

НОРМАЛИЗАЦИЯ ОРИЕНТАЦИИ ИЗОБРАЖЕНИЙ НОМЕРНЫХ ЗНАКОВ ТРАНСПОРТНЫХ СРЕДСТВ НА ОСНОВЕ НЕЙРОННОЙ МОДЕЛИ 3D ДЕТЕКЦИИ

LICENSE PLATE NORMALIZATION BASED ON 3D NEURAL NETWORK

I. Saitov
A. Filchenkov

Summary. In license plate recognition systems, an important quality aspect is the position of the cameras, which are usually located on the side or on top of the vehicles. Thus, the license plate image exhibits various distortions that affect the optical character recognition. The normalization of rotation angle distortion based on computer vision techniques is an additional step in number plate recognition algorithms. We propose an approach that uses a 3D vehicle detection model to find the object's deflection angle for subsequent perspective transformation of the number plate. A dataset of license plate images belonging to previously predominantly CIS countries is used as data. Character recognition experiments using the proposed normalization show its effectiveness, achieving an accuracy value of 0.939 on the considered data.

Keywords: neural network, license plate recognition, 3D detection, optical character recognition, orientation normalization.

Саитов Ирек Аликович

Аспирант, Национальный исследовательский университет ИТМО (г. Санкт-Петербург)

isaitov@itmo.ru

Фильченков Андрей Александрович

к.ф.-м.н., Национальный исследовательский университет ИТМО (г. Санкт-Петербург)

Аннотация. В системах распознавания номерных знаков важным аспектом качества является положение средств фотовидеофиксации, которое расположено как правило сбоку или сверху от транспортных средств. Тем самым на изображении номерных знаков проявляются различные искажения, влияющие на оптическое распознавание символов. Нормализация искажения угла поворота на основе методов компьютерного зрения является дополнительным шагом в алгоритмах распознавания номерных знаков. Нами предлагается подход, использующий модель детекции 3D транспортного средства для нахождения угла отклонения объекта для последующего перспективного преобразования номерного знака. В качестве данных используется датасет изображений номерных знаков, принадлежащих ранее преимущественно странам из СНГ. Эксперименты по распознаванию символов с использованием предложенной нормализации показывают его эффективность, достигнув на рассматриваемых данных значения точности 0.939.

Ключевые слова: нейронная сеть, распознавание номерных знаков, 3D детекция, распознавание оптических символов.

Введение

В настоящее время автоматические системы распознавания номерных знаков транспортных средств получили своё распространение во многих прикладных областях, таких как регулирование дорожного движения и парковочных мест, детекция аварийных ситуаций, а также контроль за скоростным режимом.

Распространенный метод распознавания номерных знаков состоит из трех блоков, показанных на схеме рисунка 1. Процесс включает детекцию ограничивающей рамки или сегментацию номерного знака, выделение символов, а также последующее распознавание оптических символов. Рассмотренная процедура детекции номерного знака может быть очень сложной во всем механизме из-за того, что это оказывает прямое влияние на точность и эффективность зависимых этапов. Поэтому очень критичны для процесса распознавания условия окружающей среды на месте съемки, качество и характеристики устройства фотовидеофиксации, освеще-



Рис. 1. Классическая схема распознавания номерных знаков с возможным дополнительным (1.1) шагом нормализации

щение и случаи окклюзии. Многие предыдущие методы так или иначе решают подобные проблемы.

Однако при разработке систем распознавания номерных знаков зачастую не принимается во внимание вспомогательный шаг нормализации плиты перед этапом распознавания символов. Одним из преобладающих искажений номерного знака на изображении является наклон ориентации. В ранних работах [1] решение проблемы распознавания при поворотах камеры решалось путём использования свойств свёрточных нейронных сетей, которые учитывают геометрическое расположение пикселей, а также искусственным увеличением объёма тренировочной выборки методом аугментации различных поворотов изображений. В [2] было предложено корректировать вертикальный и горизонтальный наклон номерного знака путём подбора и выбора результата по максимуму значений серого цвета, что включает множество лишних вычислений.

Процесс исправления ориентации наклона номерного знака заключается в поиске углов отклонения относительно его проекций. Сложность поиска направлений проекций номерного знака исходит из двумерности своей формы. Однако исходя из факта жесткого крепления плиты автомобильного знака к передней или задней части транспортного средства, задача поиска ориентации сводится к нахождению ориентации движения транспортного средства. В настоящее время средства компьютерного зрения шагнули далеко вперёд. В частности, глубокие нейронные средства показали свою эффективность для различных задач детекции объектов. Одной из таких задач является выделение 3D ограничивающей рамки объекта по его фотографии. Использование подобной архитектуры позволяет определить положение транспортного средства на изображении и после выполнения проецирования определить углы отклонения объекта.

В нашем исследовании мы предлагаем метод нормализации ориентации номерных знаков на основе модели 3D детекции транспортного средства, который сочетается с существующими приёмами аугментации и позволяет улучшить результат распознавания номерных знаков.

Методология

Метод определения ориентации на основе модели 3D детекции

При рассмотрении возможных визуальных искажений номерных знаков на изображении можно выделить несколько типов, которые соответствуют поворотам объекта относительно трёх ортогональных векторов. Такие повороты номерного знака определяют его ориентацию в пространстве. Задача нормализации номерного знака

заключается в преобразовании его ориентации к исходному сонаправленному с точкой обзора. Изображение или видеокادر представляет собой двумерное отображение объектов на объектив камеры. При нормализации ориентации по двумерному изображению необходимо восстановить свойства трехмерного пространства для последующего преобразования.

При фотовидефиксации происходит преобразование точек реального мира на плоскость — матрицу камеры. Данный процесс упрощенно может быть описан моделью камеры с точечной диафрагмой. При фотфиксации объекта проецирование точек X глобальной системы координат в точки x плоскости изображения определяется формулой (1).

$$x = PX, \quad (1)$$

где P является проекционной матрицей между евклидовым пространством и изображением. Матрица P определяется выражением (2) со значениями калибровочной матрицы K , матрицы вращения R , единичной матрицы I и матрицы переноса C .

$$P = KR[I \mid -C]. \quad (2)$$

При рассмотрении задачи нормализации ориентации номерного знака знание внутренних параметров (калибровочной матрицы) позволяет находить внешние параметры (матрицы поворота и трансляции), что даёт возможность соотносить точки 2D пространства изображения и исходного 3D мира. Такая задача носит название «Perspective-n-Points» [3], что зачастую в настоящее время решается методами глубокого обучения.

Модель YOLO3D [4] относится к одноименному классу нейронных сетей. В основе модели лежит глубокая нейронная модель детекции с одной стадией YOLOv5 [5], совмещающая высокую скорость с эффективностью. Отличительной особенностью является решение задачи регрессии по восстановлению 3D ограничивающей рамки по исходной 2D рамке. Ограничение возможности применения подобной архитектуры непосредственно к номерным знакам заключается в плоскости плиты, которая не позволяет непосредственно определить его 3D форму. Однако будучи жестко закрепленным к передней либо задней поверхности транспортного средства ориентация номерного знака полностью совпадает с направлением машины. Тем самым задача поиска ориентации номерного знака может быть сведена к задаче поиска направления машины по его 3D форме.

При полученном векторе направлении необходимо определить ориентацию на снимке нормального образца. В данной работе нами использовано допущение по направленности глобальной системы координат сле-

дующего образа: ось OZ и OX сонаправлены сторонам изображения, ось OY направлена перпендикулярно к наблюдающему. Угол отклонения по каждой из осей определяется исходя из свойств косинусной меры между

$$\cos(\theta_a) = \frac{(u_a, v_a)}{|u_a| |v_a|},$$

где u, v векторы ориентации транспортного средства и глобальный вектор, $a \in \{x, y, z\}$.

Алгоритм нормализации ориентации на основе модели 3D детекции

При рассмотрении задачи нормализации ориентации номерного знака возникают несколько случаев, что требует дополнительной алгоритмической обработки.

В частности, на одном фотоснимке может присутствовать несколько транспортных средств. Нами было протестировано, что модель предсказания номерного знака эффективнее работает на целевом изображении, нежели последовательном выделении объекта транспортного средства, а затем номерного знака. Кроме того, автомобиль может лишь частично быть представлен на изображении, что затрудняет его распознавание. Соотнесение объектов в таком случае может быть произведено по расположению объектов и определению степени их пересечения. Нами был разработан алгоритм, подробно описанный на Рисунке 2. На вход подается кадр съёмки. Результатом его работы является массив нормализованных изображений номерных знаков.

Эксперименты

Данные

Набор данных Nomeroff Net [6] является как нам известно единственным полноценным датасетом, включающим данные изображений транспортных средств и номерных знаков из стран, ранее преимущественно относящимся к СНГ. Одной из отличительных особенностей датасета является большое разнообразие положений камер. При отсутствии знания о калибровочной матрице (см. Рисунок 3) качество распознавания показывает неточные результаты. Существуют методы определения калибровочной матрицы по изображениям [7, 8], однако для их использования необходимо 3, а лучше сильно больше снимков с одной камеры.

Для имитации процесса распознавания номерных знаков в реальных условиях окружающей среды были взяты данные изображений номерных знаков и 3D объект машины, к которому параметрически прикреплялась плита при рендеринге и среда с определенными характеристиками калибровочной матрицы K . Всего было взято 1000 различных объектов номерных знаков.

Метрики

В работе было использовано несколько различных метрик. В ходе обработки кадров видеофиксации транспортных средств зачастую приходится сталкиваться с проблемой обнаружения нескольких объектов на од-

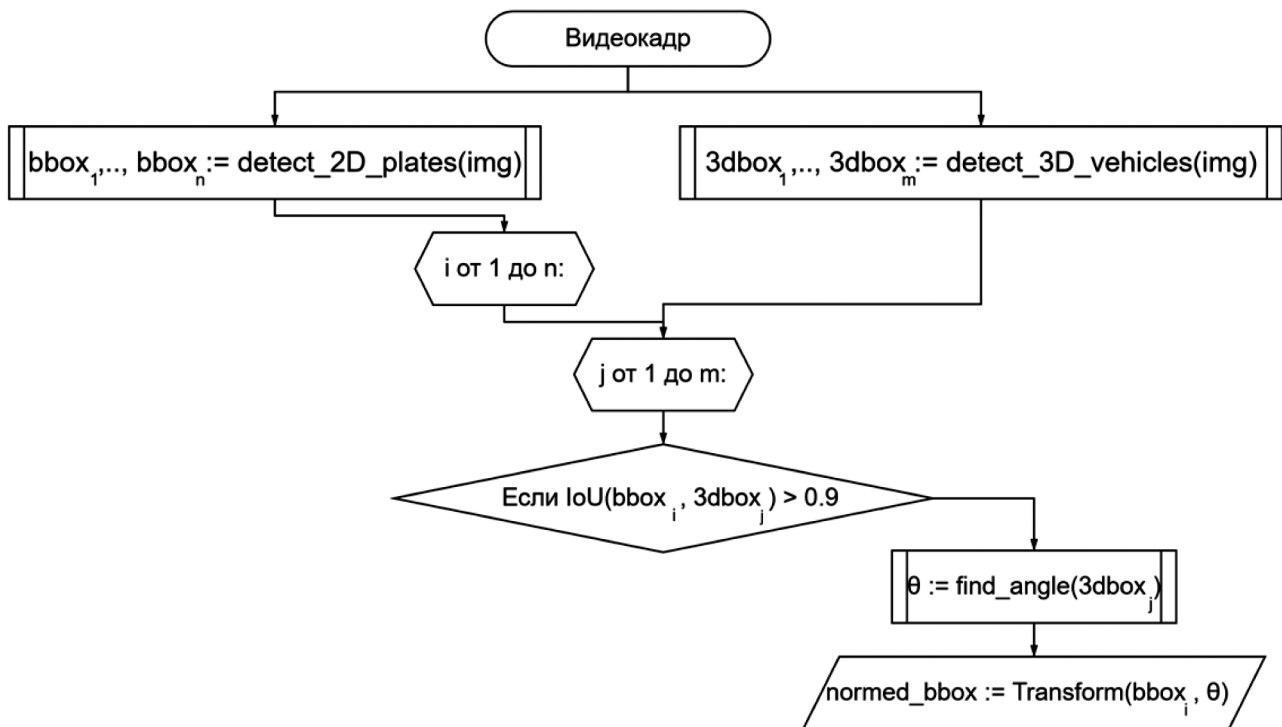


Рис. 2. Алгоритм нормализации поворота номерного знака с использованием модели 3D детекции



Рис. 3. Пример распознавания YOLO3D ограничивающей рамки на изображении. Слева представлен результат с известной геометрией исходного обучающего множества. Справа неточное распознавание по тестовому изображению из целевого датасета

ном снимке. Для этого в разработанном алгоритме учитывается сопоставление объектов. Путём сравнения положения номерного знака и лицевой стороны транспортного средства. Данное сравнение выполняется по метрике Жаккара (англ. IoU, Intersection over Union), представляющем собой следующее выражение:

$$IoU = \frac{|A \cap B|}{|A \cup B|},$$

где A является областью номерного знака, а B — областью лицевой стороны транспортного средства. Правильно сопоставленные объекты считаются при превышении порогового значения $IoU \geq 0.9$.

Для определения качества работы разработанного метода нормализации номерных знаков использовалась косвенная оценка улучшения качества распознавания символов по метрике точность. Правильно положительной считалась строка номерного знака со всеми верно определёнными символами.

Настройки экспериментов

Эксперименты были проведены с использованием GPU ускорения видеокарты NVIDIA GeForce GTX 1080 Ti объемом памяти 12 ГБ.

В качестве дополнения данных путём рендеринга была задана параметризация в условных единицах. Объект транспортного средства задавался размерами 1500x2000x3000. Номерной знак крепился к центральной фронтальной части автомобиля на высоте 150 относительно нижней границы. Размеры плиты приводились к единому размеру 260x110. Для симуляции относитель-

ного положения камеры объект транспортного средства вместе с каждым номерным знаком были повернуты случайно 10 раз с использованием аффинного преобразования внешнего вращения соответственно по осям в промежутках $\{\theta_x, \theta_y \in \mathbb{R}; 10 \leq (\theta_x, \theta_y) \leq 60\}$, $\{\theta_z \in \mathbb{R}; 5 \leq \theta_z \leq 30\}$. По окончании рендеринга при фиксированном положении камеры на удалении 10000 были взяты изображения, используемые в дальнейшем в качестве тестовых данных.

Результаты

В ходе экспериментов мы оценили время, затрачиваемое на выполнение методов Nomeroff Net и 3D нормализации ориентации номерных знаков. Результаты приведены в Таблице 1, по которым можно сделать вывод о быстродействии базового метода Nomeroff Net. Это может быть объяснено ресурсоёмкостью использования 3D модели детекции, включаемой в разработанный алгоритм 3D нормализации ориентации.

Таблица 1.

Сравнение исполнительного времени методов нормализации ориентации.	
Более эффективный результат выделен жирным шрифтом	
Метод	Nomeroff Net
Время (мс)	3D нормализация
	4.17 ± 0.114
	923 ± 30.9

Для сравнения качества разработанного метода нормализации проводились эксперименты на оригинально

сгенерированных данных без нормализации, а также с использованием методов Nomeroff Net и 3D нормализации ориентации. Как видно из Таблицы 2, методы нормализации ориентации улучшают качество распознавания символов номерных знаков. При этом разработанный алгоритм 3D нормализации превосходит по метрике точности со значением 0.939.

Таблица 2.

Результаты экспериментов с использованием методов нормализации и без на датасете Nomeroff Net с приведенной процедурой параметризации данных по метрике точности распознавания символов.

Лучший результат выделен жирным шрифтом.

Метод	Оригинальные данные	3D нормализация Nomeroff Net
Точность	0.905	0.939
	0.924	

Выводы и дальнейшая работа

Был разработан новый алгоритм нормализации ориентации номерных знаков на основе модели 3D детекции транспортного средства. Для проведения экспериментов используемый датасет Nomeroff Net был дополнен разработанной параметризованной генерацией данных. Результаты экспериментов на наборе данных Nomeroff Net показали точность распознавания. Из ограничений применимости можно выделить необходимость знания калибровочной матрицы для повышения качества точности 3D детекции транспортных средств, а также относительная большая ресурсоёмкость предложенного метода.

В дальнейшем планируется реализовать модуль определения характеристик проективной камеры сцены по нескольким снимкам. Для этого необходимо провести работу в исходном датасете Nomeroff Net по выделению снимков с одного ракурса и единой камеры. Также можно собрать данные снимков путём формирования соответствующих запросов в интернете и открытых базах данных.

ЛИТЕРАТУРА

1. Друки А.А. Применение сверточных нейронных сетей для выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном // Известия Томского политехнического университета. Инжиниринг георесурсов. — 2014. — Т. 324. — №. 5. — С. 85–92.
2. Wang Y., Chen Y., Wang M. A new vehicle license plate correction method based on Sobel operator and priori knowledge // 2016 6th International Conference on Machinery, Materials, Environment, Biotechnology and Computer. — Atlantis Press, 2016. — С. 1065–1073.
3. Lu X.X. A review of solutions for perspective-n-point problem in camera pose estimation // Journal of Physics: Conference Series. — IOP Publishing, 2018. — Т. 1087. — №. 5. — С. 052009.
4. Mousavian A. et al. 3d bounding box estimation using deep learning and geometry // Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. — 2017. — С. 7074–7082.
5. Jocher G. et al. ultralytics/yolov5: v6. 0-YOLOv5n'Nano'models, Roboflow integration, TensorFlow export, OpenCV DNN support // Zenodo. — 2021. Doi:10.5281/zenodo.3908559.
6. AUTO-RIA [Электронный ресурс] // Nomeroff Net фреймворк распознавания номерных знаков. URL: <https://github.com/ria-com/nomeroff-net>. Дата обращения 01.03.2024.
7. Clarke T.A., Fryer J.G. The development of camera calibration methods and models // The Photogrammetric Record. — 1998. — Т. 16. — №. 91. — С. 51–66.
8. Zhang Y.J. Camera calibration // 3-D Computer Vision: Principles, Algorithms and Applications. — Singapore: Springer Nature Singapore, 2023. — С. 37–65.

© Саитов Ирек Аликович (isaitov@itmo.ru); Фильченков Андрей Александрович
Журнал «Современная наука: актуальные проблемы теории и практики»