

Metody Rozpoznawania Obrazów i Podstawy Uczenia Maszynowego

k-means

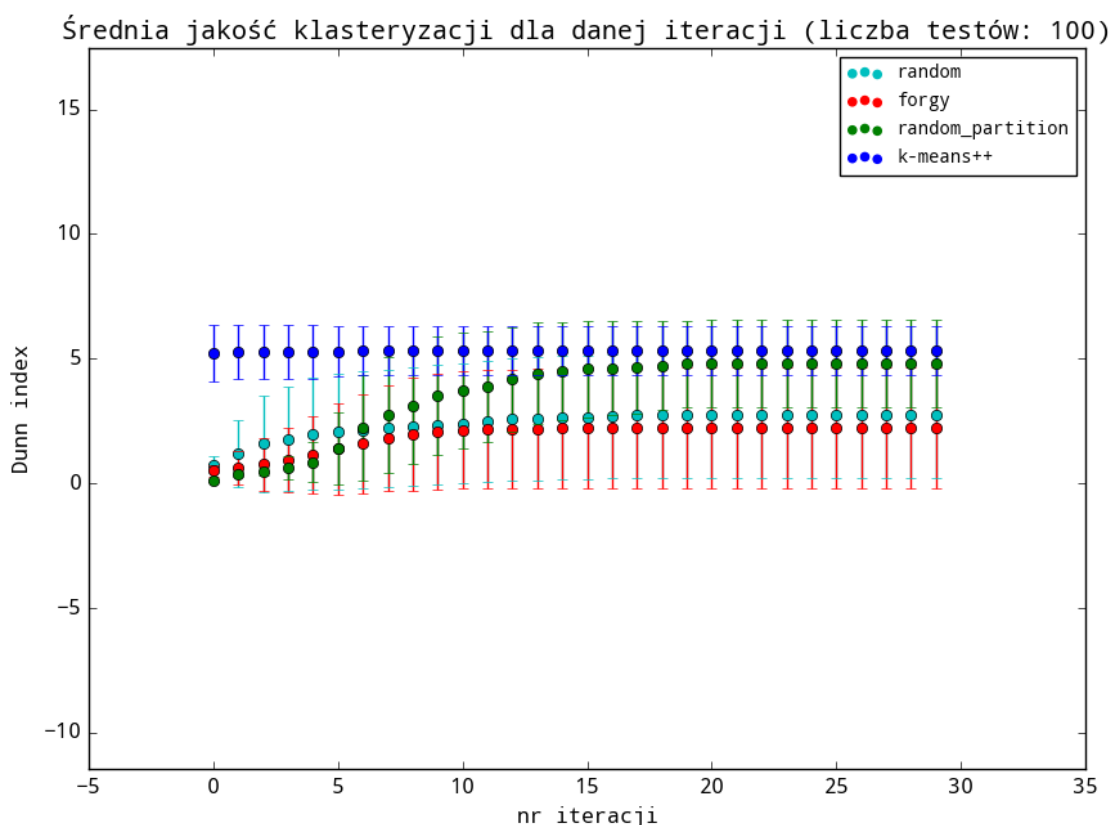
Damian Kuś

nr albumu: 267262

nr w arkuszu z punktami: 1

Część A

Jakość klasteryzacji w zależności od metody wyboru środków klastrów.



Wykres 1: Zależność między jakością klasteryzacji a numerem iteracji k-means dla różnych metod generowania początkowych punktów centralnych klastrów

Obserwacje:

W celu zbadania jakości klasteryzacji wykorzystano indeks Dunn'a o postaci:

$$D = \frac{\min_{1 \leq i \leq j \leq n} d(i, j)}{\max_{1 \leq k \leq n} d'(k)}$$

Wartość $d(i, j)$ oznacza odległość między środkami klastrów, natomiast d' średnią odległość między punktami w klastrze a środkiem tego klastra. N oznacza liczbę klastrów (w naszym przypadku przyjęto $n=9$, w skład każdego klastra wchodzi 100 próbek).

W ramach ćwiczenia wykonano 100 testów, przy czym każdy z testów składał się z 30 iteracji algorytmu k-means. Najlepsze wyniki dawało wykorzystanie metody k-means++. Zwykle umożliwiała ona znalezienie właściwego podziału już w pierwszej iteracji (stąd odpowiadający wykres ma formę funkcji stałej). Nieco gorsze wyniki uzyskano przy użyciu metody random partition (tworzenie losowych klastrów i wyznaczanie ich punktów centralnych jako średnich ze wszystkich punktów do nich należących).

Najgorszą zbieżność algorytmu zaobserwowano w przypadku dwóch najprostszych metod – random i Forgy. Zauważono, że mają one tendencję do utknięcia w minimum lokalnym, przez co mimo wykonywania kolejnych iteracji podział próbek na klastry się nie zmieniał.

Część B

Wyniki kompresji zdjęcia w formacie BMP.



Ilustracja 1: Zdjęcie oryginalne



Ilustracja 2: Zdjęcie skompresowane $k = 8$



Ilustracja 3: Zdjęcie skompresowane $k = 32$



Ilustracja 4: Zdjęcie skompresowane $k = 128$

Propozycja formatu wykorzystującego kompresję k-means

Plik składa się z trzech części:

- informacji o rozmiarze – dwie wartości heksadecymalne oddzielone spacją mówiące o szerokości i wysokości obrazu zakończone znakiem nowej linii (dla uproszczenia szacowania rozmiaru możemy przyjąć, że każda z tych wartości jest reprezentowana przez 4 znaki, co przekłada się na maksymalną wartość równą $2^{16}-1=65535$ – w praktyce powinna ona wystarczyć)
- listy kolorów dla poszczególnych klastrów – zakładamy, że informacje o kolorach są zapisywane w systemie szesnastkowym z wykorzystaniem 6 znaków np. FF00AA. Jako że liczba znaków jest stała, nie ma potrzeby, żeby oznaczenia kolorów były rozdzielone dodatkowymi znakami. Lista jest zakończona znakiem nowej linii.
- Listy kategorii dla poszczególnych punktów – zakładamy, że każdy piksel jest reprezentowany przez $\text{ceiling}(\log_{16} N)$ znaków heksadecymalnych, gdzie N to liczba używanych kolorów (odczytana na podstawie listy kolorów), przy czym dla uproszczenia zakładamy, że liczba znaków jest stała dla danego pliku. Oznaczenie takie jest liczbą odpowiadającą indeksowi koloru na liście we wcześniejszym fragmencie. Jako że w pierwszej linii pliku znajduje się informacja o rozmiarze obrazu ponownie nie ma potrzeby separowania oznaczeń kolorów dla piksela.

Przykład

10 10 (16 x 16 pikseli)
FFFF000000FF0000 (liczba znaków reprezentujących indeks koloru = $\text{ceiling}(\log_{16} 3)=1$)
0000000000000000 (znaki nowej linii dodane dla czytelności)
0011110000111100
0011110000111100
0000000000000000
0022222222222200
0002222222222000
0000022222200000
0000000000000000

Rozmiar pliku

Rozmiar pojedynczego znaku 1B

Informacja o rozmiarze: $4 + 1 + 4 + 1 = 10[B]$

Lista kolorów: $N * 6 + 1 B$

Lista pikseli: $W * H * \text{ceiling}(\log_{16} N) B$ (W – szerokość, H - wysokość)

Rozmiar całkowity: $6 N + W * H * \text{ceiling}(\log_{16} N) + 11 B$

Przykładowe wartości dla obrazka o rozdzielczości 360 X 286 (takiej, jakiej użyto w ćwiczeniu)

Uwaga: kB jest tutaj rozumiany jako $2^{10} B$

liczba kolorów	rozmiar [B]	Rozmiar [kB]	liczba znaków reprezentujących kolor
16	107	0,104	1

32	203	0,20	2
64	395	0,39	2
128	779	0,76	2
65536	393227	384,01	4
16777216	100663307	98304,01	6

Porównanie rozmiarów obrazów o 24-bitowej głębi koloru (ok. 16 mln kolorów)

Format	Rozmiar
BMP	308.9 kB
JPEG	21.4 kB
Proponowany format	96 MB

Okazuje się, że format wykorzystujący metodę k-means jest niepraktyczny ze względu na wielkość pliku dla pełnej palety kolorów. W takim przypadku etykieta przypisana pikselowi zajmuje tyle pamięci, ile potrzebne jest na podanie wartości koloru bezpośrednio, więc lista kolorów stanowi niepotrzebny narzut.