

ПРИМЕНЕНИЕ МЕТОДА ХАФА ДЛЯ ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ В ТРЕХМЕРНОЙ СЦЕНЕ

П.Ю. Якимов, С.А. Разлацкий

Самарский государственный аэрокосмический университет имени академика С.П. Королева (национальный исследовательский университет)

В последние несколько лет оборудование, обеспечивающее съемку трехмерных сцен, получило широкое распространение в областях робототехники и компьютерного зрения за счет уменьшения стоимости. Благодаря этому появилась возможность получать дополнительную информацию об окружающей трехмерной сцене в реальном времени. Вместе с появлением третьей координаты появилась и потребность в новых алгоритмах детектирования и распознавания объектов. В данной статье рассмотрен эффективный метод обнаружения заранее известных объектов в трехмерном облаке точек. Реализованный метод Хафа для трехмерных данных протестирован на данных, полученных от сенсора Microsoft Kinect. В статье также показаны результаты экспериментов.

1 Введение

Способов получения трехмерных данных существует большое множество. Однако, в целом их можно разделить на два вида: пассивные и активные (излучающие) сенсоры. Например, стереокамеры используют ту же идею, что и человеческое зрение, и, совмещая изображения с двух или более камер, получают так называемую карту глубин [1]. При этом стереокамеры используют только свет, излучаемый другими объектами. Иной принцип работы используют TOF-камеры (*time of flight*) [2]. Они получают изображение на основе расчета времени пролета лазерного луча (инфракрасный диапазон) между сенсором и объектом для каждой точки, то есть работают по принципу эхолотатора.

Существует множество методов и алгоритмов обработки и распознавания объектов в двумерных изображениях. К сожалению, далеко не все из них применимы для обработки и интерпретации трехмерных данных. Однако, очевидно, что потенциал развития трехмерных методов и алгоритмов детектирования и распознавания очень велик.

Востребованность таких методов в прикладных сферах сложно переоценить. Алгоритмы работы с трехмерными данными могут со временем полностью заменить методы обработки двумерных данных в областях, где требуется детектирование и распознавание трехмерных объектов. Появление третьего измерения может существенно повысить надежность детектирования и распознавания.

2 Постановка задачи и исходные данные

В качестве исходных данных используется трехмерное облако точек, полученное при помощи камеры Microsoft Kinect. Облако точек – это набор вершин в системе координат. Вершины определяются координатами X, Y, Z и обычно предназначаются для представления внешней поверхности объекта. Microsoft Kinect – это бесконтактный сенсорный многофункциональный контроллер, первоначально представленный для консоли Xbox 360, а впоследствии и для персональных компьютеров под управлением операционной системы Windows [3]. Состоит из двух сенсоров глубины, цветной видеокамеры и микрофонной решетки. Проприетарное программное обеспечение осуществляет полное 3-х мерное распознавание движений тела, мимики лица и голоса. Датчик глубины состоит из инфракрасного проектора, объединенного с монохромной КМОП-матрицей, что позволяет датчику Kinect получать трёхмерное изображение при любом естественном освещении. Для получения облака точек подходит также и стереокамера, однако для разработки алгоритмов детектирования объектов в трехмерной сцене было принято решение использовать данные Microsoft Kinect ввиду большей точности.

Задача заключается в том, чтобы обнаружить заранее известный объект по получаемым трехмерным данным. На рисунке 1 приведены примеры данных, полученных при помощи Kinect.

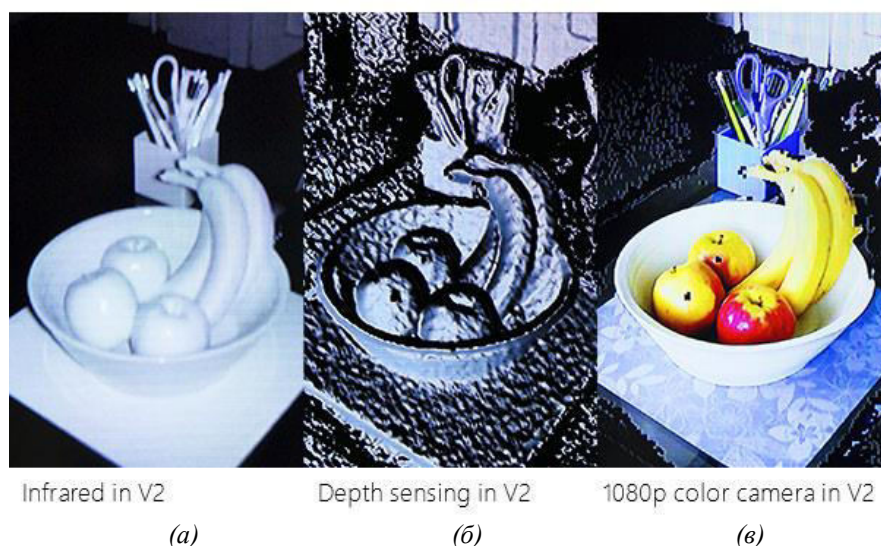


Рисунок 1 – а) инфракрасное изображение; б) карта глубин; в) FullHD камера

3 Применение метода Хафа для детектирования объектов в трехмерной сцене

За основу метода взят алгоритм преобразования Хафа (*Hough Transform*) [4]. Трехмерная версия преобразования Хафа основана на тех же принципах, что и двумерная. Одним из отличий является извлечение особых точек для уменьшения вычислительной нагрузки. Трехмерная сцена, как правило, содержит сотни тысяч точек, модель искомого объекта также может состоять из десятков тысяч точек. Использование всех точек, как в двумерном случае, на порядок повышает вычислительную нагрузку.

Для извлечения особых точек используется алгоритм кластеризации [5]. То есть множество точек, находящихся друга от друга на малом расстоянии, объединяется в одну область, которая имеет объем и рассматривается при этом как одна точка в новом облаке точек. Для получаемых сцен необходимо каждый раз извлекать особые точки. Для модели искомого объекта достаточно провести кластеризацию один раз и использовать ее в дальнейшем. Извлеченные особые точки затем пополняются информацией о точках вокруг. Имея наборы «расширенных» особых точек для сцены и модели, есть возможность построить множество соответствий особых точек.

Так же, как и в двумерном случае, использования метода голосования Хафа включает в себя так называемое аккумуляторное пространство для сбора информации о возможном присутствии в данной области объекта. Для обнаружения искомого объекта модель-эталон «прикладывается» к особым точкам в трехмерной сцене. В качестве опорной точки может быть выбрана любая точка, принадлежащая модели. В нашем случае для удобства мы выбрали центральную точку C^M .

При каждом совпадении точек модели и сцены в трехмерном аккумуляторном пространстве прибавляются так называемые голоса в той области, где произошло совпадение точек.

Одна из основных сложностей предлагаемого алгоритма – это вероятность различной ориентации сцены и искомого объекта. При этом в отличие от двумерного случая добавляется еще одна степень свободы, увеличивая на порядок количество вариантов ориентации искомого объекта в сцене. Эта проблема решается при помощи введения специальных векторов, которые описаны ниже.

Предположим все точки координат трехмерной модели и сцены представлены в одной системе координат, для каждой особой точки модели F_i^M вычисляется вектор между C^M (точка центра объекта) и F_i^M :

$$V_{i,G}^M = C^M - F_i^M \quad (1)$$

Затем, для обеспечения инвариантности к вращению и повороту, каждый вектор $V_{i,G}^M$ преобразовывается в координаты соответствующей локальной относительной системы

координат (в англоязычной литературе – local reference frame или local RF) (т.е. вычисляемой в соответствии с F_i^M) с помощью следующего преобразования:

$$V_{i,L}^M = R_{G,L}^M * V_{i,G}^M, \quad (2)$$

где «*» представляет собой произведение матриц, а $R_{G,L}^M$ является матрицей вращения, где каждая строка – единичный вектор локальной системы координат для особой точки F_i^M :

$$R_{G,L}^M = [L_{i,x}^M, L_{i,y}^M, L_{i,z}^M]^T \quad (3)$$

На рисунке 2 представлена схема такого преобразования.

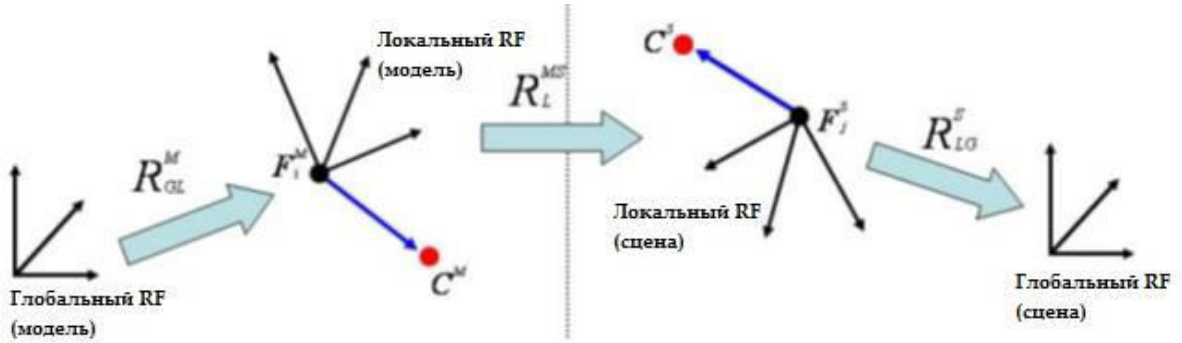


Рисунок 2 – Схема преобразования вектора соответствия точек в модели

Вычисления (1) – (3) производятся для модели один раз. Эта стадия заканчивается привязкой каждой особой точки F_i^M по своему вектору $V_{i,L}^M$.

С момента, когда соответствия между трехмерной моделью и сценой получены, как описано выше, каждая из особых точек сцены F_j^S , для которых было найдено соответствие ($F_j^S \leftrightarrow F_i^M$) голосует на своей позиции в сцене. Поскольку вычисление локальных RF для F_j^S инвариантно к поворотам и вращениям, это позволяет определить преобразование, как $R^M S_L$. Наконец, необходимо преобразовать $V_{i,L}^M$ в глобальную систему координат посредством следующего преобразования:

$$V_{i,G}^S = R_{L,G}^S * V_{i,L}^S + F_j^S, \quad (4)$$

где $R_{L,G}^S$ является матрицей вращения, полученной при помощи выстроенных в столбцы единичных векторов локальной относительной системы координат для особой точки F_j^S :

$$R_{L,G}^S = [L_{i,x}^S, L_{i,y}^S, L_{i,z}^S]. \quad (5)$$

Благодаря этим преобразованиям функция F_j^S может проголосовать в аккумуляторном пространстве Хафа с помощью вектора $V_{i,G}^S$. Признаком наличия определенного объекта являются экстремумы в аккумуляторном пространстве Хафа. Среди получившихся множественных пиков в аккумуляторном пространстве можно выделить несколько кандидатов на роль потенциального искомого объекта. Если мы уверены, что на сцене имеется лишь один экземпляр искомого объекта, мы выбираем объект на сцене набравший максимальное количество голосов. Кроме того, для придания пикам устойчивости к эффекту квантования, также задействуем соседние точки, которые уступают локальному максимуму. В частности, мы предлагаем порог пространства Хафа по добавлению к каждой точке значения шести ее соседних точек, так как из-за наличия шума правильным соответствием может быть соседняя точка. Затем для последующих стадий выбираются подмножества соответствий при том, только такие что касаются центральной точки над пороговым максимумом для повышения точности.

На рисунке 3 изображена схема метода Хафа для детектирования объектов в трехмерной сцене в виде облака точек.

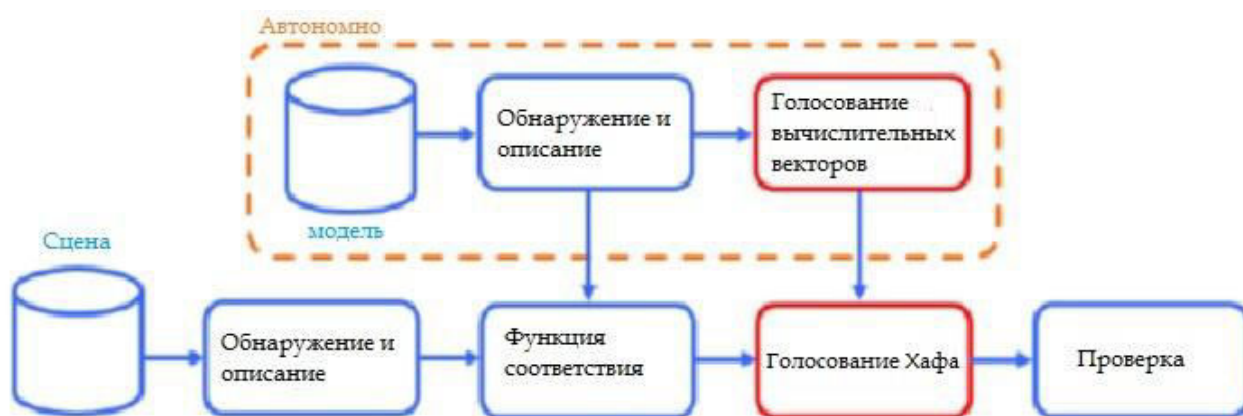


Рисунок 6.3 – Схема метода Хафа для детектирования объектов в трехмерной сцене

4 Примеры работы метода детектирования объектов в трехмерной сцене

Результаты применения алгоритма голосования Хафа представлены на рисунке 4. Как видно на рисунке, алгоритм успешно решил поставленную задачу и точно указал на область присутствия искомого объекта.

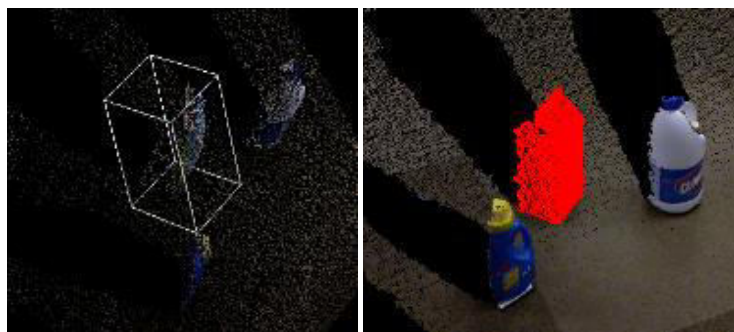


Рисунок 4 – Результаты распознавания алгоритмом Хафа

Слева на рисунке 5 изображена модель человека в виде облака точек. Справа представлена трехмерная сцена, на которой описанным в настоящем разделе методом был произведен поиск объекта.

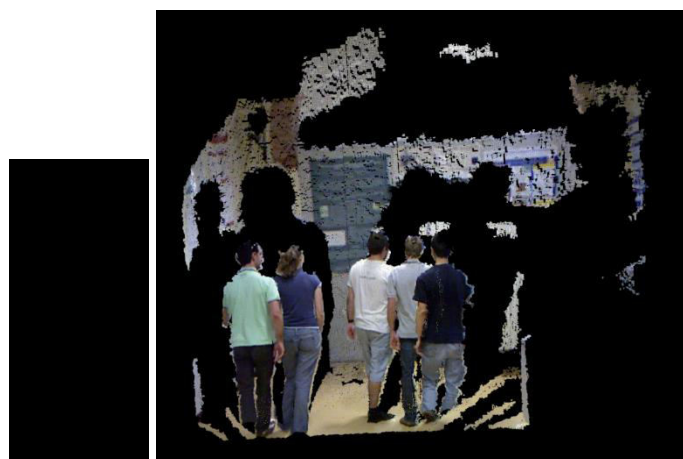


Рисунок 5 – Модель-эталон и исходная трехмерная сцена

Рисунок 6 показывает вид на трехмерную сцену с двух ракурсов с обнаруженным и отмеченным искомым объектом.



Рисунок 6 – Обнаруженный объект на трехмерной сцене

Алгоритм требует первоначального обучения, которое может занять довольно долгое время, и не может быть выполнено в реальном времени. Однако, такое обучение необходимо лишь при добавлении новых моделей-эталонов. Поиск предварительно обработанных объектов на поступающих трехмерных сценах выполняется в реальном времени.

5 Заключение

В настоящей статье описан метод детектирования различных объектов с использованием трехмерной модели окружающей дорожной обстановки. В качестве входных данных разработанный метод принимает эталонные объекты (пешеходы, животные, прочие препятствия) и поступающие трехмерные сцены в виде облаков точек. На выходе метод получает трехмерные координаты искомого объекта относительно снимающей трехмерной камеры.

Метод Хафа для детектирования трехмерных объектов в облаке точек показал эффективные результаты. В дальнейшем полученные результаты исследования предполагается использовать в задаче детектирования и распознавания пешеходов и прочих препятствиях при движении автомобиля. Также планируется модифицировать метод с целью уменьшить зависимость качества детектирования от точного совпадения трехмерной модели и искомого объекта в облаке точек. Для ускорения обработки данных будет использована технология программирования графических процессоров CUDA.

Литература

- Fursov, V.A.**, Information technology for digital terrain model reconstruction from stereo images / V.A. Fursov, Y.V. Goshin // Computer Optics. – Volume 38, Issue 2. – 2014. – P. 335-342
- Ringbeck, T.** A 3d time of flight camera for object detection / T. Ringbeck, B. Hagebecker // Range Imaging Day, Zürich. – 2007.
- Khoshelham K.** Accuracy and resolution of kinect depth data for indoor mapping applications / Khoshelham K., Elberink S. O. //Sensors. – 2012. – Т. 12. – №. 2. – С. 1437-1454.
- Duda, R.** Use of the Hough Transformation to Detect Lines and Curves in Pictures / Duda, R. O. and P. E. Hart // Comm. ACM. – 1972. – Vol. 15, P. 11–15.
- Tombari F.**, Object recognition in 3D scenes with occlusions and clutter by Hough voting / L. Di Stefano, F. Tombari// Fourth Pacific-Rim Symposium on Image and Video Technology. – 2010. – P. 2-4