

# НОG+НОF+ССS ДЕСКРИПТОР В ЗАДАЧЕ ОБНАРУЖЕНИЯ ПЕШЕХОДОВ

Малыгин Евгений Сергеевич, бакалавр, [malygin.work@yandex.ru](mailto:malygin.work@yandex.ru)  
Научный руководитель: Вахитов Александр Тимурович, к. ф.-м. н.

## Аннотация

СПИСОК (Системное Программирование, Информационные Системы, Обеспечение Качества) — периодическая научная конференция по проблемам информатики.

В этой статье рассказано о практической реализации современного дескриптора для детектирования пешеходов и о полученных результатах.

## Введение

В связи с доступностью камер и вычислительных мощностей и с разработкой беспилотных транспортных средств задача детектирования пешеходов актуальна.

Учёные предлагают признаки, которые можно получить из изображений видеоряда, с помощью которых можно успешно находить на изображениях пешеходов. В бегущем по изображению окне формируется вектор таких признаков, который затем подаётся на вход классификатору, который в простой модели относит окно к одному из двух классов: пешеход или не-пешеход.

Успешность метода детектирования характеризуется несколькими параметрами: точностью, долей найденных истинных пешеходов, долей ложных срабатываний, скоростью работы и т.д.

Возникла идея реализовать и изучить какой-либо перспективный современный метод распознавания пешеходов. В качестве такого метода был выбран метод, предложенный в статье [2] и основанный на комбинации НОG, НОF и ССС дескрипторов. Дополнительной целью работы была проверка точности этого составного дескриптора по сравнению с общеизвестным НОG-дескриптором.

## Комбинация дескрипторов

- НОG – гистограммы ориентированных градиентов. Это известный уже реализованный в OpenCV дескриптор, повторная реализация которого не несёт интереса. Гистограммы ориентированных градиентов как дескриптор извлекают из одиночного изображения информацию о цветовых контурах, вычисляя локальные связи между цветами на изображении. НОG дескриптор показывает

хорошую производительность в задаче обнаружения пешеходов, поэтому обычно с ним сравнивают новые детекторы, например, в статьях [1], [2].

- HOF – гистограммы ориентированного оптического потока. Этот дескриптор, предложенный в статье [3], извлекает из 2-х последовательных изображений информацию об относительном движении близких малых областей картинки относительно друг друга. Сначала по двум картинкам вычисляется оптический поток по какому-то уже существующему алгоритму. В данной работе для нахождения оптического потока был использован метод Farnerback из OpenCV. После нахождения оптического потока в окна поиска производится суммирование и нормализация смещений по осям OX и OY.
- CSS – самоподобие цветов, дескриптор, предложенный авторами статьи [2]. Этот дескриптор характеризует глобальные связи между цветами на изображении. Сначала исходное изображение в цветовом пространстве RGB разбивается на блоки размером 8x8. В каждом из этих блоков определяется средний цвет по каждой цветовой компоненте, в результате чего получается изображение пониженного разрешения. Затем изображение пониженного разрешения переводится в цветовое пространство HSV, которое показало наилучшие результаты в работе CSS-дескриптора. Затем между каждой парой блоков в фиксированном порядке считается нормированное евклидово расстояние в цветовом пространстве, и эти значения заносятся в вектор признаков.

## **Классификатор**

Было решено использовать метод опорных векторов (SVM) для классификации. Было опробовано 3 реализации SVM: SVMLight, OpenCV SVM и LIBLINEAR. В связи с проблемами с масштабируемостью данных для обучения SVMLight и OpenCV SVM были откинута, и в конечном итоге была взята реализация SVM LIBLINEAR.

Хорошо документированная [4] библиотека LIBLINEAR была интегрирована в код дескриптора и успешно использована.

## **Наборы данных**

Для обучения и тестирования классификатора требовалось использовать наборы данных, относящиеся к предметной области. К наборам данных были следующие требования:

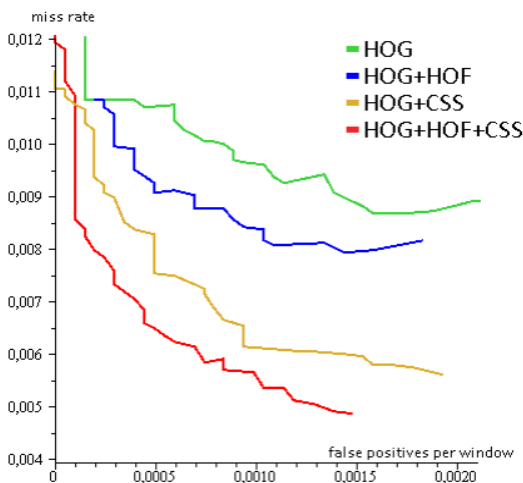
- Изображения должны были быть цветными (чтобы CSS-дескриптор не вырождался в одномерном цветовом пространстве).
- Изображения должны идти совместно, как в видеоряде.
- Набор данных должен быть размечен, то есть должна присутствовать информация о рамках с действительными пешеходами.
- Изображения должны быть представлены в удобном для работы виде.
- Число кадров и уникальных пешеходов должно быть достаточно велико.

Этим условиям удовлетворяло 2 набора данных: TUD-Brussels [5] и ETN [6].

## Обучение и тестирование

Из наборов данных были нарезаны и масштабированы к размеру окна поиска рамки с пешеходами и рамки без пешеходов для обучения SVM в составе дескриптора.

Обучение и тестирование по наборам данным производилось раздельно, и был получен следующий график для разных комбинаций использованных дескрипторов:



Этот график отражает изменения в работе детектора при смене bias у SVM. По графику видно, что miss rate (1.0 - accuracy) и доля ложноположительных срабатываний детектора у комбинаций детекторов ниже.

Также было произведено исследование скорости работы отдельных частей дескриптора.

|             | 1 окно, мс | окон в секунду | секунд на изображение |
|-------------|------------|----------------|-----------------------|
| CSS         | 0.375      | 2667           | 2.8                   |
| HOG         | 0.730      | 1370           | 5.5                   |
| HOF         | 4          | 250            | 30                    |
| HOG+HOF+CSS | 5.1        | 196            | 38                    |

По таблице видно, что самой медленной частью оказался подсчёт оптического потока.

## Заключение

Был реализован комбинированный дескриптор, в который была интегрирована библиотека для SVM. Дескриптор был обучен и протестирован, и была доказана польза комбинирования предлагаемых признаков в задаче детектирования пешеходов.

## Литература

1. Dollar P. et al. Pedestrian detection: An evaluation of the state of the art //Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2012. – Т. 34. – №. 4. – С. 743-761
2. S. Walk, N. Majer, K. Schindler, and B. Schiele, “New features and insights for pedestrian detection,” in IEEE Conf. Computer Vision and Pattern Recognition, 2010. 8, 11, 12
3. Navneet Dalal, Bill Triggs, Cordelia Schmid: Human Detection Using Oriented Histograms of Flow and Appearance. ECCV (2) 2006: 428-441
4. <http://www.csie.ntu.edu.tw/~cjlin/liblinear/>
5. <http://www.d2.mpi-inf.mpg.de/tud-brussels/>
6. <http://www.vision.ee.ethz.ch/~aess/dataset/>