

УДК 004.932.2

И. А. Горошков, С. А. Кулаков

Детектор распознавания присутствия людей в кадре

Приводится алгоритм распознавания присутствия человека в кадре. Обсуждаются преимущества выбранного алгоритма определения людей. Отмечается возможность применения данного алгоритма для обеспечения сохранности ТМЦ (товарно-материальные ценности) и безопасности на предприятиях любых сфер и направлений деятельности.

Алгоритм распознавания людей, HOG, машинное обучение, трекинг

Задача обнаружения людей на изображениях становится все более актуальной в области компьютерного зрения и распознавания образов. Это связано с возможностью применения алгоритмов обнаружения человека в системах видеонаблюдения для охраны объектов, контролем за техникой безопасности на промышленных предприятиях, наблюдением в торговых комплексах в целях изучения спроса и пр. Идеальный детектор должен обнаружить всех присутствующих на цифровом изображении людей независимо от их расположения, масштаба или позы. Однако автоматическое детектирование человека на снимке является довольно непростой задачей из-за ряда факторов: детектору приходится иметь дело с различным освещением, различным положением тела человека и одежды, сложным рельефом местности, частичным или полным перекрытием человека различными предметами или автомобилями. Также детектор должен отличать людей от подвижных механизмов, схожих по размеру с человеком (например, грейфер или магнит на мостовом кране). Кроме того детектор должен определять людей в большом диапазоне внутригрупповых изменений и при этом должен иметь как можно меньше ложных срабатываний на областях фона.

В связи с актуальностью и сложностью задачи детектирования фигуры человека на настоящий момент существует множество разнообразных подходов к ее решению. Так, предпринимались попытки осуществить детектирование сравнением предобработанных изображений с некоторыми эталонами (шаблонами) с помощью специальной метрики [1], использования детекторов и дескрип-

тов особых точек; применения сегментации изображения [2] и т. д. Одним из наиболее популярных и перспективных на настоящий момент подходов к детектированию объектов на изображении является так называемый метод бегущего окна. Данный метод основан на осуществлении экстенсивного поиска объектов определенного размера. В каждом кадре последовательно рассматриваются области изображения заданного размера и им ставится в соответствие признаковое описание, на основе которого с помощью алгоритма классификации принимается решение, содержит ли данная область объект или нет. Более подробно данный подход и особенности его реализации рассматриваются далее, здесь же отметим, что ключевыми компонентами системы детектирования, основанной на таком подходе, являются алгоритмы получения признакового описания изображения (HOG-дескриптор (Histogram of Oriented Gradients, гистограмма ориентированных градиентов)) [3] и его классификации. Данный метод был взят за основу детектора. В качестве классификатора, определяющего принадлежность объекта к группе «Человек» или «Фон», применили метод опорных векторов (SVM).

Для повышения вероятности распознавания присутствия людей в кадре алгоритм определения фигуры человека содержит ряд шагов (этапов):

1. Запускается детектор движения. Поскольку анализу подвергается изображение большого размера (1Мпиксель и более) то выполнить алгоритм распознавания человека методом HOG на всем кадре за приемлемое время практически невозможно. При этом для каждого кадра выпол-

няется сравнение с фоновым изображением и выделяются участки, в которых разница фонового изображения и текущего кадра превышает заданный порог. Затем модель фона обновляется. Участки движения сегментируются, и часть из них отбрасывается в случае, если их размер не удовлетворяет заранее заданному размеру человека.

2. Выделенные прямоугольники подаются на детектор, в котором происходит перемасштабирование исходя из заранее заданных минимального и максимального размеров человека. При этом используются данные о положении линии горизонта на изображении, а также допущение, что размеры человека линейно растут от минимальных (на дальнем плане, вблизи линии горизонта) до максимальных (на ближнем плане).

3. Для каждого окна строится вектор признаков из гистограмм направленных градиентов и производится классификация, которая определяет, находится ли в данном окне фигура человека или нет. Внутри каждой области происходит анализ на наличие человека из серии окон различного масштаба с шагом в 1%. Применение описанного подхода позволяет значительно сократить количество поисковых окон. Кроме того, за счет сужения области поиска людей уменьшается количество ложных срабатываний алгоритма, что приводит к повышению качества работы системы в целом.

4. Все детектируемые объекты записываются в буфер, в котором для каждого объекта хранятся:

- текущие координаты;
- история перемещения (трекинг) с данными о статусе («НОГ» – если на данном кадре объект определен с помощью алгоритма НОГ, или «трекинг» – если объект не подтвердился алгоритмом НОГ и сопровождался при помощи алгоритма оптического потока);
- время последней детекции алгоритмом НОГ;
- изображение объекта.

Для каждого последующего кадра происходит перебор всех ранее распознанных объектов и для каждого объекта снова происходит поиск алгоритмом НОГ в окне большего размера. Если объект снова распознан, обновляются данные по объекту в буфере, в противном случае запускается процедура трекинга и в буфере обновляются данные по объекту с указанием в статусе, что произведен трекинг.

Также с заданной периодичностью запускается детектор движения и выполняются шаги 1–3, а в буфер добавляются только новые объекты.

При анализе всех объектов, содержащихся в буфере, отображаются только объекты, для которых было несколько подтверждений распознавания алгоритмом НОГ и позиции распознаваний находились не в одном месте.

Данные ограничения позволили повысить вероятность распознавания людей на сложном рельефе (например: свалка металлолома) и исключить ложные срабатывания, неизбежно появляющиеся в таких случаях. На рис. 1 приведен один из видеок кадров во время детекции и трекинга людей.



Рис. 1

В детекторе движения применяется алгоритм вычитания фона из текущего кадра. Модель фона строится по следующей формуле:

$$F_k(x, y) = \alpha I_k(x, y) + (1 - \alpha)F_{k-1}(x, y),$$

где α – задает скорость обновления фона; $I_k(x, y)$ – значение интенсивности освещения в текущей точке; $F_{k-1}(x, y)$ – предыдущее значение модели фона.

На каждом шаге происходит сравнение интенсивностей в каждой точке модели фона и текущего кадра и при превышении формируется бинарное изображение. Далее над бинарным изображением проводятся морфологические операции эрозии и дилатации, после чего запускается алгоритм сегментации, который возвращает отдельные области.

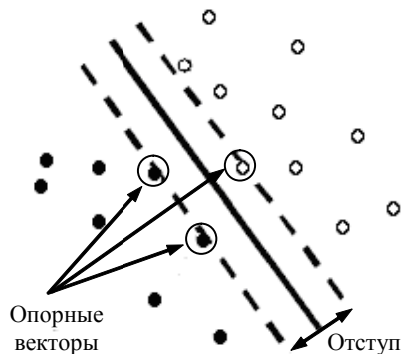


Рис. 2

В качестве алгоритма машинного обучения используется линейная машина опорных векторов (Linear SVM) [4]. Отметим, что натренированной моделью линейного SVM является плоскость, наилучшим образом разделяющая объекты от негативов в обучающей выборке (рис. 2). На рисунке отображены объекты, разделенные гиперплоскостью, которая отделяет группу объектов, имеющих различную классовую принадлежность. В этом примере объекты относятся либо к классу ЧЕРНОГО, либо БЕЛОГО цвета. Разделительная линия определяет границу, с правой стороны от которой все объекты БЕЛЫЕ, а с левой – ЧЕРНЫЕ. Любой новый объект, попадающий справа, классифицируется как БЕЛЫЙ (или ЧЕРНЫЙ, если он попадает слева от разделительной линии). Векторы, лежащие ближе всех к разделяющей гиперплоскости, называются опорными векторами (support vectors).

Для обучения классификатора требуется сформировать обучающую выборку, содержащую признаки описания изображений с объектом и без него. Так как источником таких изображений, как правило, являются фотографии, большую часть которых занимает фон, а не сам объект, большинство окон детектирования не будут содержать объект. В связи с этим требуется аккуратная балансировка прецедентов различных классов в обучающей выборке. Одним из подходов к обучению классификатора является итерационный

алгоритм добавления признаков описаний изображений фона бутстраппинг (Bootstrapping) [5]. На нулевой итерации в обучающую выборку добавляются признаки описания всех изображений с объектом, приведенных к нужному размеру, и выбранные случайным образом области изображений с фоном. С использованием полученной выборки осуществляется обучение классификатора, после чего выполняется детектирование объектов на изображениях фона, и выборка дополняется признаковыми описаниями ложных срабатываний детектора. Классификатор обучается на измененной выборке заново. Данный процесс повторяется несколько раз. Подобная процедура позволяет автоматически отрегулировать количество прецедентов в обучающей выборке, соответствующих изображениям без искомого объекта, и, тем самым, уменьшить количество ложных срабатываний полученного детектора.

Использование алгоритма HOG совместно с алгоритмами трекинга и детектором движения, а также информация о пространственных размерах фигуры человека и ограничения на возможные геометрические перемещения позволили создать детектор распознавания людей на сложном рельефе с вероятностью правильного распознавания не ниже 90 % и минимизировать ложные срабатывания, неизбежно появляющиеся в системах распознавания образов.

СПИСОК ЛИТЕРАТУРЫ

1. Viola P., Jones M. Rapid object detection using a boosted cascade of simple feature // Conference on Computer Vision and Pattern Recognition (CVPR). 2001.
2. Enzweiler M., Dariu M. Gavrilă. Monocular Pedestrian Detection: Survey and Experiments // Pattern Analysis and Machine Intelligence. 2009.
3. Dalal N., Triggs B. Histograms of oriented gradients for human detection // Conference on Computer Vision and Pattern Recognition (CVPR). 2005.
4. Vapnik V. N. The Nature of Statistical Learning Theory. Springer-Verlag, 1995.
5. Freund Y., Schapire R. E. A Short Introduction to Boosting // J. of Japanese Society for Artificial Intelligence. 1999. № 14(5). P. 771–780.

I. A. Goroshkov, S. A. Kulakov

DETECTOR OF PRESENCE OF PEOPLE IN THE DETECTION BLOCK

The article presents an algorithm for recognizing the presence of a person in the frame. The advantages of the selected algorithm for determining people. Mentions the possibility of applying this algorithm to ensure the safety and security of goods and materials to suit all types of businesses and activities.

An algorithm for recognizing people, HOG, machine learning, tracking