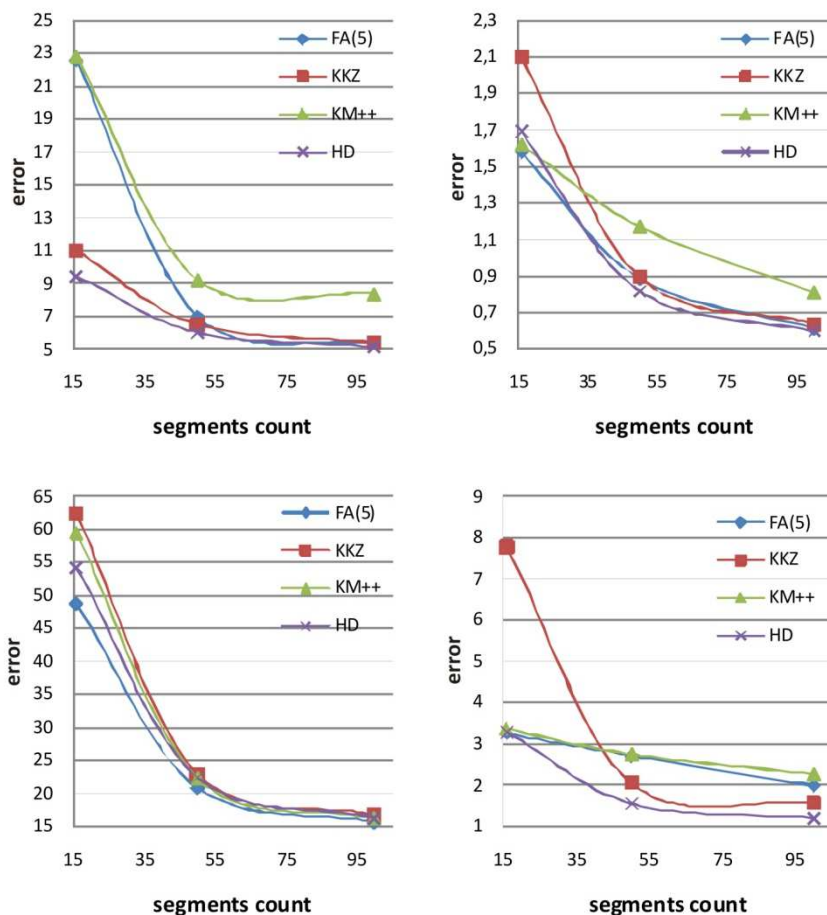
Metoda inicjalizacji uwzględniającą gęstość pikseli określona została jako High Density (HD). Na środki segmentów powinny zostać wybrane piksele o jak największej gęstości ale i jak najbardziej odległe od już wyznaczonych środków, dlatego w metodzie na pierwszy środek jest wybierany piksel o największej gęstości, natomiast pozostałe środki segmentów przyjmują wartości pikseli o maksymalnym iloczynie gęstości piksela i jego odległości od już wyznaczonych środków.

4. Porównanie metod inicjalizacji

W eksperymencie porównane zostały wyniki segmentacji czterech sekwencji obrazów uzyskanych za pomocą techniki termografii aktywnej. Wykorzystano metody inicjalizacji FA, KKZ, kmeans++ oraz nową metodę HD. W Tabeli 1 przedstawione zostały błędy wyznaczone zgodnie z (2) dla segmentacji sekwencji obrazów na 16, 50, 100 i 1024 klastrów. Rys. 1 przedstawia te wyniki w postaci wykresów, pomijając segmentację na 1024 obiektów, ze względu na podobne błędy dla wszystkich metod inicjalizacji.

Tabela 1. Błędy segmentacji przykładowych sekwencji w funkcji liczby segmentów dla różnych metod inicjalizacji

	Segments count	16	50	100	1024
Seq1	FA(5)	22,58	7,04	5,36	3,03
	KKZ	10,97	6,52	5,4	2,91
	KM++	22,83	9,22	8,37	3,09
	HD	9,44	6,02	5,15	2,95
Seq2	FA(5)	1,58	0,886	0,615	0,319
	KKZ	2,1	0,892	0,639	0,297
	KM++	1,62	1,17	0,808	0,331
	HD	1,69	0,817	0,597	0,308
Seq3	FA(5)	48,74	20,86	15,67	8,99
	KKZ	62,3	22,97	16,73	8,56
	KM++	59,48	22,36	16,17	9,08
	HD	54,2	22,39	16,31	8,83
Seq4	FA(5)	3,25	2,71	1,99	1,027
	KKZ	7,77	2,05	1,56	0,603
	KM++	3,36	2,74	2,26	1,058
	HD	3,3	1,55	1,2	0,635



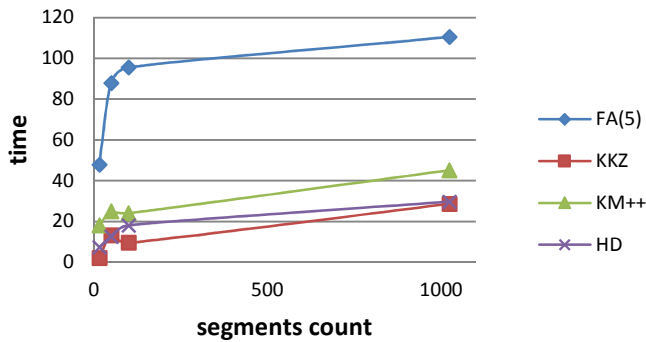
Rys. 2. Wykres błędów segmentacji sekwencji Seq1 do Seq4 dla różnych metod inicjalizacji

Środki segmentów w przypadku FA wybierane są statystycznie z najlepszej inicjalizacji losowej. W ten sposób algorytm k-means traci jednak swą największą zaletę, jaką jest szybkość znajdowania rozwiązania. W eksperymencie wybrano najlepszą segmentację losową z pięciu prób.

Metoda KKZ uniezależnia wynik segmentacji od statystyki. Kolejne środki wyznaczone są jak najdalej od znalezionych wcześniej środków, skutkując prawie równomiernym ich rozmieszczeniem.

Inicjalizacja kmeans++ stara się wykorzystać zalety inicjalizacji KKZ i FA. Ponieważ jednak, podobnie jak FA jest metodą statystyczną, wymaga w zasadzie kilku niezależnych uruchomień algorytmu, a więc kilkukrotnie dłuższych czasów przetwarzania. Również błąd segmentacji jest często wyższy niż w metodzie KKZ (szczególnie dla większej liczby segmentów). W eksperymencie, algorytm wykonywany był dla inicjalizacji kmeans++ tylko raz, stąd też błędy segmentacji przeważnie wyższe niż dla pięciu uruchomień FA.

Dla inicjalizacji HD algorytm k-means zachowuje się podobnie jak w przypadku KKZ, ale ze względu na wykorzystanie informacji o rozkładzie gęstości pikseli błąd końcowy segmentacji jest mniejszy. Liczba iteracji potrzebnych do znalezienia minimum jest również mniejsza i tym samym czas przetwarzania, mimo dodatkowych obliczeń podczas inicjalizacji, jest podobny jak w metodzie KKZ (Rys. 3).



Rys. 3. Czas segmentacji sekwencji Seq1 dla różnych metod inicjalizacji

Podsumowanie

Metoda k-means znajduje bardzo szerokie zastosowanie ze względu na szybkość i łatwość implementacji w segmentacji, kwantyzacji wektorowej i rozpoznawaniu cech. Zasadniczy wpływ na dokładność znalezionej metody ma wybór początkowy środków segmentów. W artykule przedstawiono metodę inicjalizacji, która korzystając z oszacowania gęstości pikseli, pozwala na uniezależnienie od wyboru statystycznego i uzyskanie wyników z mniejszym błędem od znanej metody KKZ. Wyniki wielokrotnej inicjalizacji metodami statystycznymi, które są standardowo używane, mogą dostarczyć podobnego rozwiązania, ale czasy przetwarzania będą wówczas wielokrotnie dłuższe.

Bibliografia

1. J. M. Pena, J. A. Lozano, P. Larranaga, *An empirical comparison of four initialization methods for the k means algorithm*, Pattern Recognition Letters 20 (10), 1027 (1999)
2. E. Forgy, *Cluster analysis of multivariate data: Efficiency versus interpretability of classifications*, Biometrics, 21, 768 (1965)
3. J. MacQueen, *Some Methods for Classification and Analysis of Multivariate Observations*, Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, 281 (1967)
4. S. P. Lloyd, *Least Squares Quantization in PCM*, IEEE Trans. Information Theory, 28, 129 (1982)
5. X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. F. M. Ng, B. Liu, P. S. Yu, Z. Zhou, M. Steinbach, D. J. Hand, D. Steinberg, *Top 10 algorithms in data mining*, Knowledge and Information Systems, 14 (1), 1 (2008)
6. J. C. Bezdek *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York (1981).
7. I. Katsavounidis, C. Kuo, Z. Zhang, *A new initialization technique for generalised Lloyd iteration*, IEEE Signal Processing Letters 1 (10), 144 (1994)
8. D. Arthur, S. Vassilvitskii, *k-means++: the advantages of careful seeding*, Proceedings of the 18th annual ACM-SIAM Symposium on Discrete Algorithms, 1027 (2007)
9. S. J. Redmond, C. Heneghan, *A method for initialising the K-means clustering algorithm using kd-trees*, Pattern Recognition Letters 28, 965 (2007)

Streszczenie

Artykuł przedstawia modyfikację inicjalizacji KKZ algorytmu k-means, uwzględniającą, oprócz wzajemnych odległości środków segmentów, również rozkład gęstości pikseli. Funkcja gęstości piksela jest sumą odwrotności odległości piksela od pozostałych i jest poddawana oszacowaniu na podstawie odległości piksela od wartości średniej i wariancji wartości pikseli. W eksperymentach segmentacji podlegały cztery różne sekwencje obrazów termicznych uzyskanych metodą termografii aktywnej. Pomimo dodatkowych obliczeń podczas inicjalizacji, metoda wykazała szybszą zbieżność algorytmu z czasami bardzo podobnymi do inicjalizacji KKZ, ale mniejszym błędem końcowym segmentacji.

Abstract

This article presents a modification for the KKZ initialization of the k-means segmentation algorithm, which, in addition to the mutual distance of segments, takes into account the density of pixels. Pixel density is expressed as a sum of the inverse of the pixel's distance to the other pixels and is subjected to estimation based on the distance from the mean and variance of the pixel values. In the experiments, four different sequences of thermal images were used, obtained using active thermography. Despite the additional calculations during initialization, method showed a faster convergence of the algorithm, with processing times very similar to the KKZ initialization, but with a lower final segmentation error.