



Р.К. Захаров

МЕТОД КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ ОСНОВАННЫЙ НА МОДИФИЦИРОВАННОМ ДЕСКРИПТОРЕ КОНТЕКСТА ФОРМЫ

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (национальный исследовательский университет)»)

Введение

В данной работе стоит задача продемонстрировать алгоритм распознавания изображений, в качестве баз данных изображений были взяты данные с международного конкурса по распознаванию изображений The PASCAL Object Recognition Database Collection Visual Object Classes – 2007 и CalTech dataset. Ниже на рис. 1 представлен пример тестового изображения.



Рис. 1. Пример тестового изображения из базы данных PASCAL VOC - 2007

В работе представлен подход для распознавания изображений. Каждый объект характеризуется набором дескрипторов, которые описывают всё изображение.

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(НОГ)[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. 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 изображения. Распознавание происходит не самого объекта, а набора дескрипторов для изображения. Для построения дескрипторов используется не контуры, а особенности изображения то есть выделяются ключевые точки и по ним строится дескриптор формы. После построения дескрипторов происходит классификация, а затем на основе данной классификации принимается решение, какому классу распознаваемых объектов принадлежит изображение.

Описание алгоритма основанного на модифицированном дескрипторе контекста формы

Предлагаемый метод основан на дескрипторе контекста формы. На рис. 2. изображена общая схема классификации для предлагаемого метода

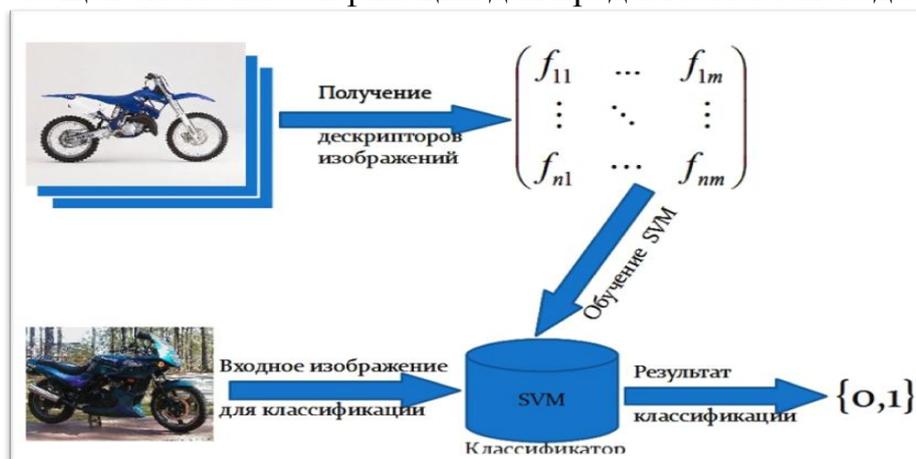


Рис. 2. Общая схема классификации на основе SVM для распознавания изображений

Далее опишем метод формирования матрицы дескрипторов. На рис. 3. схематично представлен сам метод формирования дескриптора

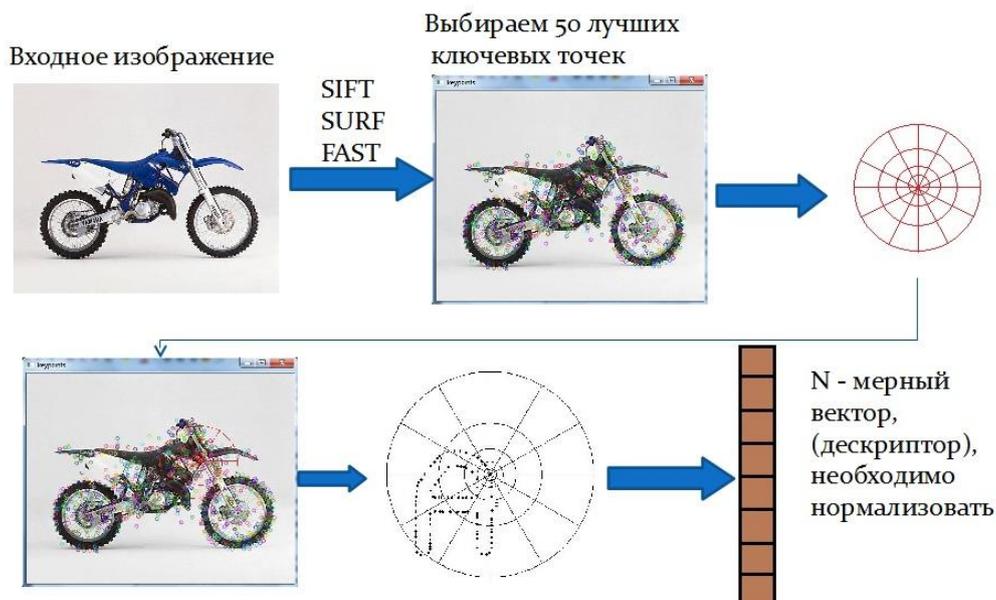


Рис. 3. формирование дескриптора для описания изображения
Разберём каждый из этапов чуть более подробно

Процедура формирования дескриптора:

1. На вход поступает изображение.
2. Строим пирамиду изображений, в данном методе было построено из одного входного изображения ещё два дополнительных. Одно имело размер в два раза меньший исходного другое же в два раза больший.
3. Выделяем ключевые особенности на изображениях (для выделения можно использовать любой алгоритм) в данной работе был использован метод SURF(Speeded Up Robust Features)
4. Выбираем 50 лучших ключевых точек. Из трёх изображений, было сформировано 3-и облака точек, сравнивая их между собой, возьмём самые лучшие точки, которые наиболее хорошо описывают объект
5. Дескриптор формы разделённый на 48 бинов вычисляется для каждой ключевой точки.
6. Объединяем все дескрипторы, полученные для ключевых точек, и получаем один выходной дескриптор, который описывает изображение.

После того как дескриптор был сформирован подаём его на вход классификатора и он его классифицирует как распознаваемый объект или объект не принадлежащий ни одному распознаваемому классу объектов.

Сравнение метода Shape Context с предлагаемым методом

Ниже представлена таблица 1. с основными отличиями в подходах в Shape Context и исследуемом методе.

Выделение ключевых точек более точно характеризует сам объект, а не всё изображение с фоном. На этапе классификации SVM работает быстрее чем Hungarian method за счёт того что скалярно перемножается вектор с весовыми коэффициентами на дескриптор, а Hungarian method производит перебор всех пар точек.



Таблица 1. Сравнение подходов в формировании дескриптора
и методов их распознавания

Shape Context	Исследуемый метод
Выделяет контуры объекта	Выделяет ключевые точки и выбирает лучшие
Распознавание происходит с помощью Hungarian method	Для распознавания обучается классификатор SVM

Результаты экспериментов

Для тестирования была выбрана база данных с международного конкурса по анализу изображений PASCAL VOC - 2007 (Pattern Analysis Statistical Modeling and Computational Learning Visual Object Classes - 2007)

Ниже приведём таблицу с лучшими результатами конкурса и сравним с исследуемым методом

Таблица 2. Сравнение методов распознавания изображений на 2-х классах

Методы / классы изображений	Автомобиль	Мотоцикл
Исследуемый метод	0.80	0.83
INRIA_Genetic	0.780	0.640
XRCE	0.754	0.585

Заключение

В работе представлен подход для распознавания изображений с использованием дескриптора контекста формы. Исследуемый метод является модификацией классического метода основанного на дескрипторе контекста формы для изображения. Метод является достаточно эффективным и надёжным 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., and 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, 2002
3. Viola and Jones Rapid object detection using a boosted cascade of simple features // Computer Vision and Pattern Recognition, 2001
4. John C. Platt Sequential minimal optimization: a fast algorithm for training support vector machines // Microsoft Research jplatt@microsoft.com Technical Report MSR-TR-98-14 April 21, 1998
5. Lowe, David G. Object recognition from local scale-invariant features // Proceedings of the International Conference on Computer Vision. 2. pp. 1150–1157. DOI:10.1109/ICCV.1999.790410, 1999
6. Dalal N., Triggs W. Histograms of oriented gradients for human detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05, 2005. – Vol. 1(3). – pp. 886 – 893