

Методика подготовки данных для восстановления трёхмерной структуры сцены***С. С. Тараненко, О. Ф. Ковалёв**

Рассматривается проблема подготовки данных для восстановления трёхмерной структуры сцены, составленной из точек, отрезков, прямых линий и т. д., используя видеоинформацию, получаемую с камеры, движущейся вокруг статической сцены. На основе анализа существующих методов представлена методика и алгоритм для выделения и подготовки данных для восстановления трёхмерной структуры сцены. За основу алгоритма взят фильтр Канни с оператором Собеля, дополненный вычислением дескрипторов из алгоритмов SURF (Speeded Up Robust Feature). Для фильтрации шумов с кадров предлагается использовать wavelet-фильтры, что позволит получить алгоритм выделения контура объекта в кадре и сопоставление его с контуром объекта на следующем кадре, что является входными данными для большинства алгоритмов восстановления трёхмерной сцены.

Ключевые слова: камера-обскура, SURF (Speeded Up Robust Feature), особые точки, трёхмерная структура, фильтр Калмана, фильтр Канни, подавление немаксимумов, сцена, градиент.

Введение. Область практического применения реалистической компьютерной графики в наши дни необычайно широка и включает в себя системы виртуальной реальности, реверс-инжиниринг, киноиндустрию, дизайн, компьютерные игры, научную визуализацию и т. д. Стремительно увеличивается число продуктов компьютерной графики, а вовлечение их в повседневную жизнь вызывает ответную реакцию потребителя, заключающуюся в непрерывном повышении требований к этим продуктам. И хотя рост вычислительных мощностей позволяет решать всё более сложные задачи визуализации трёхмерных сцен (число объектов, источников освещения, разрешение изображения), рост сложности самих задач опережает его, и актуальность этих задач растёт.

Одной из ведущих задач в компьютерной графике стало восстановление трёхмерной структуры сцены. Для восстановления структуры сцены требуется сбор информации о ней. Один из способов сбора информации о сцене — это видео- и фотосъёмка, но в полученных кадрах много шумов и второстепенной информации, кроме того информация с двух соседних кадров требует согласования и сведения кадров в стереопару. По данной задаче проводились исследования и есть результаты, но они не совершенны.

Исторический обзор. Первый учёный, который начал рассматривать проблему восстановления структуры движущегося объекта, был Ульман [1], и его работа породила много алгоритмов. Эти алгоритмы отличаются по типу ввода (например, плотные полутоновые изображения, характерные точки, строки и т. д.), числу требуемых изображений (например, два, три, неограниченное количество), моделям камеры (например, перспективная или ортогональная проекция), и принятому знанию движения камеры (например, полностью известная, постоянная скорость, неизвестная скорость и т. д.).

Многие исследователи занимались проблемой восстановления структуры сцены, составленной из множества точек ряда изображений. Лонгуст-Хиггинс и другие [2–6] показали, что структура ряда точек может быть восстановлена с двух изображений, даже когда относительная позиция этих двух камер неизвестна. Другие исследователи разработали алгоритмы, которые используют повторные изображения твёрдой сцены, чтобы произвести более точную реконструкцию, и они попадают в два класса: пакетные алгоритмы, которые обрабатывают все данные одновременно и инкрементные алгоритмы, которые поддерживают некоторое понятие состояние, которое обновляется с каждым новым кадром.

* Работа выполнена в рамках инициативной НИР.

В этой работе представлена методика подготовки данных для восстановления трёхмерной структуры сцены, используя данные видеосъёмки, полученной с движущейся камеры.

Описание алгоритма нахождения особых точек. Работа алгоритма по восстановлению трёхмерной структуры сцены заключается в определении положения точек в пространстве, основываясь на их двумерных координатах, полученных с камеры. Это возможно, так как кадры содержат объект с разных ракурсов и образуют стереопары. Для идентификации и сопоставления уникальных точек на кадрах с камеры предложена методика на основе алгоритма SURF (Speeded Up Robust Feature) [7].

Рассмотрим модель камеры и получаемых кадров.

Модель камеры. Камера-обскура (рис. 1) представляет собой светонепроницаемый ящик в центре одной из стенок которого находится маленькое отверстие. На внутренней стороне противоположной стенки ящика закреплена фоточувствительная пластина. Камера имеет единственный внутренний параметр — фокусное расстояние f , то есть расстояние от отверстия до противоположной стенки ящика.

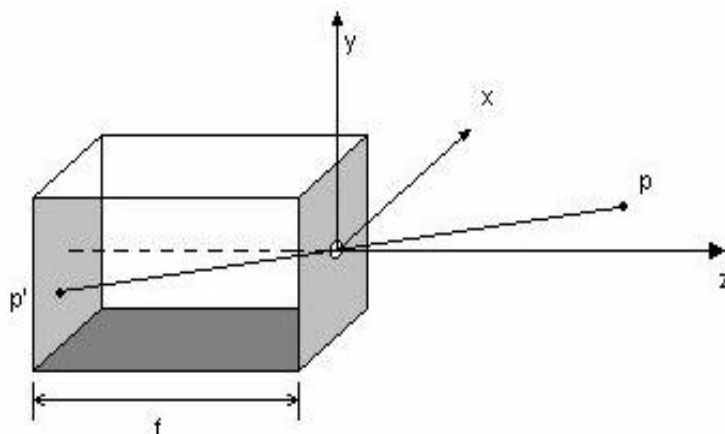


Рис. 1. Камера-обскура

При моделировании камеры-обскуры светонепроницаемый ящик заменяют системой «точка наблюдения — плоскость изображения» (рис. 2).

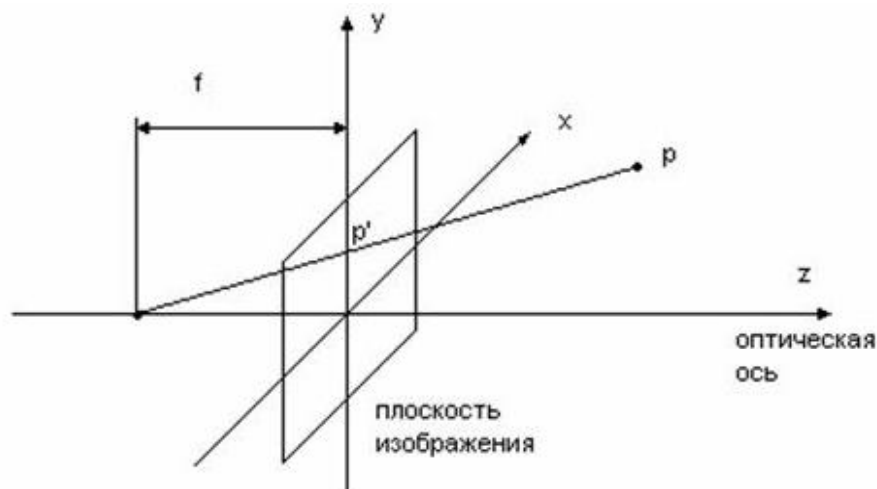


Рис. 2. Система «точка наблюдения — плоскость изображения»

Между координатами точки $P(X, Y, Z)$ в пространстве сцены и точкой (x, y) в плоскости изображения камеры-обскуры существует следующее соотношение:

$$\begin{aligned} x &= f \cdot \frac{X}{Z} + \eta_x, \\ y &= f \cdot \frac{Y}{Z} + \eta_y, \end{aligned} \quad (1)$$

где f — фокусное расстояние камеры; η_x и η_y — шумы.

Оцифрованное изображение хранится в буфере кадра, который можно рассматривать как матрицу пикселей размером $W \cdot H$. Пусть (i, j) — дискретные координаты буфера кадра с началом в верхнем левом углу, (Ox, Oy) — фокусная точка линзы (место пересечения оптической осью плоскости изображения) в координатах буфера кадра и (x, y) — координаты изображения.

Отношение координат плоскости изображения и координат буфера кадра:

$$\begin{aligned} x &= (i - Ox) \cdot S_x, \\ y &= (j - Oy) \cdot S_y, \end{aligned} \quad (2)$$

где S_x, S_y — горизонтальное и вертикальное расстояние двух соседних пикселей в буфере кадра.

Таким образом, имея оцифрованное изображение снимка с камеры, можно определить координаты каждого пикселя в плоскости изображения [8].

Выделение границ объектов, алгоритм Канни с оператором Собеля. Опираясь на результаты работы Канни, который изучил математическую проблему получения фильтра, оптимального по критериям выделения, локализации и минимизации нескольких откликов одного края, можно сказать, что предложенный им детектор реагирует на границы, но при этом игнорирует ложные, точно определяет линию границы и реагирует на каждую границу один раз, что позволяет избежать восприятия широких полос изменения яркости как совокупности границ. Канни ввёл понятие Non-Maximum Suppression (подавление немаксимумов), которое означает, что пикселями границ объявляются точки, в которых достигается локальный максимум градиента в направлении вектора градиента [9, 10].

Алгоритм состоит из пяти отдельных шагов:

1. Сглаживание. Подавление шума и размытие границ. Для оптимизации вычислений исходный кадр преобразуется в оттенки серого (с помощью цветовой модели YUV (или HSL, HSV, других)).

Для подавления шума, воспользуемся Wavelet-фильтрами (математические функции, позволяющие анализировать различные частотные компоненты данных).

2. Поиск градиентов. Максимальное значение градиента кадра отмечает границу.

Оператор Собеля — это дискретный дифференциальный оператор, вычисляющий приближённое значение градиента яркости изображения. Результатом применения оператора Собеля в каждой точке изображения является либо вектор градиента яркости в этой точке, либо его норма. Оператор Собеля основан на свёртке изображения небольшими целочисленными фильтрами в вертикальном и горизонтальном направлениях. Используемая им аппроксимация градиента достаточно грубая, особенно это сказывается на высокочастотных колебаниях изображения.

Ядра фильтров:

$$D_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}, D_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \quad (3)$$

Угол вектора квантуется по 45° .

3. Подавление немаксимумов. Только локальные максимумы отмечаются как границы, остальные подавляются.

4. Двойная пороговая фильтрация. Границы определяются порогами.

Следующий шаг — применение порога, чтобы определить находится ли граница в данной точке изображения. Чем меньше порог, тем больше границ будет находиться, но тем более восприимчив результат к шуму. Высокий порог может проигнорировать слабые края или получить границу фрагментами. Поэтому в фильтре используется два порога: верхний и нижний.

5. Трассировка области неоднозначности. Итоговые границы определяются путём подавления всех границ, несвязанных с найденными границами.

Данный алгоритм позволяет выделить границы в кадре, но перед нами стоит задача ещё и сопоставить выделенные границы (набор особых точек объекта) сопоставить на всех кадрах, где они встречаются. Для этого будем использовать дескриптор особой точки.

Вычисление дескриптора особой точки. За основу можно взять вычисление дескрипторов точек в SURF алгоритмах. Дескриптор — это массив чисел независимых от вращения и масштаба, которые позволяют идентифицировать особую точку.

Дескриптор считается как описание градиента для 16 квадрантов вокруг особой точки рис. 3.

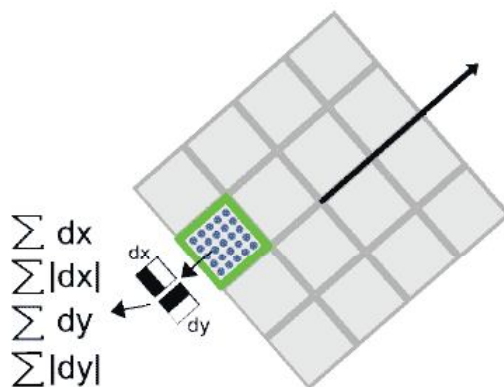


Рис. 3. Описание градиента для 16 квадрантов вокруг особой точки

Берётся регулярная сетка 5×5 и для точки сетки ищется градиент, с помощью фильтра Хаара. После нахождения 25 точечного градиента квадранта, вычисляются четыре величины, которые собственно и являются компонентами дескриптора: $\sum dx, \sum |dx|, \sum dy, \sum |dy|$.

Две из них — это суммарный градиент по квадранту, а две других — сумма модулей точечных градиентов.

На рис. 4 показано поведение этих величин для разных участков изображений.



Рис. 4. Зависимость компонентов дескриптора от градиента

Четыре компонента на каждый квадрант — 16 квадрантов дают 64 компонента дескриптора для всей области особой точки. При занесении в массив значения дескрипторов взвешиваются

на гауссиану, с центром в особой точке и с сигмой, что придаёт большую устойчивость дескриптора к шумам.

В SURF так же используется следа матрицы Гессе для различия светлых и тёмных пятен. В нашем случае этого не нужно, т. к. мы имеем границы и с помощью дескриптора сопоставляем их на кадрах.



Рис. 5. Результаты выделения особых точек и их сопоставление на двух кадрах

Заключение. Приняв во внимание то, что на кадрах из видеопотока особые точки объекта существенно переместиться не могут, получаем аксиому, что примерное положение особых точек на последующем кадре при непрерывной съёмке резко изменится не может, что позволяет описанному алгоритму получать входные данные из видеопотока для алгоритма восстановления трёхмерной сцены, а именно списки особых точек (составляющие границы объекта) с двух и более кадров, образующих стереопары. Далее используя один из алгоритмов восстановления трёхмерных объектов, например фильтр Калмана, можно получить трёхмерную модель сцены.

Для повышения количества особых точек и качества их определения (подавления шумов), имеем два порога — минимальный и максимальный, как следствие применения основы фильтра Канни. Благодаря особенностям семейств wavelet фильтров и переключая между ними, можно добиться адекватного восприятия особых точек алгоритмом в видеопотоке с разной степенью зашумлённости.

Библиографический список

1. Ullman, S. The Interpretation of Visual Motion. Cambridge : MIT Press, 1979.
2. Longuet-Higgins, H. C. A computer algorithm for reconstructing a scene from two projections. *Nature*, 1981, vol. 293, pp. 133–135.
3. Weng, J., Huang, T. S., Ahuja, N. Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. *IEEE Trans. Pattern Anal. Machine Intell.*, 1989, vol. 11, no. 5, pp. 451–476.
4. Hu, X., Ahuja, N. Motion and structure estimation using long sequence motion models. *Image and Vision Computing*, 1993, vol. 11, no. 9, pp. 549–570.
5. Horn, B. K. P. Relative orientation. *Int. J. Computer Vision*, 1990, vol. 4, pp. 59–78.
6. Weng, J., Huang, T. S., Ahuja, N., Motion and Structure from Image Sequences, Springer Series on Information Sciences. Berlin, Springer-Verlag, 1993.

7. Herbert, B., Andreas, E., Tinne, T., Luc, V. G. SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU). 2008, vol. 110, no. 3, pp. 346–359.
8. Грузман, И. С. Цифровая обработка изображений в информационных системах / И. С. Грузман [и др.]. — Новосибирск : Новосиб. гос. техн. ун-т, 2002. — 352 с.
9. Canny, J. A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 1986, vol. 8, no. 6, pp. 679–698.
10. Wyman, C. Canny Edge Detection. Advanced Computer Graphics (Advanced OpenGL Rendering), Syllabus for Spring 2008. Available at: <http://homepage.cs.uiowa.edu/~cwyman/classes/spring08-22C251/homework/canny.pdf> (accessed: 05.05.2014).
11. Kalra, P. K. Canny Edge Detection. Lectures on Digital Image Processing, 2009. Available at: <http://www.cse.iitd.ernet.in/~pkalra/csl783/canny.pdf> (accessed: 03.02.2014).
12. Moeslund, T. B. Image and Video Processing. Computer Vision and Media Technology, Aalborg University, 2009.

Материал поступил в редакцию 28.05.2014.

References

1. Ullman, S. The Interpretation of Visual Motion. Cambridge : MIT Press, 1979.
2. Longuet-Higgins, H. C. A computer algorithm for reconstructing a scene from two projections. Nature, 1981, vol. 293, pp. 133–135.
3. Weng, J., Huang, T. S., Ahuja, N. Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. IEEE Trans. Pattern Anal. Machine Intell., 1989, vol. 11, no. 5, pp. 451–476.
4. Hu, X., Ahuja, N. Motion and structure estimation using long sequence motion models. Image and Vision Computing, 1993, vol. 11, no. 9, pp. 549–570.
5. Horn, B. K. P. Relative orientation. Int. Journal of Computer Vision, 1990, vol. 4, no. 1, pp. 59–78.
6. Weng, J., Huang, T. S., Ahuja, N. Motion and Structure from Image Sequences, Springer Series on Information Sciences. Berlin, Springer-Verlag, 1993.
7. Herbert, B., Andreas, E., Tinne, T., Luc, V. G. SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU). 2008, vol. 110, no. 3, pp. 346–359.
8. Грузман, И. С. Tsifrovaya obrabotka izobrazheniy v informatsionnyih sistemah. [Digital image processing in information systems]. Novosibirsk : NGTU, 2002, 352 p. (in Russian).
9. Canny, J. A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence, 1986, vol. 8, no. 6, pp. 679–698.
10. Wyman, C. Canny edge detection. Advanced Computer Graphics (Advanced OpenGL Rendering), Syllabus for Spring, 2008. Available at: <http://homepage.cs.uiowa.edu/~cwyman/classes/spring08-22C251/homework/canny.pdf> (accessed: 05.05.2014).
11. Kalra, P. K. Canny Edge Detection. Lectures on Digital Image Processing, 2009. Available at: <http://www.cse.iitd.ernet.in/~pkalra/csl783/canny.pdf> (accessed: 03.02.2014).
12. Moeslund, T. B. Image and Video Processing. Computer Vision and Media Technology, Aalborg University, 2009.

DATA PREPARATION METHODS FOR RECOVERING THREE-DIMENSIONAL STRUCTURE OF THE SCENE*

S. S. Taranenko, O. F. Kovalev

The problem of data preparation to recover the three-dimensional structure of the scene including points, segments, straight lines, etc., using video obtained from a moving camera around a static scene is considered. Based on the analysis of the existing methods, the methodology and algorithm of the data selection and preparation for recovering the three-dimensional scene structure are provided. Canny filter and Sobel operator complemented with calculating the algorithm SURF (Speeded Up Robust Feature) descriptors is taken as a basis. To filter noise from frames, wavelet filters are proposed. This allows obtaining an object contour extraction algorithm in the frame, and its comparison with the object contour in the next frame, which is the input data for the majority of algorithms of the three-dimensional scene reconstruction.

Keywords: camera-obscure, SURF (Speeded Up Robust Feature), singular points, three-dimensional structure, Kalman filter, Canny filter, non-maxima suppression, scene, gradient.

* The research is done within the frame of the independent R&D.