Министерство образования Республики Беларусь

Учреждение образования

БелорусскиЙ государственный университет

информатики и радиоэлектроники

Факультет компьютерных систем и сетей

Кафедра информатики

ОТЧЕТ

по лабораторной работе №6

**Кластеризация**

|  |  |  |
| --- | --- | --- |
| Студент |  | И.Ю. Фалько |
| Преподаватель |  | М. В. Стержанов |

Минск 2019

СОДЕРЖАНИЕ

[Условие лабораторной работы 3](#_Toc25744885)

[Ход выполнения 3](#_Toc25744886)

[Заключение 6](#_Toc25744887)

# Условие лабораторной работы

Набор данных **ex6data1.mat** представляет собой файл формата \*.mat (т.е. сохраненного из Matlab). Набор содержит две переменные X1 и X2 - координаты точек, которые необходимо кластеризовать.

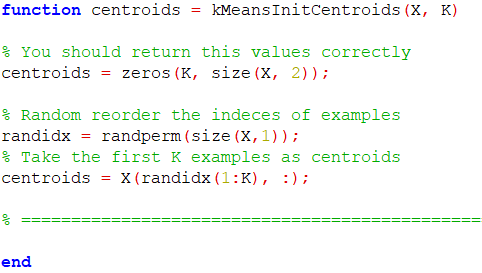
Набор данных **bird\_small.mat** представляет собой файл формата \*.mat (т.е. сохраненного из Matlab). Набор содержит массив размером (16384, 3) - изображение 128x128 в формате RGB.

# Ход выполнения

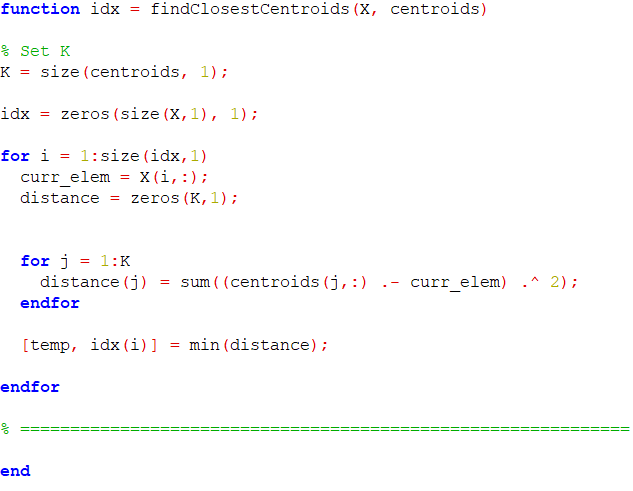
1. Загрузите данные **ex6data1.mat** из файла.



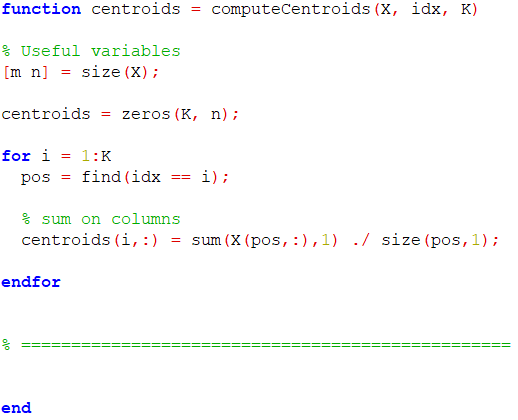
1. Реализуйте функцию случайной инициализации K центров кластеров.



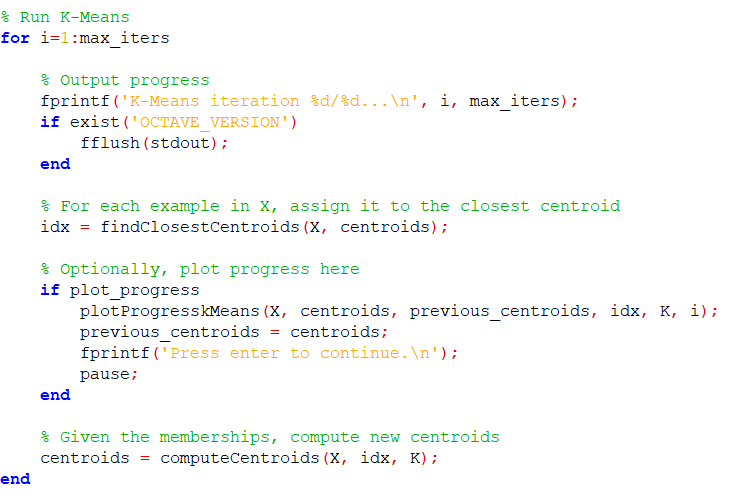
1. Реализуйте функцию определения принадлежности к кластерам.



1. Реализуйте функцию пересчета центров кластеров.



1. Реализуйте алгоритм K-средних.



1. Постройте график, на котором данные разделены на K=3 кластеров (при помощи различных маркеров или цветов), а также траекторию движения центров кластеров в процессе работы алгоритма.

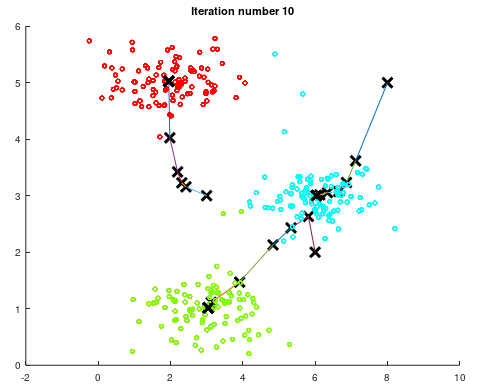


Рисунок 1 – График разделения данных на 3 кластера

1. Загрузите данные **bird\_small.mat** из файла.
2. С помощью алгоритма K-средних используйте 16 цветов для кодирования пикселей.

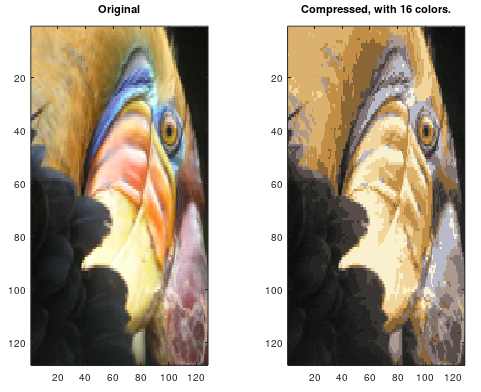


Рисунок 2 – Сравнение исходного изображения и сжатого (16 цветов)

# ЗАКЛЮЧЕНИЕ

В данной лабораторной работе были освоены навыки использования алгоритма K-средних для решения задачи классификации, а также для сжатия изображения, используя для кодировки его пикселей только *n*-ное количество цветов.