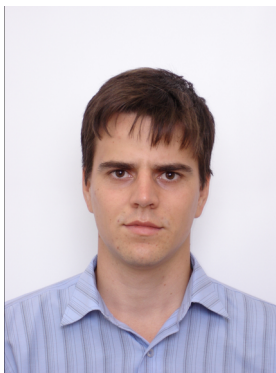


МЕТОД РАСПОЗНАВАНИЯ ОБЪЕКТОВ В ЦИФРОВЫХ СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ

Нужный С. П.

Ставропольский государственный университет
nuzhny@mail.ru



Нужный Сергей Петрович аспирант кафедры алгебры физико-математического факультета Ставропольского государственного университета. К основным научным интересам относится машинное зрение, классификация, обработка изображений.

В связи со стремительным распространением цифровых систем видеонаблюдения повышаются требования к качеству обработки и анализа видеоизображений. Сегодня недостаточно констатации движения в кадре и устойчивости детектора к дождю и снегу. Необходимо также выделять и распознавать движущиеся объекты [4, 5], отслеживать их перемещение, проводить анализ поведения.

Задачи распознавания подразумевают чаще всего поиск человеческих лиц, пешеходов и автотранспорта. Реже требуется распознавание лиц и автомобильных номеров.

В работе [3] классификация объектов (пешеходов и автотранспорта) происходит на основании заранее сформированной *библиотеке шаблонов*. Эта библиотека (codebook) содержит 400 образцов, собранных вручную в течение длительного наблюдения на разных сценах. При анализе очередного объекта используется векторное квантование и автоматическая иерархическая классификация на основании имеющихся шаблонов.

В работе [2] подробно рассмотрены *классификаторы Байеса* на основе различных распределений. Рассмотрено построение классификаторов по гистограммам классов.

Часто для распознавания объектов используются *нейронные сети*. Для этого метода характерна длительная начальная инициализация и обучение.

Для случаев построения классификатора по образцам можно использовать *машину опорных векторов* (SVM). Скорость работы данного метода может в несколько раз превосходит скорость работы нейронной сети.

Разреженная просеивающая сеть (SNoW) является особым видом нейронной сети. Данный подход к распознаванию объектов на видеоизображениях считается весьма перспективным благодаря высокой скорости работы и маленького процента ошибок на контрольной выборке.

Одни из самых лучших результатов в плане точности распознавания и быстродействия показывает *метод усиления слабых классификаторов* (classifier boosting). Он основывается на комбинировании более слабых классификаторов в один более сильный комитет. Данный метод пытается спроецировать исходные данные в пространство линейно разделимых классов.

В данной работе рассматривается метод распознавания людей и автотранспорта, основанный на библиотеке шаблонов. Каждый шаблон в библиотеке содержит идентификатор класса (человек или автомобиль), к которому он относится.

В качестве исходных данных, используемых при формировании библиотеки шаблонов, используются результаты работы алгоритма вычитания фона [1] и последующей сегментации (рис. 1, 2).



Рис. 1 Кадр в формате RGB24 уличной сцены



Рис. 2 Результат действия алгоритма вычитания фона и сегментации

Шаблон представляет собой матрицу T размера $n_t \times m_t$, элементы которой формируются по следующему алгоритму:

1. Анализируется движущийся объект, заданный прямоугольной областью $n_o \times m_o$. Каждый пиксель объекта классифицирован как передне- или заднеплановый.
2. Горизонтальная векторизация: просматривается первая строка объекта, счётчик одинаковых пикселей:
 - а) если текущий пиксель принадлежит переднему плану, то переходим к следующему и увеличиваем счётчик l на единицу, иначе переходим к пункту 2.в;
 - б) просматриваются пиксели до тех пор, пока не будет найден пиксель заднего плана;
 - в) текущий пиксель не принадлежит переднему плану, то

очередной элемент шаблона T будет равен $\frac{l}{n_o}$. Счётчик l обнуляется. Для последующих пикселей заднего плана выполняются операции, начиная с пункта 2.а, аналогичные переднеплановым;

г) в случае достижения конца строки объекта, оставшиеся элементы матрицы шаблона T заполняются нулями. Далее просматриваются все строки объекта.

3. Вертикальная векторизация: просматривается первый столбец матрицы шаблона T :

а) столбец разбивается на участки размером $\frac{m_o}{m_t}$;

б) для каждого участка вычисляется среднее арифметическое составляющих его величин;

в) полученные значения записываются в столбцы матрицы шаблона T .

Библиотека шаблонов формируется по описанному алгоритму на основе видеозаписей, полученных с разных камер при разных условиях освещённости. Размер библиотеки зависит от размерности шаблона. Например, при размере шаблона 6×12 в библиотеке достаточно 400-500 шаблонов. От количества и размеров шаблонов также зависит скорость поиска в библиотеке при штатной работе алгоритма.

Данный алгоритм является инвариантным по отношению к операциям масштабирования. Однако первоначальное формирование библиотеки является достаточно трудоёмким процессом.

Литература

1. Нужный С. П., Червяков Н. И. Метод вычитания фона в видеопотоке с возможностью детектирования оставленных предметов // Физико-математические науки на современном этапе развития СГУ: Материалы 52-й научно-методической конференции преподавателей и студентов СГУ "Университетская наука региону" - Ставрополь: Издательство СГУ, 2007.
2. Д. Форсайт, Ж. Понс Компьютерное зрение. Современный подход. : Пер. с англ. – М.: Издательский дом «Вильямс», 2004.
3. Chris Stauffer, W. Eric L. Grimson Learning Patterns of Activity Using Real-Time Tracking, 2000.
4. A. Lipton, H. Fujiyoshi, and R.S. Patil, Moving Target Classification and Tracking from Real-Time Video, Proc. IEEE Workshop Applications of Computer Vision (WACV), pp. 8-14, Oct. 1998.
5. C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, Pfunder: Real-Time Tracking of the Human Body, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 780-785, July 1997.