

УДК 681.3

МАРКИРОВКА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

М.Э. КуССуль, А.С. Сычов, Е.Г. Садовая

Институт Проблем Математических Машин и Систем НАН Украины

e-mail: owl@nntean.org.ua

Введение

В данное время создается много специализированных систем поддержки принятия решений, основанных на анализе и классификации видеопоследовательностей. Одним из важных этапов при создании таких систем является маркировка данных. Маркировка необходима, во-первых, для определения качества тестирования разрабатываемых систем, во-вторых, для использования в алгоритмах обучения или настройки на исходные данные, которые зачастую используются в таких системах. В большинстве работ обсуждению процесса маркирования не уделяют особого внимания, сосредотачиваясь на описании самих задач и алгоритмов их решения. В публикациях процесс маркировки остается, так сказать, за кадром, несмотря на то, что это очень трудоемкий процесс, к тому же он в значительной степени способен влиять как на эффективную подготовку таких систем к работе, так и конечное качество их работы.

В данной статье мы хотим обсудить маркировку видеопоследовательностей на примере маркировки данных для системы распознавания пешехода, очертить круг проблем, которые возникают при маркировании оптических данных и показать пути их решения.

1. Постановка задачи обнаружения пешехода

Задача обнаружения пешехода на дороге, основанная на видео-зрении, является частью системы поддержки принятия решений, которые уже начали устанавливаться на автомобили и помогают водителю принять правильное решение в критической ситуации. Данная задача широко обсуждается в литературе [1]-[4].

Постановка задачи определена следующим образом: портативная цветная видеокамера с разрешением 360x288 пикселей и скоростью сбора данных 24 кадра в секунду закрепляется в автомобиле. Рабочая дистанция для поиска пешехода определена в пределах от ~10 до ~35 метров при скорости автомобиля от 0 до 50 км/час. Необходимо по одному или нескольким последовательным изображениям определить наличие на них пешехода (или группы пешеходов) и указать его приблизительное местоположение.

Система должна работать в естественных условиях, а это значит, в любое время года (лето, весна/осень, зима); при произвольных погодных условиях (солнечно, пасмурно, идет дождь, идет снег, ветрено); при произвольном взаимном поведении автомобиля и пешехода (оба неподвижны, оба движутся, включая резкие повороты автомобиля, движется только либо автомобиль, либо пешеход); при различных вариантах окружающей среды (разные варианты дорожного покрытия, различные ширина и уклон дороги, например, холмы); допускается присутствие деревьев, зданий, столбов, заборов, а также, наличие других движущихся объектов, в частности, автомобилей. Естественно, сам пешеход может находиться в любом ракурсе, кроме принятого ограничения, что пешеход не может лежать или сидеть, и его изображения может находиться в произвольном месте на видеоизображении.

Очевидно, что наличие такого большого количества вариантов различных ситуаций и условий, которые должны быть представлены как в обучающей, так и в тестовой выборке, говорит о необходимости маркирования большого числа первичных данных.

2. Проблемы, возникающие при маркировке неформализованных оптических объектов

Решая задачу маркировки большого количества разнообразных данных, мы столкнулись с некоторыми проблемами. При маркировке каждый объект на изображении может быть плотно очерчен некоторым прямоугольником так, чтобы контуры объекта полностью попали в прямоугольник. Размеры окантовывающих прямоугольников, следовательно, должны соответствовать размерам объекта. Но размеры объектов на изображении имеют большой диапазон изменения и зависят от расстояния, на котором объект находится от камеры. Определив рабочую дистанцию в пределах от ~10 до ~35 метров, получаем, что линейные размеры пешехода в рабочей области системы могут отличаться друг от друга довольно сильно, а это усложняет маркировку данных, поскольку оператор должен не только определить местоположение пешехода, но и его размер. Данную проблему можно решить, если разбить весь спектр изменения размеров объекта на дискретное количество областей, или зон. Зоной будем называть некую полосу на изображении, в которую попадают пешеходы определенных размеров (их окантовывающие прямоугольники имеют соответственно близкие размеры). Каждой зоне соответствует определенное расстояние пешехода до бампера автомобиля.

При решении нашей задачи мы ограничились тремя такими зонами: первая зона – размер прямоугольника равен 57x140 пикселей, что соответствует расстоянию от 7 до 18 метров; вторая зона – размер прямоугольника 37x85 пикселей при расстоянии в 16 до 40 метров; третья зона включает пешеходов, находящихся далее 30 м.

На рисунке 1 показан кадр, в котором производится разметка. Справа виден курсор (красный цвет) в виде трех вложенных прямоугольников, при маркировании оператор должен навести данный курсор на пешехода, центрировать его и задать номер зоны, к которой принадлежит данный пешеход. После маркировки вокруг пешехода остается прямоугольник, соответствующий номеру его зоны (зеленый цвет). Преимущество данного подхода состоит в том, что центрирование рамки и задание номера зоны требует гораздо меньше времени, чем задание границ пешехода, а поскольку данная операция повторяется при маркировке данных многократно, то такая оптимизация довольно существенно сказывается на затратах рабочего времени оператора.

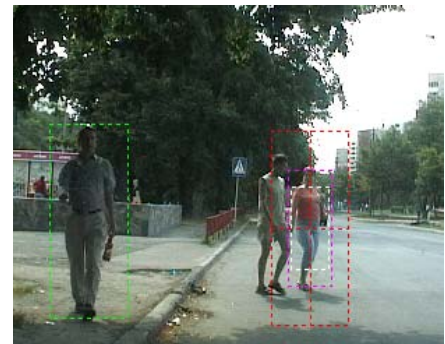


Рис. 1 Кадр с разметкой пешехода

Следующая проблема, которая возникает при маркировании, – как маркировать объект, когда он частично закрыт. Такая ситуация встречается довольно часто, поскольку пешеход бывает частично закрыт деревом, автомобилем или другим пешеходом. Возможно несколько вариантов решения данной проблемы. Например, для нашей задачи возможны такие варианты маркировки частично закрытого пешехода либо как пешехода, либо как отсутствие пешехода, либо выделить данные случаи в отдельный класс, либо не маркировать их вовсе. Выбор приемлемого варианта, как правило, зависит от дальнейшей работы алгоритмов распознавания. В нашем случае оказалось, что наилучшие результаты получаются, если не маркировать такие случаи вовсе. Во-первых, экономится время, во-

вторых, не вносится дополнительная неопределенность (оператору приходится субъективно оценивать на сколько сильно закрыт пешеход).

Еще одна проблема, на которой хотелось бы остановиться, - это нахождение контр-примеров. Для обучения нейронной сети нужны не только примеры с изображением пешехода, но также и примеры с его отсутствием. Самый легкий путь решения – это, если считать все пространство за пределами окантовывающих прямоугольников как не содержащее пешехода, то тогда можно покрыть его прямоугольниками соответствующих размеров и получить примеры отсутствия пешехода. Однако в этом случае мы получаем слишком большое количество примеров отсутствия пешехода, их будет в сотни раз больше, чем примеров, содержащих пешехода, и большинство из них не будут нести полезной информации. Это могут быть, например, десятки тысяч практически одинаковых прямоугольников с изображением асфальта, которые только засоряют обучающую последовательность. Поэтому нами был разработан простой алгоритм поиска так называемых зон интереса, который основывается на подсчете количества контуров в выбранном прямоугольнике. Если количество контуров в данном прямоугольнике превышает некоторое пороговое значение, то такой прямоугольник считается интересным и попадает в обучающую выборку как пример, не содержащий пешехода. В противном случае он пропускается. Данный подход позволяет резко сократить объем обучающей выборки без потери качества работы системы.

3. Способы автоматизация маркировки данных

Трудоемкость процесса маркирования делает актуальным вопрос о его автоматизации. Полная автоматизация маркирования данных невозможна поскольку, если бы существовал такой алгоритм, позволяющий правильно маркировать данные, то он бы и являлся решением поставленной задачи. Однако возможно создать алгоритм автоматизации, который бы работал с некоторой погрешностью, которую оператор мог бы устранить в дальнейшем во время просмотра результатов работы алгоритма. Такой алгоритм был разработан, хотя в окончательном варианте системы от него пришлось отказаться.

В основу алгоритма легла наша же система распознавания пешехода. Суть его заключается в следующем: оператор вручную размечает небольшое количество первичных данных. На их основе проводится обучение нейронной сети, и настраиваются все необходимые алгоритмы для нахождения пешехода. Когда количество правильных ответов системы становится около 60-70% (как правило, для этого необходимо маркировать небольшое количество данных - около 10-15% от всего количества), систему распознавания пешехода используют в качестве помощника оператора при маркировке последующих изображений. Естественно, оператор производит контроль и исправляет ошибки системы. После маркировки следующей части данных система переучивается, качество ее работы повышается, что соответственно увеличивает продуктивность работы оператора.

Данный метод позволяет довольно сильно ускорить процесс маркирования данных, однако, он имеет и существенные недостатки. Главным недостатком является неточность позиционирования прямоугольника относительно пешехода. Часто пешеход не попадает в центр данного прямоугольника либо частично выходит за его пределы. Если оператор будет исправлять такие ошибки, то ускорение процесса по времени не происходит. Если же не производить центрирование пешехода вовсе, то качество работы системы резко ухудшается. Как показали проведенные нами эксперименты, при использовании данных,

маркированных полностью вручную, качество работы системы возрастет более чем в 10 раз.

Однако, не взирая на отрицательный опыт, на наш взгляд, не следует полностью отказываться от данной идеи автоматизации процесса маркирования. В дальнейшем если удастся дополнить данный алгоритм простыми и надежными алгоритмами центрирования объекта, он поможет резко сократить время работы оператора без существенных потерь качества.

4. Искусственное увеличение обучающей выборки

Существует еще один путь увеличения количества маркированных данных – это искусственная генерация данных.

Чтобы увеличить число данных, мы применили метод, известный в литературе по распознаванию образов, как «подвижка» или «дисторсия» [5]. Метод заключается в небольших перемещениях уже маркированных прямоугольников вокруг места, которое маркировано оператором. Каждый предварительно маркированный прямоугольник должен быть смещен относительно начальной позиции на $\pm 1-2$ пикселя по вертикали и ± 1 пиксель по горизонтали, включая взаимные (диагональные) сдвиги. Тогда число маркированных данных увеличивается в 15 раз. При таком увеличении размеров обучающей выборки качество работы системы увеличивается в 2-3 раза, что было проверено экспериментально.

Выводы

Мы очертили ряд вопросов возникающих в процессе маркирования данных и показали пути их решения. Показали насколько значительную роль может иметь процесс маркирования в системах распознавания оптических данных. Он не только занимает значительную часть времени в процессе разработки, но и может существенно влиять на качество работы всей системы.

Таким образом, благодаря качественному маркированию данных удалось улучшить работу системы обнаружения пешехода более чем в 20 раз, что позволило получить результат до 99.95% правильных ответов.

Литература

1. М.Э. Куссуль, О.Г. Садовая, А.С. Сычев Обзор алгоритмов и решений задачи оптического распознавания пешехода // Математические машины и системы.- 2005.- №4.- С. 110
2. D.M. Gavrila, J. Giebel and S. Munder. Vision-Based Pedestrian Detection: The PROTECTOR System. Intelligent Vehicles Symposium, 2004 IEEE, 16-18, 2004.
3. I.Gat, M.Benady and A.Shashua. A Monocular Vision Advance Warning System for the Automotive Aftermarket. SAE World Congress & Exhibition 2005.
4. Shashua, A.; Gdalyahu, Y.; Hayun, G. Pedestrian detection for driving assistance systems: single-frame classification and system level performance. IEEE Intelligent Vehicles Symposium, pp.1-6, 2004.
5. Y. LeCun, L. Bottou, G. Orr, and K. Muller, "Efficient BackProp," in *Neural Networks: Tricks of the trade*, (G. Orr and Muller K., eds.), 1998.