МИНОБРНАУКИ РОССИИ

Федеральное государственное автономное образовательное

учреждение высшего профессионального образования

«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

**Инженерно-технологическая академия**

**Институт компьютерных технологий и информационной безопасности**

**Кафедра Математического обеспечения и применения ЭВМ**

**Реферат**

на тему «Обработка стереоизображений и облаков точек»

по курсу «Компьютерное зрение»

Выполнили:

ст. группы КТмо1-3

Шепель И.О.

Проверил:

доцент кафедры МОП ЭВМ

Селянкин В.В.

Оценка

«\_\_\_\_» \_\_\_\_\_\_\_\_\_\_ 2017 г.

Таганрог 2017

**СОДЕРЖАНИЕ**

ВВЕДЕНИЕ 3

1. СТЕРЕОСОПОСТАВЛЕНИЕ 4

#### 1.1. Модель проективной камеры 4

#### 1.2. Калибровка пары камер 7

#### 1.3. Проекция трехмерной точки на изображениях стереопары 9

#### 1.4. Триангуляция точек 12

#### 1.5. Построение карт глубины 13

#### 1.6. Альтернативные способы получения облаков точек 16

2. ОБРАБОТКА ОБЛАКОВ ТОЧЕК В ЗАДАЧАХ СИНТЕЗА МОДЕЛИ ПРОХОДИМОСТИ ОКРУЖАЮЩЕЙ СРЕДЫ 19

ЗАКЛЮЧЕНИЕ 24

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ 25

**ВВЕДЕНИЕ**

В настоящее время большой сектор задач компьютерного зрения занимают задачи получения, обработки и анализа трехмерной информации об окружающем мире. В частности это задачи, связанные с получением стереоизображений, и из них облаков точек – массивов трехкоординатных измерений. Современная техника позволяет получать облака точек и напрямую, минуя этап обработки двух плоских изображений для получения дальнометрической информации.

С помощью измерения расстояний до пикселей на изображении можно получить не только цветовую, но и геометрическую информацию. Это позволяет решать множество задач, которые невозможно было решить посредством обработки цветных изображений. Более того, такие методы лишены многих недостатков цветных изображений – меньше зависят от условий освещенности, работают вне зависимости от времени года и погодных условий. Таким образом стало возможно распознавать объекты, в том числе и замаскированные, определять препятствия любых видов, в том числе деревья, стены, отрицательные препятствия (ямы и рвы), лужи, динамические препятствия (пешеходов, велосипедистов и автомобили), что сделало возможным широкое и активное производство автономных мобильных систем.

В этой работе пойдет речь о методах стереосопоставления и получения трехмерных данных, а также о методах обработки облаков трехмерных точек для решения частной задачи области робототехники и автономной навигации – построения модели проходимости.

**1. СТЕРЕОСОПОСТАВЛЕНИЕ**

**1.1. Модель проективной камеры**

Прежде чем углубляться в методы стереосопоставления (англ. – stereo matching), необходимо рассмотреть модель проективной камеры (англ. – projective camera, pinhole camera) [1]. Эта модель представлена на рисунке 1.

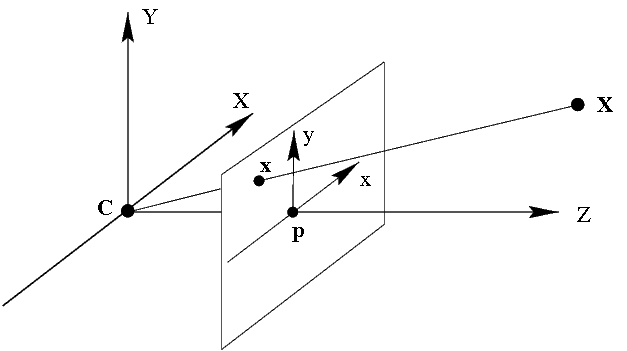


Рис. 1. Модель проективной камеры.

Проективная камера определяется центром камеры, главной осью — лучом, начинающимся в центре камеры и направленным туда, куда камера смотрит (ось *CZ* на рисунке), плоскостью изображения — плоскостью на которую выполняется проецирование точек, и системой координат на этой плоскости. В такой модели, произвольная точка пространства *X* проецируется на плоскость изображения в точку *x* лежащую на отрезке *CX*, который соединяет центр камеры *C* с исходной точкой *X*.

Формула проецирования имеет простую математическую запись в однородных координатах:

где *X* – однородные координаты точки пространства, *x –* однородные координаты точки плоскости, *P* – матрица камеры размера 3 × 4 [2].

Матрица *P* выражается следующим образом:

где *K* — верхняя треугольная матрица внутренних параметров камеры размера 3 × 3 (конкретный вид приведен ниже), *R* — ортогональная матрица размера 3 × 3, определяющая поворот камеры относительно глобальной системы координат, *I* — единичная матрица размера 3 × 3, вектор *c* — координаты центра камеры, а

Стоит отметить, что матрица камеры определена с точностью до постоянного ненулевого множителя, который не изменит результатов проецирования точек. Однако этот постоянный множитель обычно выбирается так, что бы матрица камеры имела нижеописанный вид.

В самом простейшем случае, когда центр камеры лежит в начале координат, главная ось камеры сонаправлена оси *CZ*, оси координат на плоскости камеры имеют одинаковый масштаб (что эквивалентно квадратным пикселям), а центр изображения имеет нулевые координаты, матрица камеры будет равна , где

Эта модель является теоретическим упрощением. У реальных камер пикселы обычно незначительно отличаются от квадратных, а центр изображения имеет ненулевые координаты. В таком случае матрица внутренних параметров примет вид:

Коэффициенты *f*, , — называются фокусными расстояниями камеры (соответственно общим коэффициентом и коэффициентами вдоль осей x и y).

Помимо этого, в силу неидеальности оптики, на изображениях, полученных с камер, присутствуют искажения-дисторсии (англ. – distortion). Данные искажения имеют нелинейную математическую запись:

где *k1, k2, p1, p2, k3* — коэффициенты дисторсии, являющиеся параметрами оптической системы;

,

*(x', y')* — координаты проекции точки относительно центра изображения при квадратных пикселях и отсутствии искажений; *(x″, y″)* — искаженные координаты точки относительно центра изображения при квадратных пикселях.

Дисторсии не зависят от расстояния до объекта, а зависят только от координат точек, в которые проецируются пиксели объекта. Соответственно для компенсации дисторсий обычно выполняется преобразование исходного изображения полученного с камеры. Это преобразование будет одним и тем же для всех изображений, полученных с камеры, при условии постоянства фокусного расстояния (математически — одной и той же матрицы внутренних параметров).

В ситуации, когда известны внутренние параметры камеры и коэффициенты дисторсии говорят, что камера откалибрована.

**1.2. Калибровка пары камер**

Данные, которые может передать одна камера – двухмерные, плоские. Для того чтобы получить дальнометрические координаты точек, другими словами трехмерное изображение, необходима система камер (минимум 2). Для того чтобы иметь возможность находить расстояние прежде необходимо узнать калибровку этой системы камер.

Пусть имеются две камеры, заданные своими матрицами *P* и *P'* в некоторой системе координат. В таком случае говорят, что имеется пара откалиброванных камер. Если центры камер не совпадают, то эту пару камер можно использовать для определения трехмерных координат наблюдаемых точек.

Зачастую, система координат выбирается так, что матрицы камер имеют вид

Это всегда можно сделать, если выбрать начало координат совпадающее с центром первой камеры, и направить ось *Z* вдоль ее оптической оси.

Калибровка камер обычно выполняется за счет многократной съемки некоторого калибровочного шаблона [3], на изображении которого можно легко выделить ключевые точки, с известными относительными положениями в пространстве. Далее составляются и решаются (приближенно) системы уравнений, связывающие координаты проекций, матрицы камер и положения точек шаблона в пространстве [4].

Калибровочные шаблоны (англ. – calibration pattern) должны иметь четкую структуру, для того чтобы по ним проще было найти характерные точки. Обычно при этом они выполнены в двух цветах с наибольшей разницей в контрастности – черном и белом. Примеры таких шаблонов приведены на рисунке 2.

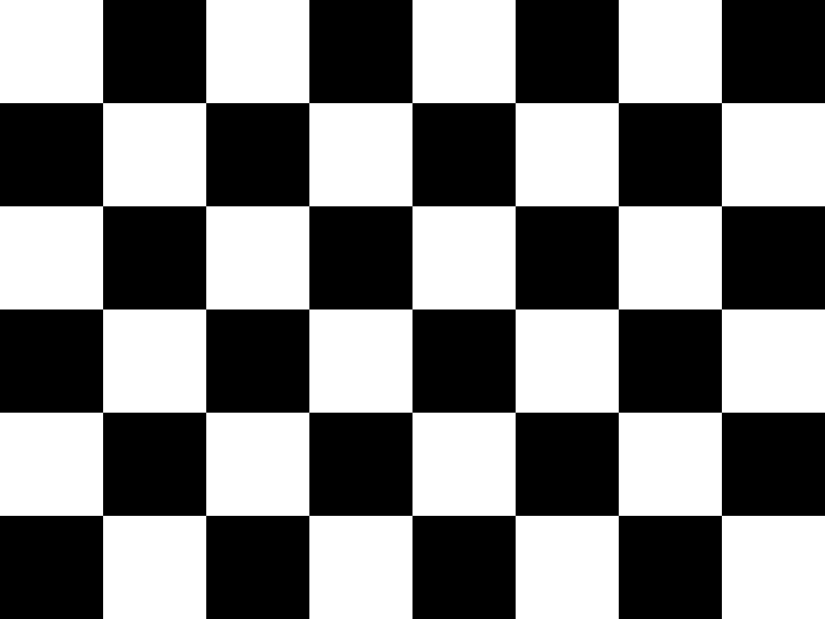
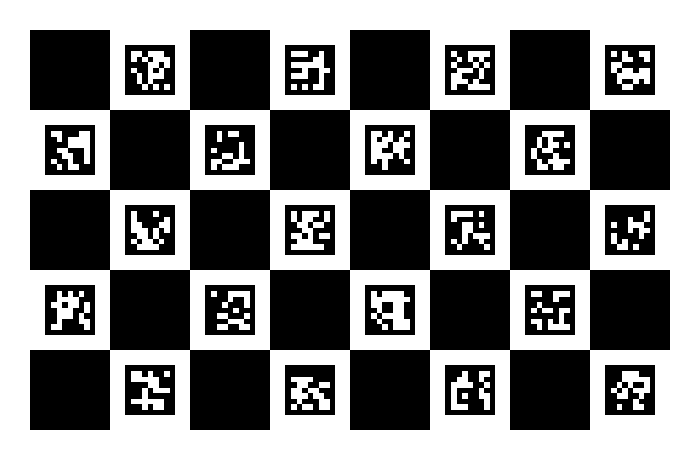
 

Рис. 2. Калибровочные шаблоны.

Слева – наиболее часто используемый паттерн – шахматная доска. Ключевые точки на ней – пересечения клеток. Алгоритм получения калибровки по такому паттерну очень прост, однако необходимо чтобы весь паттерн всегда попадал в область видимости камеры. Этого недостатка лишен второй паттерн – в каждую белую клетку встроен уникальный QR-код, который позволяет однозначно и очень точно установить связь между пикселем изображения и точкой на шаблоне, даже если его часть не видна.

**1.3. Проекция трехмерной точки на изображениях стереопары**

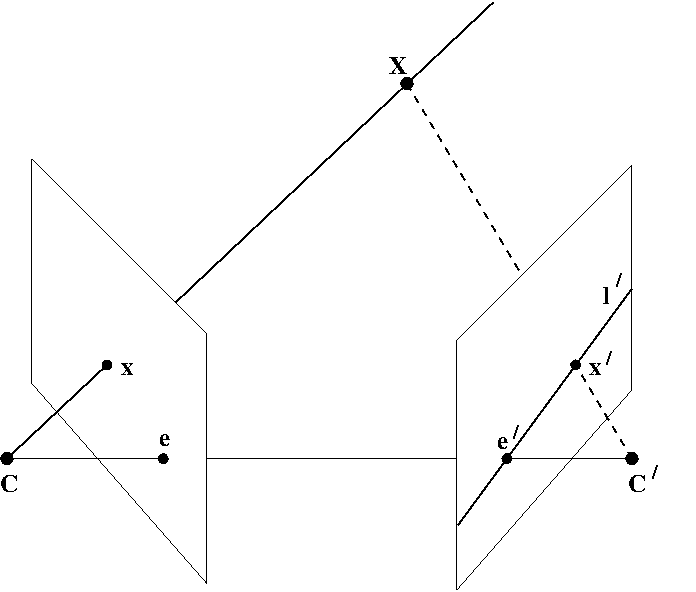


Рис. 3. Поиск стереопары.

Пусть имеются две камеры, как изображено на рисунке 3. *C* — центр первой камеры, *C'* — центр второй камеры. Точка пространства *X* проецируется в *x* на плоскость изображения левой камеры и в *x'* на плоскость изображения правой камеры. Прообразом точки *x* на изображении левой камеры является луч *xX*. Этот луч проецируется на плоскость второй камеры в прямую *l'*, называемую эпиполярной линией. Образ точки *X* на плоскости изображения второй камеры обязательно лежит на эпиполярной линии *l'*.

Таким образом, каждой точке x на изображении левой камеры соответствует эпиполярная линия *l'* на изображении правой камеры. При этом пара для *x* на изображении правой камеры может лежать только на соответствующей эпиполярной линии. Аналогично, каждой точке *x'* на правом изображении соответствует эпиполярная линия *l* на левом.

Эпиполярную геометрию используют для поиска стереопар, и для проверки того, что пара точек может быть стереопарой (т.е. проекцией некоторой точки пространства).

Эпиполярная геометрия имеет очень простую запись в координатах. Пусть имеется пара откалиброванных камер, и пусть *x* — однородные координаты точки на изображении одной камеры, а *x'* — на изображении второй. Существует такая матрица *F* размера 3 × 3, что пара точек *x*, *x'* является стереопарой тогда и только тогда, когда:

Матрица *F* называется фундаментальной матрицей (англ. – fundamental matrix). Ее ранг равен 2, она определена с точностью до ненулевого множителя и зависит только от матриц исходных камер *P* и *P'*.

В случае, когда матрицы камер имеют вид фундаментальная матрица может быть вычислена по формуле:

где для вектора *e* *[e]x* вычисляется как:

С помощью фундаментальной матрицы вычисляются уравнения эпиполярных линий. Для точки *x*, вектор, задающий эпиполярную линию, будет иметь вид , а уравнение самой эпиполярной линии: . Аналогично для точки *x'*, вектор, задающий эпиполярную линию, будет иметь вид .

Помимо фундаментальной матрицы, существует еще такое понятие, как существенная матрица (англ. – essential matrix):

В случае, когда матрицы внутренних параметров будут единичными существенная матрица будет совпадать с фундаментальной. По существенной матрице можно восстановить положение и поворот второй камеры относительно первой, поэтому она используется в задачах, в которых нужно определить движение камеры.

**1.4. Триангуляция точек**

Перейдем к тому, как определить трехмерные координаты точки по координатам ее проекций. Этот процесс в литературе называется триангуляцией (англ. – triangulation).

Пусть имеются две откалиброванные камеры с матрицами *P1* и *P2*. *x1* и *x2* — однородные координаты проекций некоторой точки пространства *X*. Тогда можно составить следующую систему уравнений:

.

На практике для решения этой системы применяется следующий подход. Векторно умножают первое уравнение на *x1*, второе на *x2*, избавляются от линейно зависимых уравнений и приводят систему к виду , где A имеет размер 4 × 4. Далее, можно либо исходить из того, что вектор *X* является однородными координатами точки, положить его последнюю компоненту равной 1 и решать полученную систему из 3-х уравнений с тремя неизвестными. Альтернативный способ — взять любое ненулевое решение системы , например вычисленное, как сингулярный вектор, отвечающий наименьшему сингулярному числу матрицы *A*.

**1.5. Построение карт глубины**

Карта глубины (англ. – depth map) — это изображение, на котором для каждого пикселя, вместо цвета, храниться его расстояние до камеры. Карта глубины может быть получена с помощью специальной камеры глубины (например, сенсор Kinect является своего рода такой камерой), а так же может быть построена по стереопаре изображений.

Идея, лежащая в основе построения карты глубины по стереопаре очень проста. Для каждой точки на одном изображении выполняется поиск парной ей точки на другом изображении. А по паре соответствующих точек можно выполнить триангуляцию и определить координаты их прообраза в трехмерном пространстве. Зная трехмерные координаты прообраза, глубина вычисляется, как расстояние до плоскости камеры.

Парную точку нужно искать на эпиполярной линии. Соответственно, для упрощения поиска, изображения выравнивают так, что бы все эпиполярные линии были параллельны сторонам изображения (обычно горизонтальны). Более того, изображения выравнивают так, что бы для точки с координатами (*x0, y0*) соответствующая ей эпиполярная линия задавалась уравнением , тогда для каждой точки соответствующую ей парную точку нужно искать в той же строчке на изображении со второй камеры. Такой процесс выравнивания изображений называют ректификацией (англ. – rectification). Обычно ректификацию совершают путем ремэппинга изображения и ее совмещают с избавлением от дисторсий.

Пример изображения и соответствующей ему карты глубины приведен на рисунке 4, картинки взяты из базы [5].

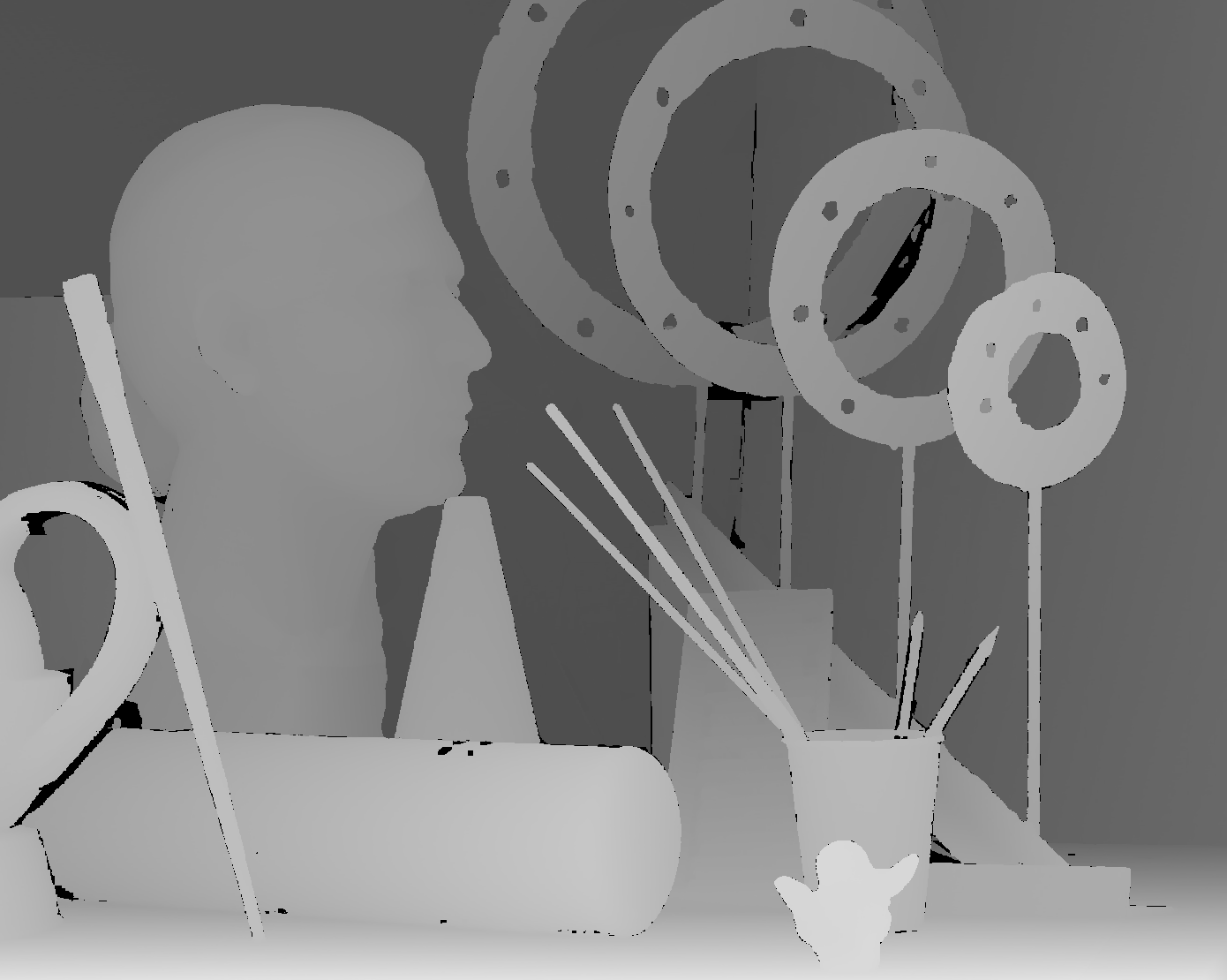
 

Рис. 4. Исходное изображение левой камеры (слева) и карта глубины (справа).

После того как изображения ректифицированы, выполняют поиск соответствующих пар точек. Самый простой способ проиллюстрирован на рисунке 5 и состоит в следующем. Для каждого пикселя левой картинки с координатами (*x0, y0*) выполняется поиск пикселя на правой картинке. При этом предполагается, что пиксель на правой картинке должен иметь координаты (*x0* — *d*, *y0*), где *d* — величина называемая несоответствие/смещение (англ. – disparity). Поиск соответствующего пикселя выполняется путем вычисления максимума функции отклика, в качестве которой может выступать, например, корреляция окрестностей пикселей [6].

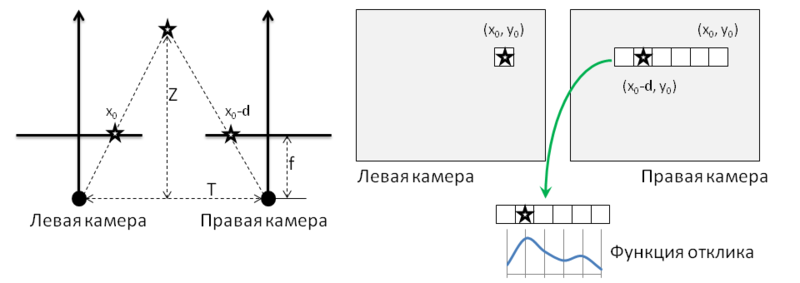


Рис. 6. Вычисление карты глубины.

Собственно значения глубины обратно пропорциональны величине смещения пикселей. Если использовать обозначения с левой половины рисунка 6, то зависимость между disparity и глубиной можно выразить следующим способом:

Из-за обратной зависимости глубины и смещения, разрешающая способность систем стерео зрения, которые работают на основе данного метода, лучше на близких расстояниях, и хуже на далеких.

**1.6. Альтернативные способы получения облаков точек**

С развитием современной техники появилось множество новых способов снимать трехмерные изображения. Например, существует целый ряд ToF-датчиков (англ. – time of flight) или времяпролетных датчиков. Они измеряют время между выпуском луча и приходом отраженного от объекта луча.



Рис. 6. ToF-датчик Mesa Imaging SR4500.

Еще один подкласс 3D датчиков - различные лазерные сканирующие системы. Принципы их работы отличаются лишь тем, что на объект сканирования наносятся специальные светоотражающие маркеры, и такой способ сканирования хоть и очень точный, но работает только со статичными объектами.

Наиболее широко используются так называемые лидары (англ. – light detection and ranging). В них используются лазерные лучи, и для замера дальности также вычисляется время пролета и разность фаз. Отдельные лидары могут генерировать до 2.2 млн точек в секунду (например, Velodyne HDL-64E).



Рис. 7. Лидар Velodyne HDL-64E.

Такие инструменты позволяют точно измерять дальность на расстояниях до 300 метров с полным круговым обзором, это достигается за счет высокой частоты вращения головки, на которой вертикально размещены 64 лазера. Поэтому такие датчики активно используются в робототехнике.

На рисунке 8 приведен пример облака точек заснятого лидаром HDL-64E.

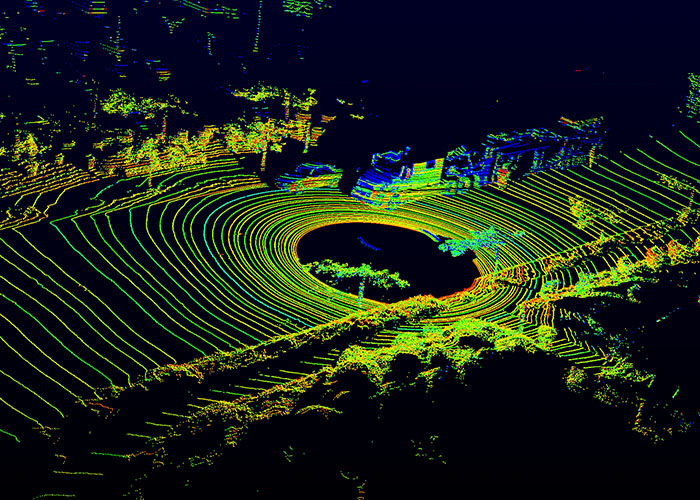


Рис. 8. Облако точек, сгенерированное лидаром Velodyne HDL-64E.

**2. ОБРАБОТКА ОБЛАКОВ ТОЧЕК В ЗАДАЧАХ СИНТЕЗА МОДЕЛИ ПРОХОДИМОСТИ ОКРУЖАЮЩЕЙ СРЕДЫ**

В современных робототехнических комплексах часто возникает задача автономной навигации робота. Для передвижения необходимо построить карту окружающей местности, которая будет содержать в себе информацию о том, может ли робот передвигаться по ней, координаты препятствий и труднопроходимых участков ландшафта. С этой информацией робот сможет строить траекторию своего передвижения так, чтобы не получить никаких повреждений и функционировать корректно.

Существует несколько подходов к построению моделей проходимости. Все они используют так называемые облака точек в качестве входного параметра. Пример облака точек приведен на рисунке 9.

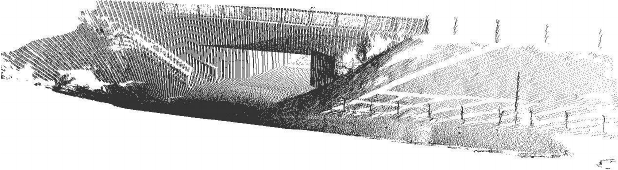


Рис. 9. Пример облака точек с различными типами проходимости – свободные для проезда участки, склоны, ступени (шаги), вертикальные препятствия и нависающие препятствия.

В большинстве современных статей на тему построения моделей проходимости выделяют следующие типы препятствий [7]:

1. Вертикальные (англ. - positive obstacles) – объекты, на которые платформа не может заехать – стены, деревья, столбы.
2. Наклонные (англ. – slope) – участки земли, которые платформа не может преодолеть из-за слишком высокого уклона.
3. Шероховатые (англ. – rough terrain) - крупногравийные дороги, жесткая трава, бугристая поверхность.
4. Шаги (англ. - step) – ступени, высота которых непреодолима для платформы.
5. Отрицательные (англ. – negative obstacles) – ямы и канавы.

Это основные типы препятствий, присущие наиболее простым моделям, но существуют и более сложные классификации.

Простейший из методов построения таких моделей – двумерная карта проходимости. Все облако точек разбивается на двумерные ячейки одинакового размера. В каждой ячейке находится центроид – среднее арифметическое координат всех точек в ячейке. В настоящее время такой подход не используется, так как он обладает очень низкой точностью, и поэтому фактически неприменим на подавляющем большинстве ландшафтов.

Наиболее точный метод синтеза информационной модели окружающей среды – построение трехмерных воксельных сеток (англ. - 3D voxel grid) [8]. Все облако точек разбивается на трехмерные ячейки. Если в ячейке присутствует необходимое количество точек, то она заменяется на куб. Основной недостаток таких сеток – они требуют большого места в памяти и больших вычислительных затрат, поэтому их применение в реальном времени на мобильных и автономных системах невозможно. На рисунке 10 изображена воксельная сетка по облаку точек кролика.

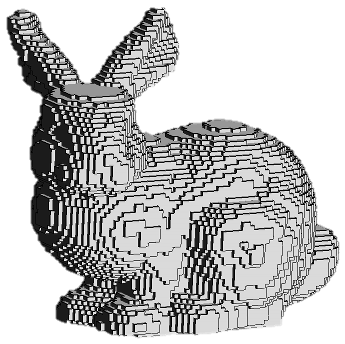


Рис. 10. Воксельная сетка по облаку точек снятому с кролика.

Наиболее широко используемый подход построения модели проходимости – карты высот (англ. - elevation map) или 2.5D карты. Облако точек также сперва разбивается на ячейки двумерной дискретной решетки. Но в отличие от двумерных карт, каждая ячейка содержит в себе много других параметров, таких как: средняя высота точек, среднеквадратичное отклонение координат точек от центроида, дисперсию распределения точек, собственные вектора и сингулярные числа матрицы с координатами точек и может быть дополнена другими геометрическими параметрами [7]. Эти параметры позволяют с высокой точностью принимать решение о проходимости ячейки, и в то же время работа с ними не требует таких больших вычислительных затрат как с трехмерными картами. Пример такой карты приведен на рисунке 11.

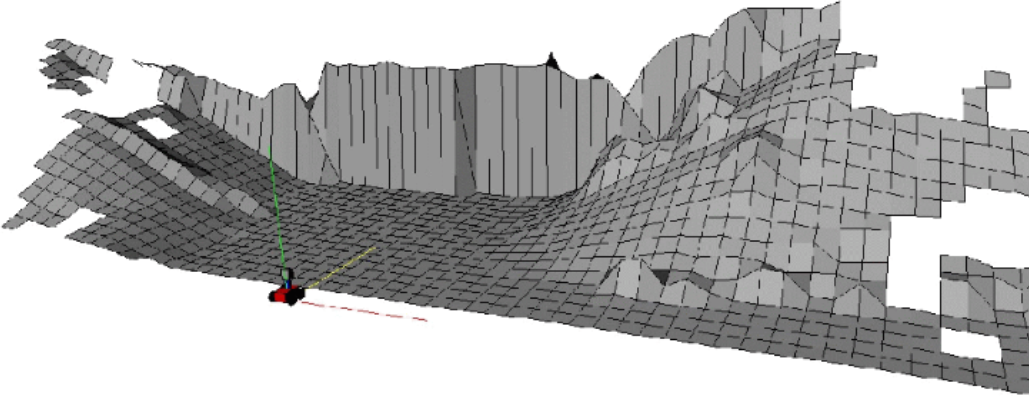


Рис. 11. Карта высот, построенная по облаку точек с рисунка 9.

Одна из проблем таких карт – нависающие препятствия не отделяются от вертикальных. На рисунке 11 видно, что проходимая поверхность под мостом определилась как вертикальное препятствие. Чтобы разделять такие виды препятствий используются расширенные карты высот (англ. - extended elevation map) [9]. В них точки сортируются по высоте, и если разница между двумя соседними по высоте точками превышает высоту робота, то все точки выше отбрасываются. Пример расширенной карты высот приведен на рисунке 12.

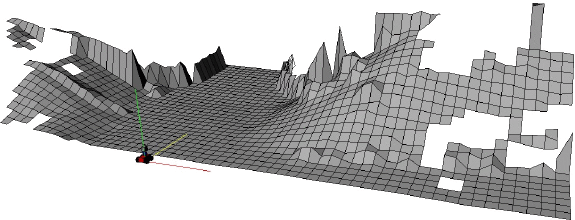


Рис. 12. Расширенная карта высот, построенная по облаку точек с рисунка 9. Поверхность под мостом корректно пометилась как проходимая.

Такой способ обработки данных работает только в том случае, если вся карта строится по одноуровневой поверхности. В городских условиях повсеместно встречаются мосты, разъезды, многоуровневые парковки. Для того чтобы учесть несколько уровней поверхности в одной ячейке используется еще одна надстройка карт высот – так называемая многоуровневая карта поверхностей (англ. - multi-level surface map) [10]. Такие карты также фильтруют точки в ячейке, но те точки, что превышают высоту робота не удаляются, а образуют новую поверхность, и так сколь угодно раз. Пример такой карты представлен на рисунке 13.

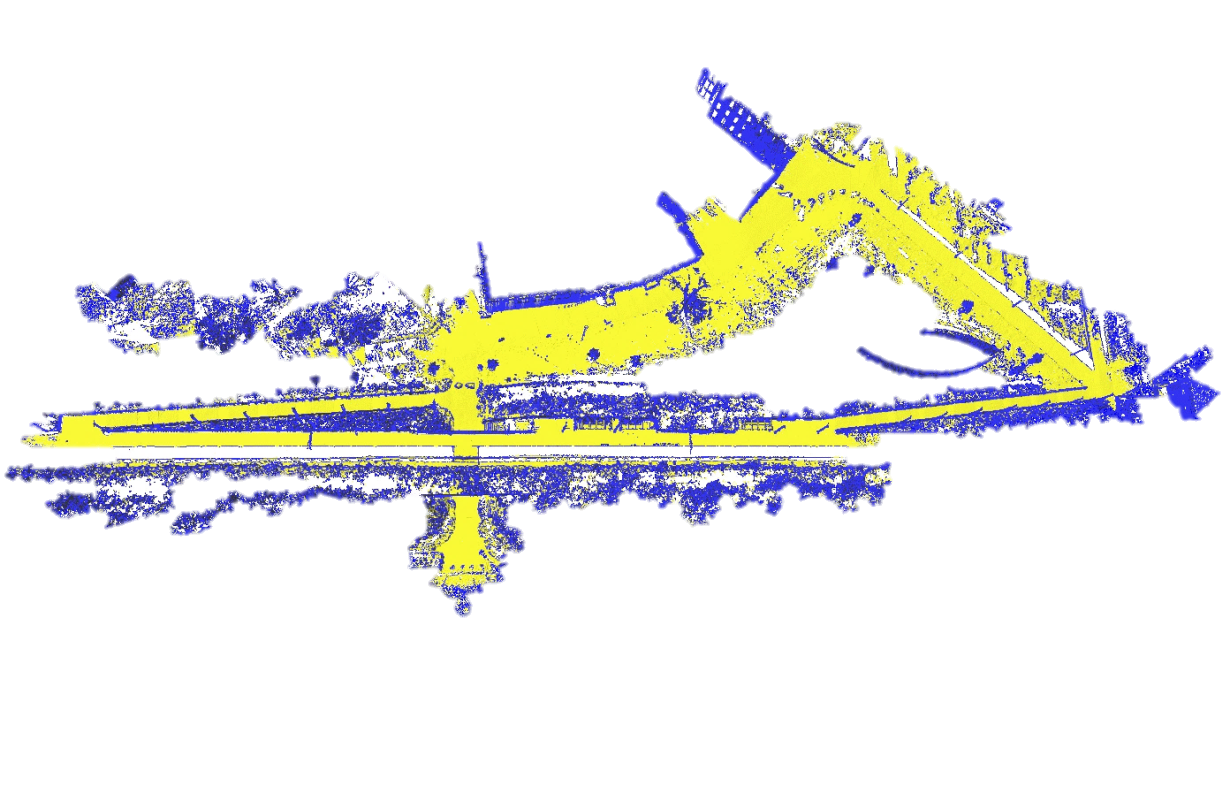


Рис. 13. Многоуровневая карта поверхностей окрестностей университета Фрейбурга. Желтым помечена проходимая поверхность. Как дорога под мостом, так и сам мост принадлежат этой поверхности.

**ЗАКЛЮЧЕНИЕ**

В современном мире для задач робототехники используются именно способы снятия трехмерных изображений, так как подобная информация менее зависит от внешних условий и может предоставить полную картину окружающей местности или объектов. Современные достижения в технике позволяют с большой частотой и точностью получать такие данные. Таким образом, именно трехмерные изображения или облака точек являются наиболее активно развивающейся областью компьютерного зрения.

**СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Kaehler A., Bradski G. Learning OpenCV. – O'Reilly Media, Inc., 2014.
2. Camera calibration and 3D reconstruction. OpenCV API Reference. [Электронный ресурс] – Режим доступа к статье: <http://docs.opencv.org/2.4.8/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html>
3. Киреев О. С. Автоматизированная калибровка стереопары в лабораторных условиях // Математические машины и системы. – 2004. – №. 1.
4. Porrill J., Pollard S. Curve matching and stereo calibration // Image and Vision Computing. – 1991. – Т. 9. – №. 1. – С. 45-50.
5. База изображений по сравнению различных методов построения карты глубины. [Электронный ресурс] - Режим доступа к статье: <http://vision.middlebury.edu/stereo/>
6. Scharstein D., Szeliski R., Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms // Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings. IEEE Workshop on. – IEEE, 2001. – С. 131-140.
7. Fleischmann P., Berns K. A Stereo Vision Based Obstacle Detection System for Agricultural Applications // Field and Service Robotics. – Springer International Publishing, 2016. – С. 217-231.
8. Karmakar N. et al. Construction of 3D orthogonal cover of a digital object // International Workshop on Combinatorial Image Analysis. – Springer Berlin Heidelberg, 2011. – С. 70-83.
9. Pfaff P., Triebel R., Burgard W. An efficient extension to elevation maps for outdoor terrain mapping and loop closing // The International Journal of Robotics Research. – 2007. – Т. 26. – №. 2. – С. 217-230.
10. Triebel R., Pfaff P., Burgard W. Multi-level surface maps for outdoor terrain mapping and loop closing // 2006 IEEE/RSJ international conference on intelligent robots and systems. – IEEE, 2006. – С. 2276-2282.