

АЛГОРИТМ РЕКОНСТРУКЦИИ ТРЁХМЕРНЫХ ОБЪЕКТОВ СЦЕНЫ СЛОЖНОЙ ФОРМЫ ПО СЕРИИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

© 2018 А. В. Шапаев, Ю. П. Преображенский

Воронежский институт высоких технологий (г. Воронеж, Россия)

В данной статье рассматривается применение алгоритмов обработки цифровых изображений для восстановления трёхмерной модели сцены. Данная статья рассказывает о таких видах алгоритмов обработки цифровых изображений, как реконструкция трёхмерной модели сцены. В современном состоянии, многие такие алгоритмы нуждаются в доработках, в данной статье предложен один из вариантов усовершенствования точности такой реконструкции.

Ключевые слова: трёхмерная модель, стереопара, алгоритмы реконструкции объектов сложной формы, обработка изображений.

Стереоскопическое зрение – это тип зрения, при котором вероятно понимание формы, объемов и дистанции до объекта. Наиболее обычный образец подобного зрения – это бинокулярное видение, для которого применяется два глаза, однако число «глаз» может быть и более.

Головной мозг приобретает два разных изображения, поступающих в него от любого глаза, а принимает их как одно трёхмерное изображение. Невзирая на то, что изображение объектов на сетчатках глаза двумерное, человек принимает общество трехмерным. На базе этого принципа и основывается компьютерное преобразование.

В геометрии стерео зрения существенную значимость представляет проективная геометрия. В проективной геометрии имеется ряд раскладов: геометрический (аналогично Евклидовой геометрии можно ввести представление геометрических предметов, аксиом и из данного извлекать все без исключения качества проективного пространства), аналитический (расценивать все в координатах, как в аналитическом раскладе к Евклидовой геометрии), алгебраический.

На сегодняшний день, как в области практического применения, так и исследовательских задач, часто возникает проблема восстановления модели трехмерного объекта по двумерным изображениям (стереопарам).

Построение моделей вручную сопряжено со значительными затратами, а использо-

вание специальной аппаратуры для сканирования трёхмерных объектов не всегда возможно, так как глубина сканирования такими устройствами ограничена а также стоимость такой аппаратуры является достаточно большой.

Большинство современных систем реконструкции производят восстановление модели, в которую попадают только особые точки. Количество особых точек на некоторых фотографиях может быть очень малым (и всегда существенно ниже количества пикселей), а распределение их неравномерным. В результате чего трехмерная модель объекта получается малоинформативной.

В случае, когда реконструкция производится по фотографиям, полученным с помощью стереокамеры, можно использовать алгоритм на основе карты глубины. Данный алгоритм позволяет получить объект, где для каждого пикселя на карте глубины производится восстановление трехмерных координат соответствующей точки модели.

Данный метод имеет свои отрицательные стороны – невозможно восстановить с необходимой точностью геометрические характеристики объекта-оригинала в трехмерном пространстве. Модель получается малоинформативной.

Для исключения данных недостатков был разработан алгоритм, который совмещает в себе преимущества двух данных методов.

Особенностью данного алгоритма является то, что входными данными являются не разрозненные фотографии сцены, а видеоряд, который по существу является последовательностью фотографий. Поэтому каждые

Шапаев Александр Викторович – ВИВТ-АНОО ВО, аспирант, alieksandr.shapaiev.94@mail.ru.
Преображенский Юрий Петрович – ВИВТ-АНОО ВО, к. т. н., профессор, petrovich@vivt.ru.

две фотографии объекта можно рассматривать как стереопару.

Для того чтобы получить карту глубины, необходимо восстановить внутренние и внешние параметры для камеры, с которой производилась съемка. Чем точнее будет найдено положение камеры в момент съемки каждого кадра, тем достовернее получится объект после реконструкции.

Процесс калибровки камеры заключается в съемке специального калибровочного шаблона, для которого параметры заранее известны. Для восстановления внутренних параметров камеры можно воспользоваться готовым решением – например библиотекой OpenCV, поэтому этот этап в данной статье подробно рассматриваться не будет.

Для того чтобы произвести реконструкцию, необходимо знать относительное положение камеры в момент съемки кадра. Эти положения можно восстановить, если рассматривать соседние фотографии в серии как стереопары. Процесс восстановления относительного положения основан на нахождении соответствий на изображениях в стереопаре. В общем случае большинство алгоритмов сопоставления изображений работает по следующей схеме:

1. На изображениях акцентируются основные точки и их дескрипторы.
2. Согласно совпадению дескрипторов акцентируются соответствующие друг другу основные точки.

Алгоритм «SIFT» (Scale Invariant Feature Transform) единственный из более зачастую применяемых алгоритмов описания особых пунктов. Точки, приобретенные с помощью метода, инвариантны к масштабированию и поворотам изображения, устойчивы к переменам освещения, шумам и переменам позиции наблюдающего.

Алгоритм «SURF» (Speeded Up Robust Features) базируется в поиске специальных пунктов и формировании дескрипторов, инвариантных к масштабированию и вращению с поддержкой матрицы Гессе. При этом для любой точки считается вектор наибольшего изменения яркости и показатель масштабирования согласно матрице Гессе.

Дескриптор «GLOH» (Gradient location-orientation histogram) считается моделью SIFT-дескриптора, который создан с целью увеличения прочности. Рассчитывается SIFT дескриптор, однако применяется полярная сетка разбиения окрестности на бины. 3 ра-

диальных блока с радиусами 6, 11 и 15 пикселей и 8 разделов. В следствии выходит вектор, включающий 272 элементы, какой проецируется в место размерности 128 с помощью применения рассмотрения основных элементов (PCA).

Дескриптор «BRIEF» (Binary Robust Independent Elementary Features). Задача формирования этого дескриптора заключалась в том, чтобы гарантировать определение схожих зон изображения, которые были убраны с различных точек зрения.

При этом устанавливалась цель предельно сократить число исполняемых вычислений. Метод определения объединяется к концепции ненамеренного леса (randomize classification trees) либо наивного Байесовского классификатора на определенном тренировочном множестве изображений и дальнейшей систематизации зон испытательных изображений.

В простом варианте может применяться способ близкого соседа с целью розыска более схожего патча в учебной выборке. Незначительное число действий поддерживается за счет представления вектора свойств в виде бинарной строки, а как следствие, применения в свойстве меры сходства расстояния Хэмминга.

В результате сопоставления информации дескрипторов было выявлено, то что при схожих обстоятельствах выполнения исследований в отдельных испытательных изображениях достоверность детектирования с помощью BRIEF практически в 1.5 раза выше, нежели с применением SURF-дескрипторов и в среднем выше, чем у абсолютно всех других.

Матрица F называется фундаментальной матрицей (fundamental matrix). Ее разряд равен 2, она установлена с правильностью вплоть до ненулевого множителя и находится в зависимости только от матриц начальных камер P и P' .

В случае, если матрицы камер имеют тип:

$$P = K[l|0], \\ P' = K'[R|t].$$

Тогда фундаментальная матрица может быть вычислена по формуле:

$$F' = K'^{-1T} [RK''t]x,$$

где для вектора e обозначение $[e]_k$ вычисляется как:

$$[e]k = \begin{bmatrix} 0 & -c_z & -c_y \\ c_z & 0 & -c_k \\ -c_y & c_k & 0 \end{bmatrix}.$$

Соответственно, если фундаментальная матрица и внутренние параметры камеры известны то, решив уравнение, мы можем вычислить расположение камеры в момент съемки.

Карта глубины (depth map) – это изображение, на котором для любого пикселя, вместо расцветки, содержится его расстояние вплоть до камеры. Схема глубины может быть получена с помощью особой камеры глубины (к примеру, сенсор Kinect является своего семейства такого рода камерой), а так же может быть создана согласно стереопаре изображений.

Идея, лежащая в базе построения карты глубины согласно стереопаре весьма легка.

Для каждой точки на одном изображении производится отбор парной ей точки в другом изображении. А согласно паре определенных пунктов возможно осуществить возобновление местоположение точки в многомерном пространстве. Зная многомерные координаты прообраза, глубина рассчитывается, как дистанция до плоскости камеры.

Парную точку необходимо находить в эпиполярной линии. Соответственно, с целью упрощения поиска, изображения выравнивают таким образом, чтобы все эпиполярные направления были параллельны граням изображения (как правило, горизонтальны).

Более этого, изображения равняют так, чтобы для точки с координатами (x_0, y_0) соответствующая ей эпиполярная линия задавалась уравнением $x = x_0$, в то время для любой точки соответствующую ей парную точку необходимо находить в той же строке на изображении со второй камеры.

Такой ход выравнивания изображений именуют ректификацией. Как правило, ректификацию делают посредством ремеппинга изображения и ее соединяют с освобождением от дисторсий.

После этого как изображения ректифицированы, осуществляют отбор определенных пар точек.

Самый обычный метод складывается в следующем. Для любого пикселя левой картинке с координатами (x_0, y_0) производится отбор пикселя на левой картинке. При этом подразумевается, то что пиксель в левой

картинке должен иметь координаты $(x_0 - d, y_0)$, где d – величина называемая несоответствие/смещение (disparity).

Поиск соответствующего пикселя производится посредством вычисления максимума функции отклика, в свойстве которой может выступать, например, корреляция окрестностей точек. В следствии выходит схема смещений (disparity map) либо схема глубины (depth map).

После получения всех необходимых параметров можно произвести реконструкцию вексельной модели объекта из полученной карты глубины и матрицы перспективного преобразования. Данное действие также можно выполнить с помощью библиотеки OpenCV.

ЛИТЕРАТУРА

1. Варакин Л. Е. Системы связи с шумоподобными сигналами / Л. Е. Варакин. – М.: Радио и связь, 2002.
2. Гуляев Ю. В. Новый класс сигналов для передачи информации: широкополосные хаотические сигналы. / Ю. В. Гуляев, В. Я. Кислов, В. В. Кислов // ДАН. – 2005. – Т. 359. – № 6. – С. 750-754.
3. Зяблов Е. Л. Разработка лингвистических средств интеллектуальной поддержки на основе имитационно-семантического моделирования / Е. Л. Зяблов, Ю. П. Преображенский // Вестник Воронежского института высоких технологий. – 2009. – № 5. – С. 024-026.
4. Казаков Е. Н. Разработка и программная реализации алгоритма оценки уровня сигнала в сети Wi-Fi / Е. Н. Казаков // Моделирование, оптимизация и информационные технологии. – 2016. – № 1 (12). – С.13.
5. Максимов И. Б. Принципы формирования автоматизированных рабочих мест / И. Б. Максимов // Вестник Воронежского института высоких технологий. – 2014. – № 12. – С. 130-135.
6. Максимов И. Б. Классификация автоматизированных рабочих мест / И. Б. Максимов // Вестник Воронежского института высоких технологий. – 2014. – № 12. – С. 127-129.
7. Паневин Р. Ю. Структурные и функциональные требования к программному комплексу представления знаний / Р. Ю. Паневин, Ю. П. Преображенский // Вестник Воронежского института высоких технологий. – 2008. – № 3. – С. 061-064.

THE ALGORITHM OF RECONSTRUCTION OF THREE DIMENSIONAL SCENE OBJECTS OF COMPLEX SHAPE BY A SERIES OF DIGITAL IMAGES

© 2018 A. V. Shapaev, Yu. P. Preobrazhenskiy

Voronezh Institute of High Technologies (Voronezh, Russia)

This paper discusses the use of digital image processing algorithms to restore the three-dimensional model of the scene. This article talks about these kinds of digital image processing algorithms as the reconstruction of three-dimensional model of the scene. In the present state, many such algorithms need to be improved in this paper proposes one of the options to improve the accuracy of such reconstruction.

Key words: three-dimensional model, a stereo pair, reconstruction algorithms of complex objects, image processing.