

А. А. ЗАХАРОВ

АВТОМАТИЧЕСКИЙ СИНТЕЗ ПРОТЯЖЕННЫХ ТРЕХМЕРНЫХ СЦЕН С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Рассмотрен способ автоматического синтеза протяженных трехмерных сцен по видеопоследовательности, основанный на поиске точечных соответствий. Представлены примеры реконструкции по изображениям.

Ключевые слова: автоматический синтез, трехмерная реконструкция, видеопоследовательность.

Одной из основных задач компьютерного зрения является восстановление формы объектов по изображениям. Выделяют активные и пассивные оптические методы трехмерной реконструкции. Активные методы основаны на использовании дорогостоящих дальнометров. Пассивные методы трехмерной реконструкции включают в себя обработку двумерных изображений. К пассивным методам [1, 2] относится реконструкция: по стереоизображениям (shape from stereo); по закраске изображения (shape from shading); по текстуре (shape from texture); по фокусировке (shape from focus); по контуру (shape from contour (silhouettes)); по перемещению (shape from motion).

В последнее время активно ведется разработка методов трехмерной реконструкции по перемещению (shape from motion) [1—4]. В этом случае происходит восстановление формы сцены из последовательности изображений, полученных с разных позиций съемки в некоторые моменты времени. Часто положения камеры тоже считаются неизвестными и восстанавливаются в ходе решения задачи. Для создания системы трехмерной реконструкции предлагается использовать метод, основанный на перемещении камеры в пространстве. Этот метод получил название „реконструкция по видеопоследовательности“ (shape from video), однако восстановление геометрии объектов сцены в подобных системах затруднено и требует значительного участия оператора. Поэтому необходимо разрабатывать новые методы и программные средства автоматической реконструкции трехмерных сцен с геометрией различной сложности.

Автоматическая генерация трехмерных сцен. Из всего набора изображений выбираются два соседних кадра для нахождения соответствий с соседним кадром последовательности.

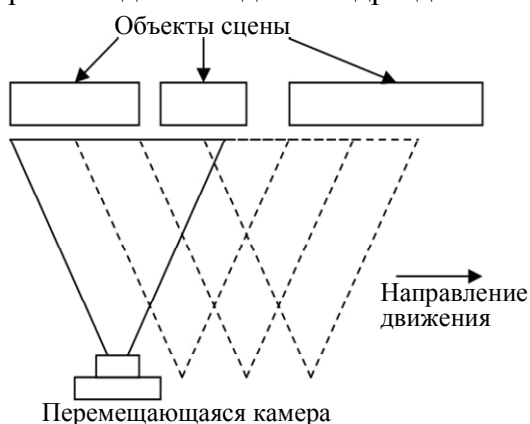


Рис. 1

По изображениям в кадрах вычисляются характерные опорные точки трехмерной сцены. Последующие изображения обрабатываются уже относительно предыдущих кадров. На каждом этапе первоначальная реконструкция уточняется и расширяется.

На начальном этапе происходит получение набора растровых изображений при помощи цифровой видеокамеры, съемка объекта ведется с разных позиций. Направление оптической оси камеры строго перпендикулярно направлению движения (рис. 1). Текущие координаты камеры определяются с помощью системы спутниковой навигации. Погрешность определения координат камеры не накладывает существенных ограничений на точность реконструкции, так как движение происходит по прямой линии.

Погрешность определения координат камеры не накладывает существенных ограничений на точность реконструкции, так как движение происходит по прямой линии.

Далее происходит выделение точечных особенностей, т.е. точек, окрестность которых отличается от окрестностей других точек изображения по выбранной мере. Одним из наиболее известных алгоритмов нахождения точечных особенностей является детектор Харриса, обладающий высоким быстродействием [6].

На следующем этапе находятся соответствия между точечными особенностями различных изображений одной и той же сцены. Для этого используются дескрипторы — идентификаторы ключевой точки, выделяющие ее из группы точек. По совпадению дескрипторов выделяются соответствующие друг другу ключевые точки. Дескрипторы должны обеспечивать инвариантность нахождения соответствия между особыми точками относительно преобразований изображений. В настоящей работе был использован дескриптор SIFT (Scale Invariant Feature Transform).

Затем вычисляются координаты опорных точек, проекции которых были выделены на изображении детектором. Два снимка с изображениями одного и того же участка местности, полученные с двух точек пространства, называются стереоскопической парой снимков (стереопарой). Для вычислений используется понятие эпиполярной конфигурации [5—8]. В простейшем случае две одинаковые параллельные плоскости изображений параллельны базе стереосистемы (рис. 2).

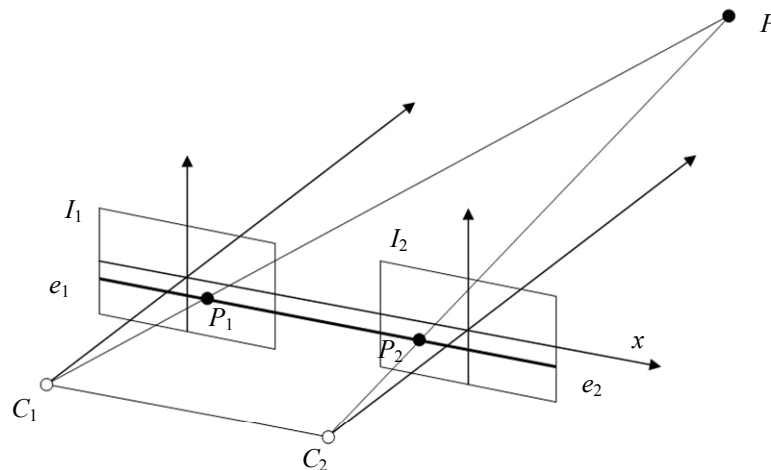


Рис. 2

Эпиполярной называется плоскость, которой принадлежит трехмерная точка P , две ее проекции P_1 и P_2 на плоскости изображений, оптические центры двух камер C_1 и C_2 . Эпиполярными называются две прямые e_1 и e_2 , которые являются пересечениями эпиполярной плоскости с двумя плоскостями изображений I_1 и I_2 . Расстояние между оптическими центрами камер называется базой b . Зная параметры эпиполярной конфигурации, можно вычислить пространственные координаты точки P .

Камера заранее калибруется, чтобы определить, как расположены пикселы на плоскости изображения данной камеры относительно точек трехмерного пространства, изображения которых требуется получать с помощью камеры. Выделяют внутренние и внешние параметры камеры. К внутренним относятся точка пересечения оптической оси и плоскости изображения, масштабные множители для горизонтального и вертикального размеров пиксела, фокусное расстояние f , множитель дисторсии объектива.

Внешние параметры описывают местоположение и ориентацию системы координат камеры в трехмерном пространстве. К ним принадлежат параметры переноса и параметры поворота.

В настоящей работе рассматривается частный случай: известны внутренние и внешние параметры камеры. Так как оптическая ось камеры всегда направлена перпендикулярно движению, то параметры поворота не учитываются.

В рассматриваемой системе координаты проекций точки P на плоскости изображений можно вычислить следующим образом:

$$x_{P_1} = f(x_P + b/2) / z_P, \quad (1)$$

$$x_{P_2} = f(x_P - b/2) / z_P, \quad (2)$$

$$y_{P_1} = y_{P_2} = fy_P / z_P. \quad (3)$$

Из первых двух соотношений следует, что

$$z_P = fb / (x_{P_1} - x_{P_2}). \quad (4)$$

Таким образом, можно вычислить координаты x и y точки P

$$x_P = b \frac{x_{P_1} + x_{P_2}}{2(x_{P_1} - x_{P_2})}, \quad y_P = b \frac{y_{P_1} + y_{P_2}}{2(x_{P_1} - x_{P_2})}. \quad (5)$$

На последнем этапе происходит построение трехмерной модели по вычисленным координатам с использованием полигональной геометрии. Этот этап является одним из наиболее простых в реализации, так как существуют эффективные алгоритмы построения поверхности по известным координатам вершин.

Результаты реконструкции. При реализации алгоритма необходимо находить компромисс между размером поля зрения, точностью вычисления трехмерных координат и количеством найденных точечных соответствий. При малой базе небольшие исходные погрешности в определении координат могут привести к значительным ошибкам вычисления глубины. Большая база может привести к трудностям в сопоставлении точечных особенностей из-за взаимного перекрытия объектов на изображениях и изменения изображения сцены.

Были проведены экспериментальные исследования алгоритма на тестовом объекте. По полученным результатам съемки сцены с нескольких позиций были получены зависимости абсолютной погрешности измерения трехмерных координат и определения количества найденных точечных пар от угла смещения относительно исходного положения камеры. Анализ зависимостей показал, что наибольшая точность вычислений координат обеспечивается при угле смещения камеры $25\text{--}45^\circ$. Однако с увеличением угла смещения значительно снижается количество автоматически найденных точечных соответствий. Таким образом, для системы были выбраны параметры, при которых угол смещения относительно текущего положения был равен 25° . Была проведена реконструкция трехмерной сцены с использованием реальных снимков (рис. 3).



Рис. 3

Следует, однако, отметить, что качество синтезируемой сцены во многом зависит от сложности растрового рисунка. Алгоритм не всегда корректен для обработки изображений, на которых представлены объекты сложной геометрической структуры или присутствуют взаимные перекрытия.

СПИСОК ЛИТЕРАТУРЫ

1. *Hartley R.* Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.
2. *Szeliski R.* Computer Vision: Algorithms and Applications. NY: Springer-Verlag, 2010.
3. *El-Hakim S.F.* Semi-automatic 3d reconstruction of occluded and unmarked surfaces from widely separated views // Proc. ISPSRS Symp. Corfu, 2002. P. 143—148.
4. *Pollefeys M.* Detailed realtime urban 3D reconstruction from video // Inten. J. of Computer Vision. 2008. Vol. 78 (2—3). P. 143—167.
5. *Март Д.* Зрение. Информационный подход к представлению и обработке зрительных образов. М.: Радио и связь, 1987.
6. *Форсайт Д.* Компьютерное зрение. Современный подход. М.: Вильямс, 2004.
7. *Хорн Б.К.П.* Зрение роботов. М.: Мир, 1989.
8. *Шапиро Л.* Компьютерное зрение. М.: БИНОМ. Лаборатория знаний, 2006. 752 с.
9. *Harris C.G.* A combined corner and edge detector // Proc. 4th Alvey Vision Conf. Manchester, 1988. P. 147—151.

Сведения об авторе

Алексей Александрович Захаров — канд. техн. наук; Муромский институт Владимирского государственного университета им. А. Г. и Н. Г. Столетовых, кафедра информационных систем; E-mail: aa-zaharov@ya.ru

Рекомендована Юго-Западным
государственным университетом

Поступила в редакцию
24.10.11 г.