

# РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ С ПРИМЕНЕНИЕМ МОДИФИЦИРОВАННОГО АЛГОРИТМА, ОСНОВАННОГО НА ДЕСКРИПТОРЕ КОНТЕКСТА ФОРМЫ

*Р.К. Захаров*

*Самарский государственный аэрокосмический университет  
им. С.П. Королёва*

Предлагается метод распознавания изображений, основанный на дескрипторе контекста формы. Предложенный алгоритм является улучшением классического метода, основанного на применении дескриптора контекста формы. Алгоритм находит на изображении объекты заданного подкласса (к примеру, колёса, двери, крышу для класса автомобилей) и принимает решение, являются ли эти объекты объектом класса автомобиль.

## **Введение**

В данной работе стоит задача продемонстрировать алгоритм распознавания изображений, в качестве базы данных изображений была взята база данных The PASCAL Object Recognition Database Collection. На рис. 1 представлен пример тестового изображения.



Рис. 1. Пример тестового изображения из базы данных PASCAL

В работе представлен подход к поиску визуальных слов (BoW Bag of Words) на изображении, их распознаванию и принятию решения, является ли данный объект автомобилем.

## **Постановка задачи**

Общая постановка задачи распознавания образов следующая. Предполагается, что имеется  $M$  изображений каждого из  $K$  объектов. Каждое изображение представляется вектором  $x = [x_1, x_2, \dots, x_N]^T$  размерности  $N$ , где  $x_1, x_2, \dots, x_N$  – признаки. Векторы, соответствующие изображениям одного объекта, составляют класс. Совокупность векторов признаков всех классов образует обучающую выборку. Решение задачи распознавания состоит в конструировании решающей функции  $f: R^N \rightarrow \{0, 1, \dots, K\}$ , которая каждому вектору  $x$  ставит в соответствие некоторый класс. Для уменьшения числа неправильных классификаций вводится также класс с номером 0, соответствующий отказу в распознавании.

Качество распознавания зависит от выбора системы признаков. Наряду с выбором системы признаков большую роль играют также используемая при распознавании мера близости и построенное на ее основе решающее правило.

Для решения задачи распознавания и поиска объектов на изображении широко используется машина опорных векторов [4], а в качестве признаков наиболее популярными являются гистограммы ориентированных градиентов (HOG) [6], также часто используют метод Viola-Jones [3], который показывает хорошие результаты для отслеживания человеческих лиц в кадре, и различные методы, основанные на комбинировании алгоритмов.

В настоящей работе ставится задача провести исследования модифицированного алгоритма, основанного на дескрипторе контекста формы, для базы данных автомобилей. Метод, основанный на дескрипторе контекста формы, впервые ввели Serge Belongie и Jitendra Malik в работе [2] в 2000 году. Данный дескриптор широко используется в задачах распознавания цифр, букв, он показывал высокую надёжность распознавания на базе данных MNIST.

Новизна работы заключается в модификации классического алгоритма, основанного на дескрипторе контекста формы, к задаче распознавания изображений. Осуществляется поиск и распознавание не самого объекта, а визуальных слов, описывающих данный объект. Далее принимается решение, является ли эта совокупность визуальных слов распознаваемым объектом.

### **Описание модифицированного алгоритма, основанного на дескрипторе контекста формы изображения**

Приведём описание алгоритма распознавания изображений, предложенного в работе, на рис. 2.



Рис. 2. Общая схема модифицированного алгоритма распознавания изображений

На рис. 3 представлена более подробная схема первого и второго этапов алгоритма, анализа найденных визуальных слов.

Рассмотрим подробнее каждый из этапов, изображённых на схеме.

1. Поиск визуальных слов на изображении. Реализуется через метод скользящего окна по изображению.

2. Анализ найденных визуальных слов. Осуществляется по схеме на рис. 3.
  - а. Для каждого окна находятся ключевые точки на изображении. Алгоритм поиска ключевых точек произвольный: SIFT [5,7], SURF [8], FAST [9].
  - б. Вычисляем дескриптор формы [1, 2] для каждой ключевой точки на изображении.
  - с. Распознаём, является ли текущее окно визуальным словом. Для распознавания в работе используется машина опорных векторов SVM [4] для увеличения скорости принятия решения данный шаг алгоритма был распараллелен. Одно окно проверялось одновременно на сопоставление всем визуальным словам.
3. Процедура принятия решения, принадлежит ли объект распознаваемому классу, строится на основе количества найденных визуальных слов, их взаимного расположения и типов найденных визуальных слов.

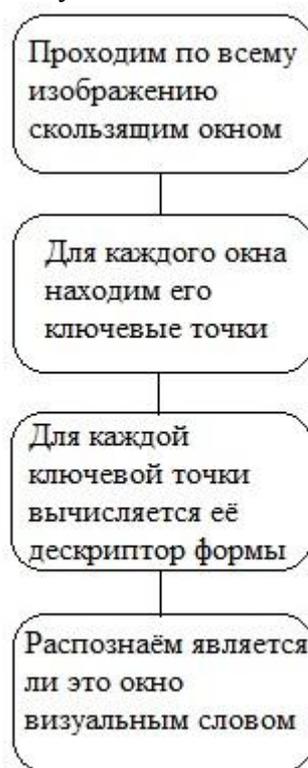


Рис. 3. Схема поиска и распознавания визуальных слов на изображении

### Заключение

В работе представлен подход к распознаванию изображений с использованием дескриптора контекста формы. Предложенный алгоритм является модификацией классического метода, основанного на дескрипторе контекста формы для изображения. Алгоритм является достаточно эффективным и надёжным методом для распознавания объектов на изображениях.

### Литература

1. Belongie S., Malik J. Matching with Shape Contexts // IEEE Workshop on Contentbased Access of Image and Video Libraries (CBAIVL-2000). 2000.
2. Belongie S., Malik J., Puzicha J. Shape Matching and Object Recognition Using Shape Contexts // IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (24):509-521. DOI : 10.1109 / 34.993558. April 2002.

3. Viola, Jones. Rapid object detection using a boosted cascade of simple features // Computer Vision and Pattern Recognition. 2001.
4. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines John C. Platt Microsoft Research jplatt@microsoft.com Technical Report MSR-TR-98-14 April 21. 1998 © 1998 John Platt.
5. Lowe D.G. Object recognition from local scale-invariant features // Proceedings of the International Conference on Computer Vision. 1999. 2. P. 1150–1157. DOI:10.1109 / ICCV.1999.790410.
6. Dalal N., Triggs W. Histograms of Oriented Gradients for Human Detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05. 2005. V. 1(3). P. 886-893.
7. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. 2004. 60, 2. P. 91-110.
8. Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features // Computer Vision and Image Understanding (CVIU). 2008. Vol. 110. No. 3. P. 346–359.
9. Trajkovic M., Hedley M. Fast corner detection // Image and Vision Computing. 1998. 16 (2). P. 75–87.