

УДК 004.896:621.865

ОБ ОДНОМ МЕТОДЕ РАСПОЗНАВАНИЯ ОБЪЕКТОВ С НЕ ПОЛНОСТЬЮ ОПРЕДЕЛЕННЫМИ ПРИЗНАКАМИ

А.Д. Московский (*moscowskyad@gmail.com*)
НИЦ «Курчатовский институт»

Аннотация. В работе описывается комплексный метод распознавания объектов на изображении в условиях неполноты входных данных, для задач навигации мобильных роботов. Представлен разработанный программный модуль для ROS, представлены результаты экспериментов, эмулированных в среде Gazebo¹.

Ключевые слова: мобильная робототехника, распознавание, техническое зрение, реконструкция сцен, недоопределенные модели, доопределение атрибутов.

Введение

Современная мобильная робототехника не обходится без задач технического зрения. Это связано, как с доступностью таких сенсоров, как камеры, так и с огромным потоком полезной информации, предоставляемым этими сенсорами. Разработано много методов, позволяющие распознавать разные элементы окружающего пространства. Среди всех можно выделить несколько основных подходов. Первая группа методов занимается извлечением из изображений трехмерных данных. Изначально это были подходы, рассматривающие стерео зрение [Horn и др., 1988], а затем так называемое SfM (Structure-from-Motion – Строение из Движения) [Huang и др., 1989]. Последний подход сосредотачивается на сравнении «близких» друг к другу изображений, сделанных во время движения камеры. Похожий подход активно был использован в целом классе навигационных алгоритмов SLAM (simultaneous localization and mapping – одновременная локализация и картирование) [Davison и др., 2007] и [Engel и др., 2014]. Также в подобных методам относится метод извлечения трехмерных данных по изображениям снятым с разных точек наблюдения [Nistér, 2004]. Другая большая группа методов сосредоточена на детектировании объектов в окружающем пространстве.

¹ Работа выполнена при частичной поддержке гранта РФФИ 15-07-07483

Данные методы стараются ответить на вопросы «что?» и «где?» [Fallis, 2013]. Среди них тоже можно выделить несколько основных групп методов, одна занимается вопросами распознавания а классификации, кластеризации и категоризации объектов [Dickinson и др., 2010]. Большой интерес для данной работы представляют методы, выделяющие непосредственно объекты на изображении, большая группа посвящена методам, разбивающие объект на отдельные геометрические примитивы и детектирующие их, например алгоритмы RBC (Recognition-by-Components) [Biederman, 1985] и ACRONYM [Zisserman и др., 1995], данные подходы были созданы в попытке воссоздать человеческий подход к распознавание в техническом зрении. Очень актуальной в своё время являлась задача распознавания лиц, породившая такие методы, как распознавание образов [Turk и др., 1991] и каскадами Хаара [Viola и др., 2001]. Еще одна наиболее интересная группа методов занимающаяся выделением ключевых точек на изображении при помощи дифференциальных и градиентных методов [Lowe, 2004]. Наиболее известными и широко распространёнными методами являются алгоритмы SIFT [Lowe, 1999] и SURF [Bay и др., 2006]. Помимо выделения ключевых точек, существуют и алгоритмы выделения ключевых областей [Matas и др., 2004].

Все перечисленные методы обычно работают по принципу детектирования, т.е. либо находят данных объект на изображении, либо утверждают, что его нет. Решение об отсутствии объекта может быть вынесено ложно, например, по причине плохой видимости искомого объекта. Факторами, мешающими распознаванию, могут быть шум, перекрытие другими объектами, необычное освещение, неудачный ракурс наблюдения и др. Иными словами входная информация неполная (или недоопределенная). Однако методы работы с изображениями в условиях неопределенности достаточно редки, или специфичны для какого-то определенного класса объектов. Цель данной работы – предложить механизм обнаружения объектов на изображении (используя в основе существующие подходы к распознаванию), с помощью которого можно было бы выдвигать предположения о нахождении объектов, в условиях неопределённости.

1. Постановка задачи

Для многих задач мобильной робототехники была бы полезна информация о возможном нахождении искомого объекта, нераспознанного из-за неполноты наблюдаемых данных. Для того, что бы детектировать объекты в условиях недоопределённости, предлагается рассматривать каждый объект как совокупность неких признаков, которые в свою

очередь, могут быть отдельно от объекта быть распознаны неким существующим методом распознавания. Признаки должны подбираться таким образом, что бы частично видимый объект можно было потенциально (т.е. с некоторой оценкой уверенности) идентифицировать. Конечно, такой подход не сможет гарантировать нахождение объекта в случае его частичной видимости, однако должен позволить сделать заключение вида «возможно требуемый объект здесь» и сопоставить с этим некую числовую оценку уверенности. При таком подходе, однако, будут и отрицательные эффекты: увеличатся случаи ложного распознавания. Избежать этой проблемы можно расширяя информацию об объекте путем добавления информации не визуального характера (раз речь идет о мобильной робототехнике, то это всевозможные дополнительные сенсоры), а также понятие «контекста» объекта. Дальнейшие действия зависят от глобальной задачи робота, в каких-то случаях придется «поверить» методу, в каких-то «захватить» объект и следить за ним, до тех пор, пока точность распознавания не увеличится. В дальнейшем этот процесс будет описан как доопределение атрибутов.

2. Описание подхода

2.1 Объекты и атрибуты

Метод распознавания работает с понятием объекта. Объектом называется некая сущность, которую требуется распознать посредством технического зрения. Обычно объект отождествляется с предметами в окружении робота, например «стул», «чашка», «дверь» и т.п. Распознать объект означает выделить на получаемом от камеры изображении область, соответствующую реальному положению объекта в пространстве.

Как говорилось, идея метода «разложить» объект на общие для всех объектов составляющие. Данные составляющие совершенно естественно описываются признаками объекта, или же атрибутами. В качестве атрибутов объекта выступают такие понятия как форма, цвет и узор (текстура), а также ряд пространственных атрибутов, таких как отношение линейных размеров и положение в пространстве. Здесь приведен рабочий перечень атрибутов, расширять данный список можно какими угодно признаками, например, добавить анализ динамики объекта (подвижный\неподвижный) или рассматривать тени, отбрасываемые объектами. Каждый атрибут принадлежит своему пространству. Например, атрибут цвета может характеризоваться точкой или областью в трехмерном пространстве RGB или HSV, в то время как форма может принимать бесконечно много значений, и поэтому пространство представляется списком элементов, внесенных в каждую конкретную

реализацию. Для распознавания отдельных атрибутов, могут использоваться различные методы распознавания. Также следует отметить, что одни атрибуты могут лучше характеризовать объект, чем другие, для этого введем коэффициент значимости для каждого атрибута. Таким образом, каждый объект характеризуется набором пар атрибут-коэффициент:

$$O_i = \{(a_1, k_1) \dots, (a_N, k_N)\}, a_j \in A_j, k_j \in \mathbb{R}, \quad (1)$$

где O_i – объект, a_i – значение атрибута из подпространства атрибутов A_i , k_i – соответствующий атрибуту коэффициент из действительных чисел.

Однозначно распознанным в рамках данного подхода объект называется таким, у которого все атрибуты были распознаны в одной и той же области изображения. При этом, чем больший набор атрибутов будет у объекта, тем сложнее будет спутать этот объект с другими.

2.1 Комплексные объекты и сцены

Однако подобное описание объекта не удовлетворяет всем нуждам задач в робототехнике. Некоторые объекты не подвергаются подобному описанию, т.к. содержат несколько разных значений одного и того же атрибута, например объект многоцветный. Для этого было введено понятие комплексного объекта, представляющего собой набор простых объектов и отношений между ними. Отношениями называются связи между объектами, в основном используются пространственные отношения [Варосян и др., 1982], а также отношения линейных размеров. Такое расширение позволяет описать больший спектр объектов. И при этом есть возможность также объединять отношениями комплексные объекты, такие объединения согласно [Minsky, 1975] будем называть сценами.

2.2. Проблема частичной видимости

Подходя к проблеме частичной видимости, возникают два случая. Рассмотрим случай, когда частично виден простой объект. В данном случае могут распознаться в одной области не все атрибуты, например форму не удалось распознать, т.к. она закрывается другим объектом, а цвет наоборот удалось, как и примерный размер, не смотря на перекрытие. Два из трех атрибутов совпадет, далее введем коэффициент уверенности CC определяемый как отношение суммы коэффициентов распознанных атрибутов, к сумме всех атрибутов объекта.

$$CC_o = \frac{\sum_{detected} k}{\sum_{total} k}, \quad (2)$$

где, k – коэффициенты атрибутов объекта из формулы 1.

Коэффициент уверенности принимает значения от нуля до единицы,

соответственно, если сила равна единице, объект считается однозначно распознанным.

Второй случай описывает ситуацию, когда комплексный объект или сцена частично видимы. В данном случае один или несколько объектов могут быть находиться в частичной видимости или отсутствовать вовсе. В таком случае коэффициент уверенности комплексного объекта или сцены определяются как отношение суммы по всем связям полусумм коэффициентов уверенности каждого объекта к общему количеству связей:

$$CC_s = \frac{\sum_{ratio} 0.5(CC_{Oj} + CC_{Oi})}{N_{ratio}}, \quad (3)$$

где CC_{oi} – коэффициенты уверенности простых объектов, рассчитанных по формуле 2, N_{ratio} – общее количество связей в комплексном объекте. Аналогично простому объекту коэффициент уверенности принимает значения от нуля до единицы и равен единице в том случае, когда все составляющие объекты однозначно распознаны, как и состоят в заданных отношениях друг с другом.

2.3 Недоопределенные модели

Как можно было заметить, объект представляет собой элемент много размерного пространства, причем каждая размерность может характеризоваться разными подпространствами, как было сказано раньше подпространство цвета – трехмерное, подпространство формы – набор дискретных значений. Однако для описания каждого атрибута требуется прибегнуть к унифицированному механизму. В роли такого механизма выступают недоопределенные модели (далее Н-модели) Нариньяни [Narin'yani и др., 1997]. Согласно парадигме Н-моделей объект на стадии распознавания является недоопределенной переменной, т.е. заданной не точно, а каким то интервалом. Данный интервал должен сужаться на каждом следующем шаге работы алгоритма, стремясь к точному значению.

Т.к. данные методы предполагается использовать в задачах мобильной робототехники, то не исключается возможность улучшения условий наблюдения на последующих шагах функционирования робота. Иными словами методы детектирования можно применять только к областям (или их расширениям), показавшим положительную силу на предыдущем шаге, при этом коэффициент у верности каждого следующего детектирования должен быть не хуже предыдущего.

$$f_t : D_{image} = D_t, CC_{Dt} > 0 \quad (4)$$

$$f_{t+1} : D_t = D_{t+1}, CC_{Dt+1} \geq CC_{Dt}, \quad (5)$$

где f_t – функция детектирования на шаге t , D – область изображения. Такой подход назовем доопределением атрибутов.

3. Эксперименты

Чтобы проверить применимость данного метода была создана система распознавания, включающая в себя детекторы для определения следующих атрибутов:

- Попиксельный детектор цвета.
- Гистограммный детектор цвета.
- Детектор формы каскадами Хаара [Viola и др., 2001].
- Детектор контуров Кэнни [Canny, 1986].
- Детектор линейных размеров объектов.

Также система способна распознавать пространственные отношения между объектами. Система написана на языке C++ под ROS (<http://wiki.ros.org/indigo>). Архитектура построена таким образом, что база детекторов легко расширяема. Эксперименты проводилась на модели, построенной в среде Gazebo (<http://gazebo.org>). В качестве распознаваемых объектов были использованы двухкомпонентные цветные маркеры (как пример комплексного объекта) и морские навигационные знаки сторон света - вежи.

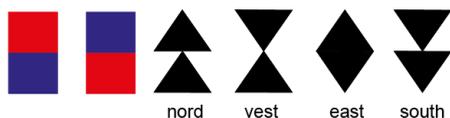


Рис. 1. Распознаваемые объекты

Окружение представляло собой модель проходного холла с дверьми колоннами и текстурным полом.

На рисунке 2 представлено распознавание вежи «восток» (см. рис. 1). Слева дан рисунок с коэффициентом уверенности ограниченным снизу значением 0.6, справа - значением 0.8.

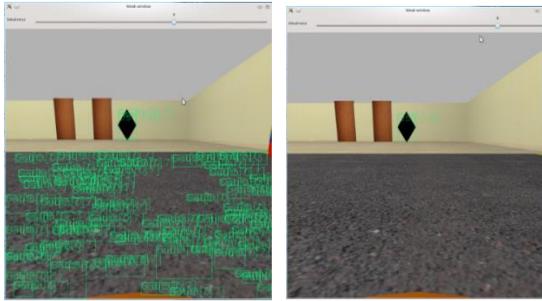


Рис. 2. Результат распознавания вехи «восток» при ограничении снизу коэффициента уверенности 0.6 (слева) и 0.8 (справа)

Описание вехи «восток» состояло из четырех атрибутов: характерная ромбовидная форма, черный цвет, продолговатый в вертикальной оси силуэт и средний размер (предполагается, что робот не должен видеть объект близко или вдали). Как видно из рисунка, метод определяет текстуру пола как искомые объекты. Коэффициент доверия у этих ложных объектов сильно ниже, чем у верно распознанного знака. Также эти объекты носили «случайный» характер и являлись ошибкой применения цветовой гистограммы к изображению. Это удалось исправить, введя еще один атрибут – повторяемость объекта. Данный атрибут давал свой вклад, только если на предыдущем шаге алгоритма этот же объект был обнаружен примерно в этой же области и с похожими другими атрибутами. Добавление нового атрибута увеличило способность распознавать объекты данного типа, уменьшив коэффициент доверия ложных объектов примерно на 0,15. Текстуриный пол дает большое количество ложных объектов, на рис. 3 показан сравнительный график, поэтому важно ограничивать для каждого объекта нижний порог распознавания.

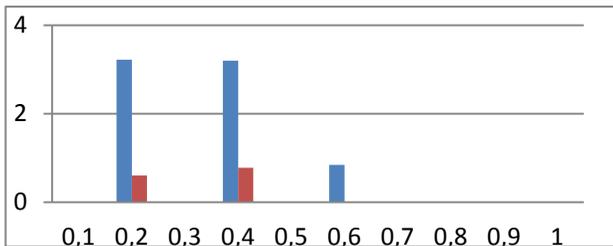


Рис. 3. Сравнительная гистограмма наличия ложных объектов при видимом текстурином полу (синий ряд) и при его отсутствии (красный ряд). По оси абсцисс отложен коэффициент уверенности, по оси ординат логарифм количества ложных объектов на изображении

На рисунке 4 представлена ситуация с частичным перекрытием распознаваемого объекта элементами окружения. Веха «восток» оказалась наполовину закрыта колонной, и распознать ее форму каскадом Хаара не удалось. Тем не менее, система предположила, что данный объект может находиться в данной области. Однако аналогичное представление было сделано касательно вехи «юг» также представленной на рисунке, это следствие того, что у этих двух объектов различимый атрибут только форма.

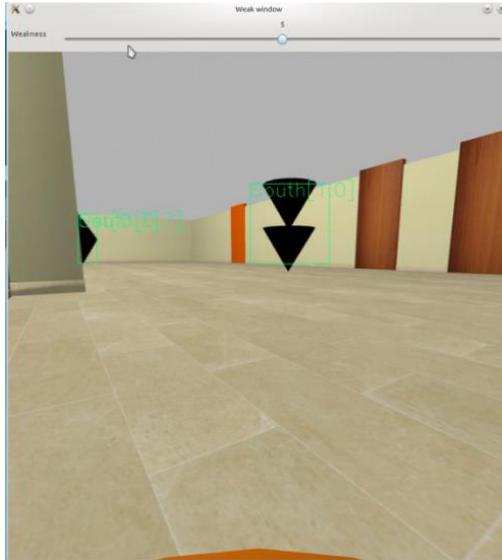


Рис. 4. Частичная видимость вехи «восток»

Аналогичная ситуация представлена на рисунке 5. Двухкомпонентный маркер состоит из двух простых объектов, соединенных пространственным отношением «над». Каждый составляющий объект является прямоугольником определённого цвета. На рисунке слева в окне представлены найденные методом простые объекты, справа – комплексные объекты. В этом случае, комплексный объект распознается с коэффициентом уверенности 0.5, согласно формуле 3.

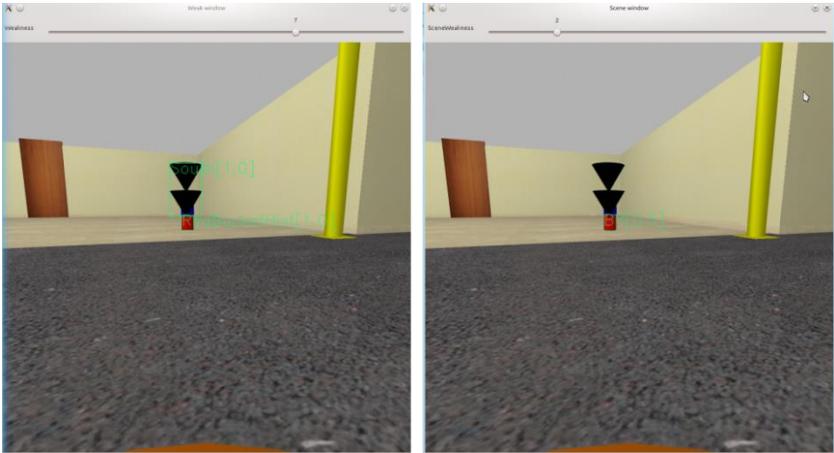


Рис. 5. Перекрывание комплексного объекта

Данные эксперименты показали потенциальную применимость метода для задач распознавания, однако робастность распознавания сильно зависит от того, какими атрибутами описывается объект и какие методы распознавания заложены в конкретной реализации. Также сильное влияние имеет выбор порогового значения для коэффициента доверия. На данный момент настройка данных параметров может осуществляться только вручную.

Дальнейшие эксперименты будут направлены на работу с более сложными объектами и на расширение базы распознавательных методов, заложенных в систему. Также планируется интеграция системы с другими системами мобильного робота, в частности с системой навигации, для получения «контекста» и использование его в качестве дополнительной информации при расчете коэффициентов уверенности найденных объектов.

Заключение

Был разработан и протестирован на лабораторных объектах комплексный метод распознавания объектов. Объект в рамках метода представляется в виде набора атрибутов, распознаваемых отдельными методами. По количеству и композиции распознанных атрибутов делаются предположения о возможном нахождении объекта, даже в условиях частичной видимости. Приведенные эксперименты показали возможность применимости метода в задачах робототехники, в т. ч. в задаче навигации [Московский, 2015].

Список литературы

- [Narin'yani и др., 1997] A.S.Narin'yani, S.B.Borde, D.A.Ivanov. Sub-Definite Mathematics and Novel Scheduling Technology Programs // Artif. Intell. Eng. 1997. Т. 11.
- [Bay и др., 2006] Bay H., Tuytelaars T., Gool L. Van. SURF: Speeded up robust features // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). , 2006. С. 404–417.
- [Biederman, 1985] Biederman I. Human image understanding: Recent research and a theory // Comput. Vis. Graph. Image Process. 1985. Т. 32. С. 29–73.
- [Canny, 1986] Canny J. A Computational Approach to Edge Detection // IEEE Trans. Pattern Anal. Mach. Intell. 1986. Т. PAMI-8. № 6. С. 679–698.
- [Davison и др., 2007] Davison A.J. и др. MonoSLAM: Real-Time Single Camera SLAM // IEEE Trans. Pattern Anal. Mach. Intell. 2007. Т. 29. № 6. С. 1–16.
- [Dickinson и др., 2010] Dickinson S.J. и др. Object Categorization, Computer and Human Vision Perspectives // J. Electron. Imaging. 2010. Т. 19. № 3. С. 039901.
- [Engel и др., 2014] Engel J., Sch T., Cremers D. LSD-SLAM: Large-Scale Direct Monocular SLAM // Eccv. 2014. С. 1–16.
- [Fallis, 2013] Fallis A.. Springer Handbook of Robotics. , 2013. 1689-1699 с.
- [Horn и др., 1988] Horn B.K.P., Hilden H.M., Negahdaripour S. Closed-form solution of absolute orientation using orthonormal matrices // J. Opt. Soc. Am. A. 1988. Т. 5. № 7. С. 1127.
- [Huang и др., 1989] Huang T.S., Faugeras O.D. Some Properties of the E Matrix in Two-View Motion Estimation // IEEE Trans. Pattern Anal. Mach. Intell. 1989. Т. 11. № 12. С. 1310–1312.
- [Lowe, 1999] Lowe D.G. Object recognition from local scale-invariant features // Proceedings of the Seventh IEEE International Conference on Computer Vision. , 1999. С. 1150–1157.
- [Lowe, 2004] Lowe D.G. Distinctive image features from scale-invariant keypoints // Int. J. Comput. Vis. 2004. Т. 60. № 2. С. 91–110.
- [Minsky, 1975] Marvin Minsky. A Framework for Representing Knowledge // Psychol. Comput. Vis. 1975.
- [Matas и др., 2004] Matas J. и др. Robust wide-baseline stereo from maximally stable extremal regions // Image and Vision Computing. , 2004. С. 761–767.
- [Nistér, 2004] Nistér D. An efficient solution to the five-point relative pose problem // IEEE Trans. Pattern Anal. Mach. Intell. 2004. Т. 26. № 6. С. 756–770.
- [Turk и др., 1991] Turk M., Pentland A. Eigenfaces for Recognition // J. Cogn. Neurosci. 1991. Т. 3. № 1. С. 71–86.
- [Viola и др., 2001] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // Comput. Vis. Pattern Recognit. 2001. Т. 1. С. I-511–I-518.
- [Zisserman и др., 1995] Zisserman а. и др. Class-based grouping in perspective images // Proc. IEEE Int. Conf. Comput. Vis. 1995. С. 183–188.
- [Варосян и др., 1982] Варосян С.О., Пospelов Д.А. Немеетрическая пространственная логика // Техническая кибернетика. 1982. Т. №5. С. 86–99.
- [Московский, 2015] Московский А.Д. Система навигации автономного мобильного робота на основе метода реконструкции сцен // Робототехника и техническая кибернетика. 2015. Т. 4. № 9. С. с.47–55.