УДК 004-932-2

ПРИМЕНЕНИЕ МЕТОДОВ ЭПИПОЛЯРНОЙ ГЕОМЕТРИИ ДЛЯ ВЫЧИСЛЕНИЯ ТРЕХМЕРНЫХ КООРДИНАТ ТОЧЕК

Б. Ю. ЧЕМЕРИЧКО, Н. И. ЖАРИН
Научные руководители Е. Л. СТАРОВОЙТОВА, канд. пед. наук, доц.;
И. И. МАКОВЕЦКИЙ, канд. физ.-мат. наук, доц.
Белорусско-Российский университет
Могилев, Беларусь

Многие слышали про автопилот в Тесле. Однако задавался ли кто-нибудь вопросом про то, как он работает, как машина «видит» дорогу, «видит» окружающее ее пространство, других участников дорожного движения.

Но перед этим рассмотрим поставленную задачу, которая приведет к ответу на ранее заданный вопрос «Как по паре стереоизображений определить расстояние до некоторого объекта (точки) в сцене?». Что такое стереоизображение? Простыми словами – это изображение, полученное с двух камер, которые снимают одну и ту же сцену с разных точек зрения (рис. 1). В нашем случае мы получаем пару стереоизображений, используя стереокамеру (рис. 2).

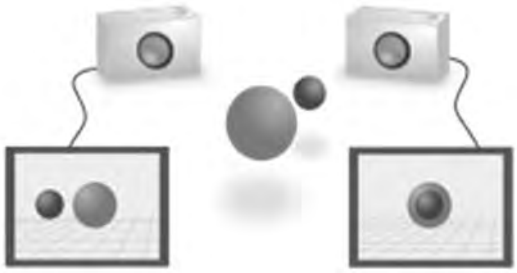

Рис. 1

Рис. 2

Приступаем к решению задачи. Разделы, которые необходимо разобрать для последующего решения, следующие.

1. **Компьютерное стереозрение** – один из методов извлечения информации со сцены при помощи изображений с двух камер (стереопары).
2. **Эпиполярная геометрия** – это геометрия стереозрения.
3. **Триангуляция** – процесс определения положения точки в трехмерном

пространстве по ее проекциям на два (или более) изображений.

Рассмотрим раздел «Эпиполярная геометрия» подробнее, а также введем некоторые основные определения, которые потребуются нам далее.

Базовая модель камеры-обскуры. Рассмотрим центральную проекцию точек в пространстве на плоскость (рис. 3). Пусть центром проекции является начало евклидовой системы координат, а плоскостью $Z = f$, которая называется плоскостью изображения или фокальной плоскостью, является плоскость $Z = f$, которая называется фокальной плоскостью. В модели камеры-обскуры точка в пространстве с координатами $X = (X, Y, Z)^T$ сопоставляется с точкой на плоскости изображения, где линия, соединяющая точку X с центром проекции, пересекается с плоскостью изображения.

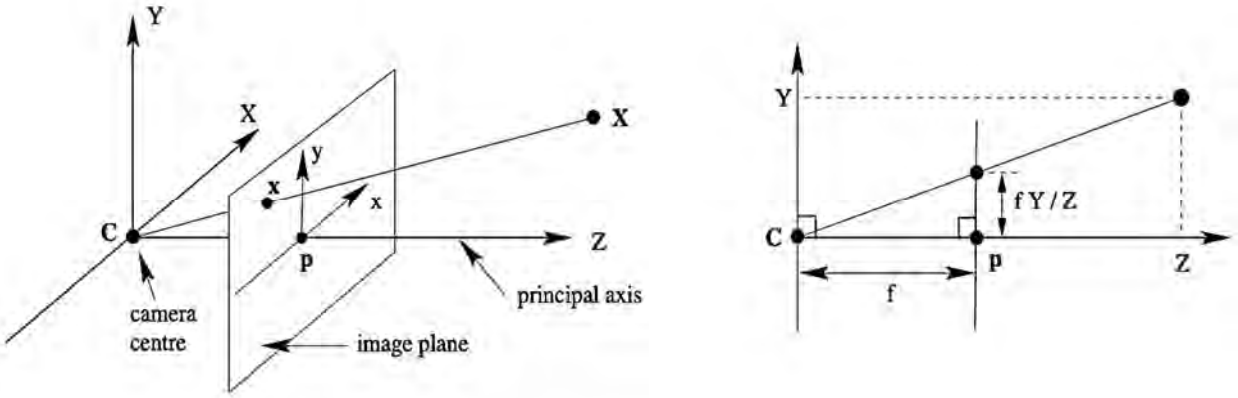

Рис. 3

С помощью похожих треугольников можно быстро вычислить, что точка $(X, Y, Z)^T$ сопоставляется с точкой $\left(\frac{fX}{Z}, \frac{fY}{Z}, f\right)^T$ на плоскости изображения. Игнорируя конечную координату изображения, мы видим, что описывает отображение центральной проекции из мира в координаты изображения. Это отображение из евклидова 3-го пространства R3 в евклидово 2-е пространство IR2.

Теперь непосредственно перейдем к определению положения точки в пространстве.

Рассмотрим модель стереокамеры (рис. 4), где P – это некоторая точка на исследуемом объекте; O_L и O_R – оптические центры двух камер, а точками отображения точки P на фотоприемных матрицах двух камер являются P и P' (плоскость изображения камеры – фокальная плоскость камеры поворачивается и располагается перед объективом); f – фокусное расстояние камеры; B – межосевое расстояние между оптическими центрами двух камер; Z – информация о глубине, которую мы хотим получить.

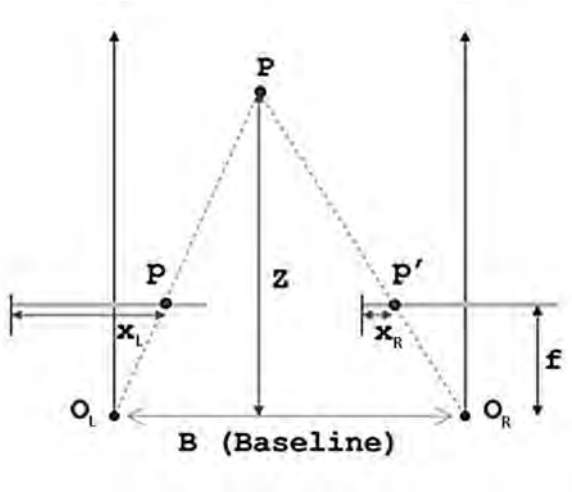

Рис. 4

Необходимо установить расстояние от точки P до точки P' равным dis, затем по принципу подобных треугольников

$$dis = B - (X_L - X_R) \text{ и } \frac{B - (X_L - X_R)}{B} = \frac{Z - f}{Z},$$

отсюда

$$Z = \frac{fB}{X_L - X_R}.$$

В формуле фокусное расстояние f и межосевое расстояние B камеры могут быть получены путем калибровки. Следовательно, если получено значение $X_L - X_R$ (т. е. параллакс d), можно получить информацию о глубине.

Для нахождения значения несоответствия d существует множество способов и алгоритмов. В данной задаче использован наиболее распространенный алгоритм полуглобального блочного сопоставления (Semi-Global Block Matching – SGBM). Данный алгоритм по паре стереоизображений возвращает карту несоответствий (disparity map), используя которую можно получить положение любой точки сцены. В качестве входных данных используются исправленные изображения в оттенках серого. Под исправленными изображениями понимаются изображения, которые выровнены по горизонтали. Если этого не делать, то горизонтальные линии будут искажены и сопоставление пикселей на левом и правом изображениях в таком случае затруднено и не дает очень значимых результатов [1].

После выполнения алгоритма сопоставления получим контраст, который обозначает, что ближайшие к нам пиксели имеют наибольшее смещение

пикселей и поэтому отображаются красным цветом, а объекты, расположенные дальше всего, – синим.

Теперь, когда у нас есть карта несоответствий, мы готовы сгенерировать облако точек. Это можем сделать с помощью функции реконструкции (триангуляции). Функция сцены оценивает точную глубину для каждой точки изображения в исправленном изображении и возвращает массив координат точек трехмерного мира, которые восстанавливают сцену по карте несоответствий и стереопараметрам.

Также следует обратить внимание, что при калибровке стереокамеры для получения точной карты несоответствий необходимо иметь низкую погрешность в пикселях, как правило, ошибки перепроекции менее одного пикселя являются допустимыми для стереозрения. Однако этот параметр зависит от многих особенностей сцены. Для того чтобы можно было находить и различать объекты, изображения должны обладать достаточной детализацией, а сцена – достаточной текстурой или неоднородностью, например, создать облако точек по изображениям гладкой монотонной стены без очевидных особенностей будет достаточно сложно [1].

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Корк, П.** Машинное зрение. Основы и алгоритмы с примерами на Matlab : пер. с англ. / П. Корк. – Москва: ДМК Пресс, 2023. – 584 с.: ил.