

## **МОДЕЛИРОВАНИЕ МЕТОДОВ КЛАСТЕРИЗАЦИИ ДЛЯ МНОЖЕСТВ ХАРАКТЕРНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ**

Дунаевская М. Д., Струненко В. А.

Научный руководитель – д.т.н., проф. Гороховатский В.А.  
Харьковский национальный университет радиоэлектроники  
E-mail: mary.dunaevskay@gmail.com

Современные робототехнические системы с использованием компьютерного зрения решают задачи распознавания визуальных объектов и применяются в машиностроении, медицине, автоматизации документооборота, и других. Для ускорения процессов распознавания образов и более результативного решения прикладных задач используют методы кластеризации [1,3].

Цель данной работы - сравнительный анализ эффективности применения различных методов кластерного анализа – к-средних и разностного группирования в задачах структурного распознавания изображений по множеству характерных признаков (ХП).

В качестве ХП использованы дескрипторы (рис. 1), полученные с помощью программного моделирования метода SURF [2], которые представляют собой числовые векторы, содержащие 64 или 128 компонентов, и отображают изменения градиента яркости в точке координат ХП.

Метод К-средних (K-means) разбивает множество объектов на заданное число кластеров. Метод разностного группирования (РГ) реализует процесс самоорганизации множества векторов. Основное различие методов заключается в способе задания входных данных: количество кластеров в методе К-средних задается в виде параметра, а в методе РГ определяется автоматически в ходе кластеризации. Погрешность построения кластерного описания (векторного квантования) вычислена в виде

$$E = \frac{1}{s} \sum_{i=1}^k \sum_{z \in C_i} \rho^2(z, c_i), \quad (1)$$

где  $z \in C_i$  – точка данных кластера  $C_i$  с центром  $c_i$ ,  $\rho(z, c_i)$  – расстояние между центром и точкой (как правило – евклидово);  $k$  – число кластеров;  $s$  – мощность обучающего множества  $Z$ .



Рис. 1. Изображение и множество ХП

В табл. 1 представлены результаты вычисления погрешности квантования для обоих методов при разном количестве кластеров. В методе РГ число кластеров зависит от величины априорно заданного порога эквивалентности дескрипторов в евклидовой метрике.

Таблица 1 - Значения погрешности квантования

Число кластеров	Метод РГ	Метод К-средних
822	0.078	0.012
225	0.346	0.062
50	0.388	0.108
14	0.382	0.154
4	0.356	0.199

Проанализировав результаты проведенного моделирования, делаем вывод, что при использовании метода К-средних погрешность квантования меньше, чем при применении метода РГ. Например, при числе кластеров, равном 14, погрешность для метода К-средних меньше в 2,5 раза. В то же время метод РГ автоматически путем самоорганизации определяет число кластеров в зависимости от параметра эквивалентности дескрипторов.

#### Литература:

1. Шапиро Л. Компьютерное зрение/ Л. Шапиро, Дж. Стокман; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
2. Bay H. Surf: Speeded up robust features/ H. Bay, T. Tuytelaars, L. Van Gool // European Conference on Computer Vision. – 2006. – P.404-417.
3. Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении: монография / В.А. Гороховатский. – Х.: Компания СМІТ, 2014. – 316с.