

# **КЛАСТЕРИЗАЦИЯ ЗНАЧЕНИЙ ПИКСЕЛЕЙ ИЗОБРАЖЕНИЙ С УЧЕТОМ МЕТРИКИ ЦВЕТОВОГО РАЗЛИЧИЯ CIEDE2000**

**В.В. Плетнев, студент группы 529-СТ  
Н.Н. Пономаренко, д.т.н., профессор кафедры 504**

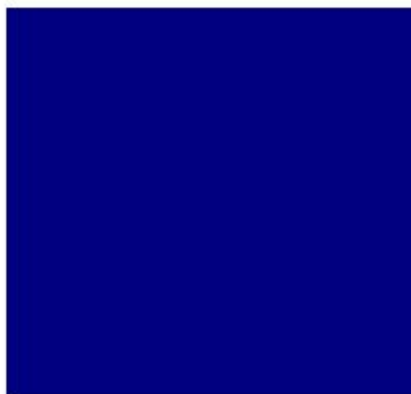
**Кафедра «Приема, передачи и обработки сигналов»  
Национальный аэрокосмический университет  
им. Н.Е. Жуковского «ХАИ»  
Украина, Харьков**

# Цель и актуальность исследования

Где контрастность выше?



(0, 128, 0)  
(0, 255, 0)



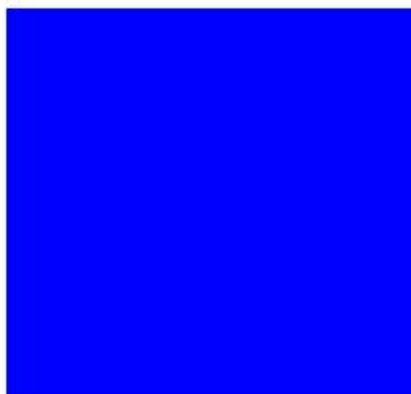
(0, 0, 128)  
(0, 0, 255)



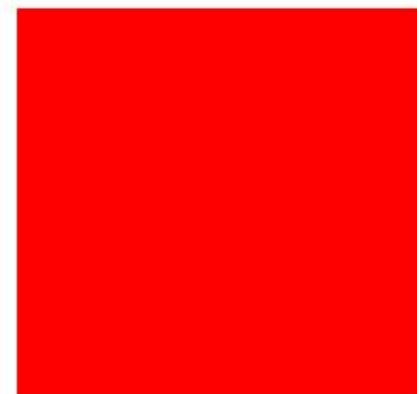
(128, 0, 0)  
(255, 0, 0)



34, 5



16,5



25,8

Разработана международным комитетом CIE  
(фр. Commission Internationale de l'Eclairage) в 2000 году

$$\Delta L' = L_2^* - L_1^*$$

$$\bar{L} = \frac{L_1^* + L_2^*}{2} \quad \bar{C} = \frac{C_1^* + C_2^*}{2}$$

$$a_1' = a_1^* + \frac{a_1^*}{2} \left( 1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}} \right) \quad a_2' = a_2^* + \frac{a_2^*}{2} \left( 1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}} \right)$$

$$\Delta h' = \begin{cases} h_2' - h_1' & |h_1' - h_2'| \leq 180^\circ \\ h_2' - h_1' + 360^\circ & |h_1' - h_2'| > 180^\circ, h_2' \leq h_1' \\ h_2' - h_1' - 360^\circ & |h_1' - h_2'| > 180^\circ, h_2' > h_1' \end{cases}$$

$$\bar{C}' = \frac{C_1' + C_2'}{2} \text{ and } \Delta C' = C_2' - C_1' \quad \text{where } C_1' = \sqrt{a_1'^2 + b_1'^2} \quad C_2' = \sqrt{a_2'^2 + b_2'^2}$$

$$h_1' = \text{atan2}(b_1', a_1') \text{ mod } 360^\circ, \quad h_2' = \text{atan2}(b_2', a_2') \text{ mod } 360^\circ$$

$$\Delta H' = 2\sqrt{C_1' C_2'} \sin(\Delta h'/2), \quad \bar{H}' = \begin{cases} (h_1' + h_2' + 360^\circ)/2 & |h_1' - h_2'| > 180^\circ \\ (h_1' + h_2')/2 & |h_1' - h_2'| \leq 180^\circ \end{cases}$$

$$T = 1 - 0.17 \cos(\bar{H}' - 30^\circ) + 0.24 \cos(2\bar{H}') + 0.32 \cos(3\bar{H}' + 6^\circ) - 0.20 \cos(4\bar{H}' - 63^\circ)$$

$$S_L = 1 + \frac{0.015 (\bar{L} - 50)^2}{\sqrt{20 + (\bar{L} - 50)^2}} \quad S_C = 1 + 0.045 \bar{C}' \quad S_H = 1 + 0.015 \bar{C}' T$$

$$R_T = -2 \sqrt{\frac{\bar{C}'^7}{\bar{C}'^7 + 25^7}} \sin \left[ 60^\circ \cdot \exp \left( - \left[ \frac{\bar{H}' - 275^\circ}{25^\circ} \right]^2 \right) \right]$$

## Задача работы

Относительно недавно был разработан метод кластеризации по граничному элементу множества:

*Пономаренко Н.Н. "Быстрая кластеризация в многомерном пространстве для задач поиска подобия", Системы обработки информации, Вып. 2 (76), 2009, С. 79-82.*

Основное достоинство этого метода: способность при кластеризации учитывать любую функцию расстояния, заданную в виде «черного ящика», например, CIEDE2000.

Задача исследования: оценить, насколько это позволяет уменьшить погрешность кластеризации по сравнению с стандартной кластеризацией Ллойда?



# Кластеризация Ллойда

Дано:  $n$  точек  $x_i$  в многомерном пространстве, которые нужно разбить на  $k$  кластеров.

1. Выбираем  $k$  случайных (их координаты заданы случайно) центров кластеров  $c_j, j=1..k$ .
2. Для каждой точки  $x_i$  находим наиболее близкий к ней центр кластера  $c_m$ . Считаем, что точка  $x_i$  попадает в кластер  $m$ .
3. Для каждого кластера  $j$  подсчитываем, сколько точек попало в него. Вычисляем новое значение  $c_j$  усреднением координат всех  $x_i$ , которые попали в этот кластер.

Шаги 2-3 повторяются до тех пор, пока точки не перестанут переходить от одного кластера к другому (координаты центров кластеров перестанут изменяться).

Если какой-либо из кластеров оказывается пустым, то для него выбирается новый случайный центр, либо самый большой кластер разбивается на две половины.

## Кластеризация по расстоянию до граничного элемента $O(n)$

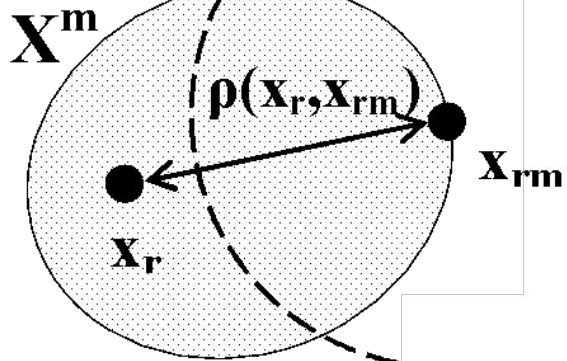
Дано:  $n$  точек  $x_i$  в  $M$ -мерном пространстве, которые нужно разбить на  $k$  кластеров.

1. Все точки относим к одному кластеру.
2. Просматриваем все кластеры и определяем кластер  $j$  с наибольшим средним расстоянием его пикселей до его центра
3. Выбираем случайную точку  $x_r$  в кластере  $j$ .
4. В кластере  $j$  находим точку  $x_{rm}$  наиболее удаленную от  $x_r$ . Она будет граничной (лежать на границе кластера).
5. Упорядочиваем все точки кластера  $j$  по расстоянию до точки  $x_{rm}$ . Находим медиану этих расстояний  $med$ .
6. Разбиваем кластер  $j$  на два кластера так, что в один кластер идут точки с расстоянием до  $x_{rm}$  меньше или равно  $med$ , а в другой – остальные.

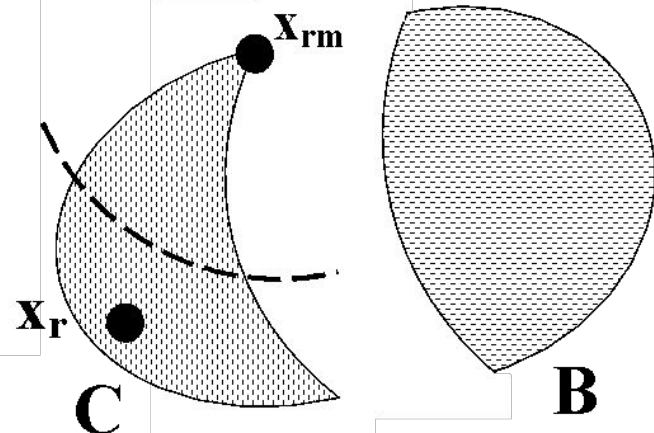
Шаги 2-6 повторяются, пока число кластеров меньше  $k$ .

# Кластеризация по граничному элементу множества

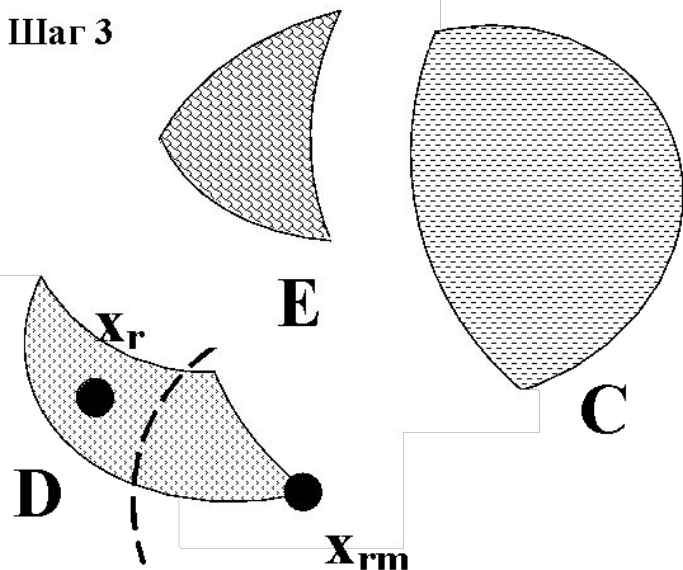
Шаг 1



Шаг 2

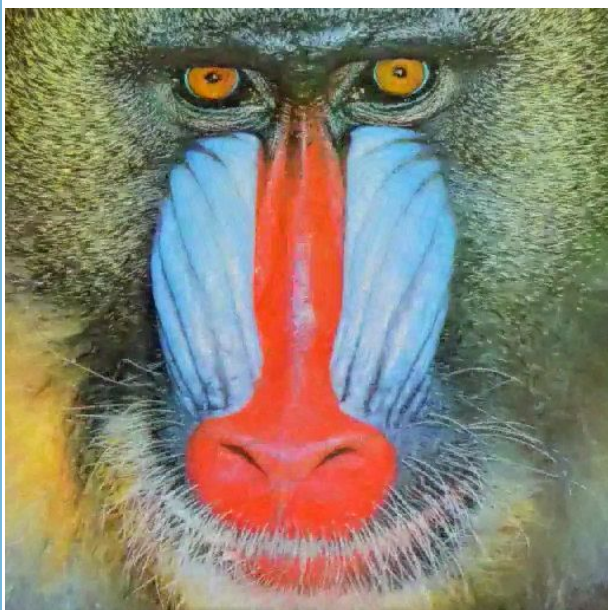


Шаг 3



Пояснение к  
 кластеризации по  
 расстоянию до  
 граничного элемента

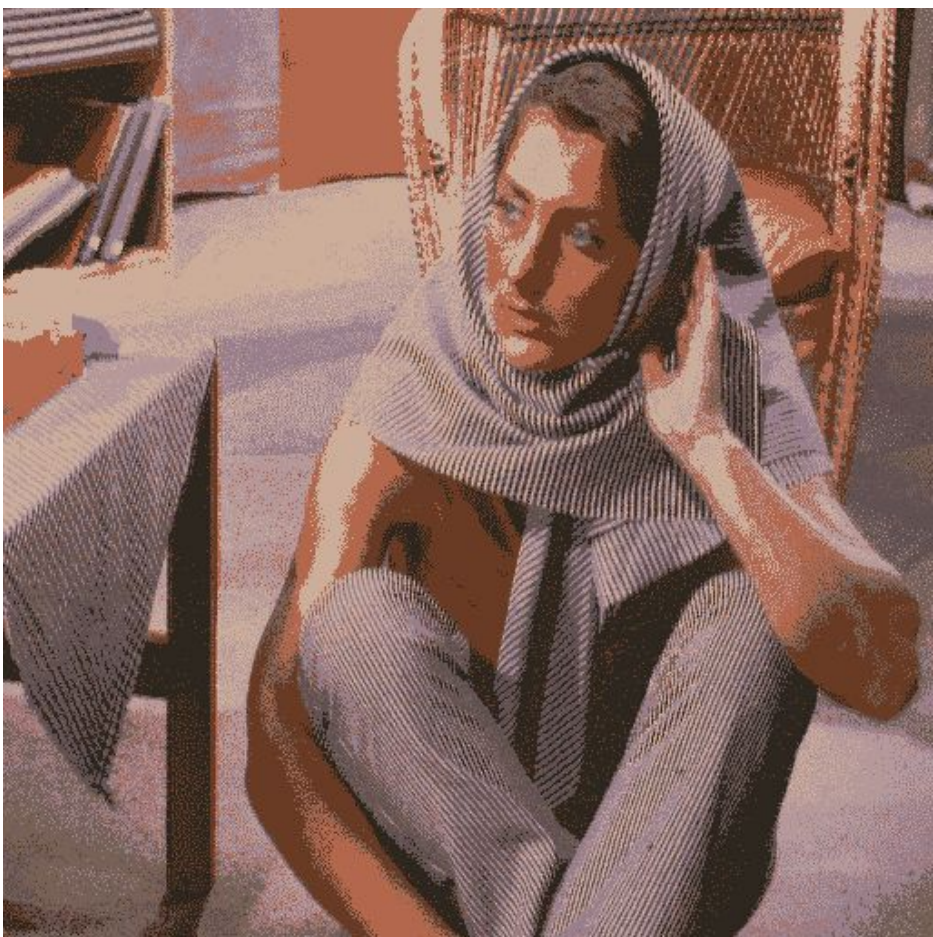
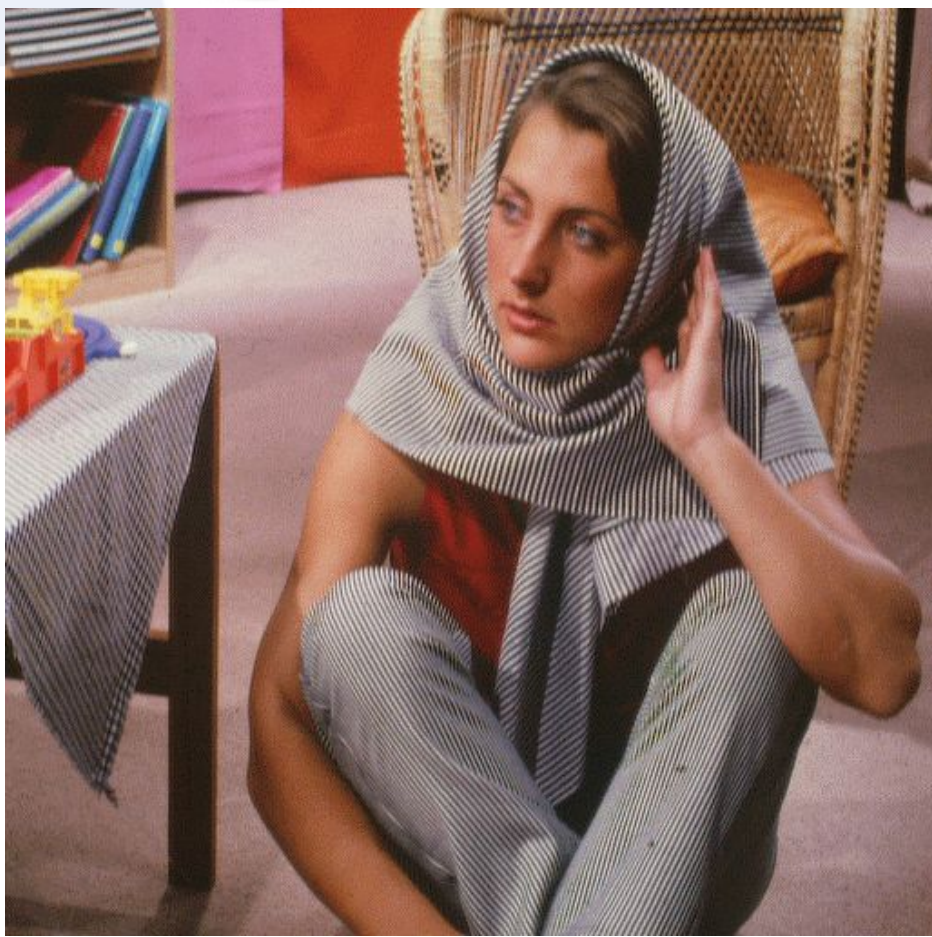
# Тестовые изображения



Стандартные цветные тестовые изображения Baboon, Barbara и Goldhill  
(512x512 пикселей)



# Пример кластеризации



Граничная кластеризация с метрикой CIEDE2000 на 8 кластеров

# Результаты

Стандартные цветные тестовые изображения Baboon, Barbara и Goldhill  
(512x512 пикселей)

Метод	Baboon		Barbara		Goldhill	
	256 класт.	1024 класт.	256 класт.	1024 класт.	256 класт.	1024 класт.
Класт. Ллойда	18,1	8,0	4,1	12,1	11,8	5,5
Класт. по гран. эл. множества	13,6	5,7	3,6	9,5	8,9	4,0
% выигрыша	25%	29%	12%	21%	25%	27%

## Выводы

- В данном исследовании был протестирован метод оценки разности цветов CIEDE2000 на двух методах кластеризации Ллойда и метод кластеризации по граничному элементу множества
- Показано преимущество метода по граничному множеству в точности нахождения оптимальных расстояний с сохранением меньшей разности между кластерезованным и исходным изображением.
- К тому же метод пограничного множества имеет преимущество в скорости обработки, так как он не требует большого количества пересчётов для оптимальной кластеризации.