

М. С. ЛАРИН

МОДЕЛИРОВАНИЕ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Рассматривается задача моделирования дополненной реальности и основные проблемы, связанные с обработкой изображений и извлечением пространственных координат.

Ключевые слова: виртуальное пространство, дополненная реальность.

Теоретические основы моделирования виртуального пространства. Виртуальное пространство моделируется с использованием аффинных преобразований [1], когда вершина для отображения ее на экране преобразуется из трехмерных координат в двухмерные. Входными данными являются координаты, а выходными — изображение. При построении дополненной реальности — наложении виртуальных объектов на изображение, фиксируемое веб-камерой [2, 3], — необходимо решить обратную задачу, где входными данными будет изображение, а выходными — координаты. Подобные преобразования используются в фотограмметрии для определения размеров объектов. Основная проблема при построении дополненной реальности заключается в определении относительных систем координат сцены (изображения) и веб-камеры, т.е. в локализации веб-камеры. Эта задача состоит из нескольких подзадач: сопоставить неподвижную систему координат со сценой; выбрать эквивалент виртуальной камеры и установить ее положение, при этом положение камеры определяется ее вращением и смещением относительно центра координат, так чтобы проекции изображений, фиксируемых веб- и виртуальной камерами, совпадали.

Как известно [4], виртуальную камеру можно задать двумя 4×4-матрицами:

$$A = \begin{bmatrix} a & 0 & f & 0 \\ 0 & b & g & 0 \\ 0 & 0 & c & d \\ 0 & 0 & e & 0 \end{bmatrix}, \quad (1)$$

$$B = \begin{bmatrix} RY_x & RY_y & RY_z & T_x \\ RP_x & RP_y & RP_z & T_y \\ RR_x & RR_y & RR_z & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

Коэффициенты a, b, c, d, e, f, g матрицы A (1) образуют фрустум [1] — пирамидальную видимую область, ограниченную шестью плоскостями, которые строятся исходя из значений угла зрения камеры, отношения высоты виртуального окна к его ширине и значения глубины Z -буфера.

Матрица B (2) состоит из 3×3 -матрицы R вращения вокруг трех осей (RY — yaw, ось направлена вверх; RP — pitch, ось направлена вправо; RR — roll, ось направления) и вектора смещения T вдоль осей x, y, z соответственно.

Проекция вершины на ближнюю плоскость экрана строится в соответствии с формулой

$$\begin{bmatrix} S_x \\ S_y \end{bmatrix} = \begin{bmatrix} a & 0 & f & 0 \\ 0 & b & g & 0 \\ 0 & 0 & c & d \\ 0 & 0 & e & 0 \end{bmatrix} \cdot \begin{bmatrix} RY_x & RY_y & RY_z & T_x \\ RP_x & RP_y & RP_z & T_y \\ RR_x & RR_y & RR_z & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} V_x \\ V_y \\ V_z \\ 1 \end{bmatrix}, \quad (3)$$

которая преобразует трехмерный вектор $[V_x, V_y, V_z, 1]^T$ виртуального пространства в двухмерный вектор $[S_x, S_y]^T$ экранного пространства. При решении задачи локализации камеры 3D- и 2D-векторы в формуле (3) известны, а матрицы A и B , состоящие из внешних и внутренних параметров веб-камеры, неизвестны, эти параметры и необходимо найти.

Матрица B содержит внешние параметры (Extrinsic Parameters) камеры: угол ее поворота и смещение относительно центра сцены. Задача нахождения матрицы B сводится к сопоставлению проекций, фиксируемых реальной и виртуальной камерами [5, 6]. Как только матрица B становится известной, задачу локализации камеры можно считать решенной.

Проблемы реализации модели дополненной реальности. Технические проблемы реализации связаны не только с производительностью компьютера, на котором выполняется приложение, но и с техническими показателями используемой камеры, такими как размер получаемого изображения и его качество. Минимальные требования к вычислительной системе невысоки, программа может быть реализована на мобильных компьютерах или даже на современных мобильных телефонах после соответствующей компиляции для выбранной системы. Относительно требований к камере, следует учитывать, что чем больше и четче область фиксируемого изображения, тем выше точность расчетов. Помимо качества изображения, необходимо чтобы скорость его фиксирования была не менее 30 кадров/с, так как от этого зависит максимальная скорость работы программы в целом.

Некоторые недостатки камеры можно компенсировать, в отличие от размера получаемого изображения и скорости работы используемой камеры, которые изменить невозможно. Например, такой недостаток, как дисторсия (рис. 1), возникающая в оптических системах, где используются выпуклые линзы, можно компенсировать пересчетом координат изображения. Для этого необходимо найти координаты центра дисторсии и ее значение; эти данные, называемые внутренними параметрами камеры (Intrinsic Parameters), определяют отношение идеальных и реальных параметров веб- и виртуальной камер:

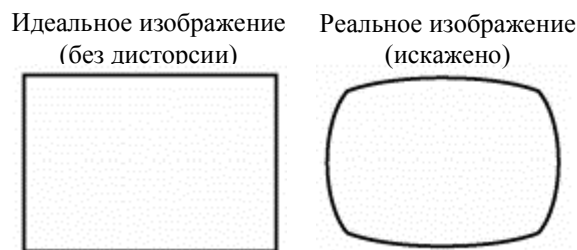


Рис. 1

$$\left. \begin{aligned} d^2 &= (x_i - x_0)^2 + (y_i - y_0)^2; \\ p &= (1 - fd^2); \\ x_d &= p(x_i - x_0) + x_0, \quad y_d = p(y_i - y_0) + y_0, \end{aligned} \right\} \quad (4)$$

здесь f — значение дисторсии; (x_d, y_d) — координаты центра дисторсии; $(x_0, y_0), (x_i, y_i)$ — координаты точек изображения, фиксируемого веб-камерой.

При переводе изображения из реального в идеальное необходимо учесть, что вследствие компенсации дисторсии изображение получается не прямоугольным (рис. 2), и по его краям появляются незаполненные области (которые могут негативно влиять на нахождение маркера).

Чтобы устранить эти области, необходимо изменить масштаб изображения, используя коэффициент масштаба (s).

Тогда система уравнений (4) преобразуется к следующему виду:

$$\left. \begin{aligned} x &= s(x_i - x_0), \quad y = s(y_i - y_0); \\ d^2 &= x^2 + y^2; \\ p &= (1 - fd^2); \\ x_d &= px + x_0, \quad y_d = py + y_0. \end{aligned} \right\} \quad (5)$$

Внутренние параметры камеры (или калибровка) определяются один раз и используются как константа, тогда как внешние параметры изменяются при каждом кадре в зависимости от положения и ориентации камеры.

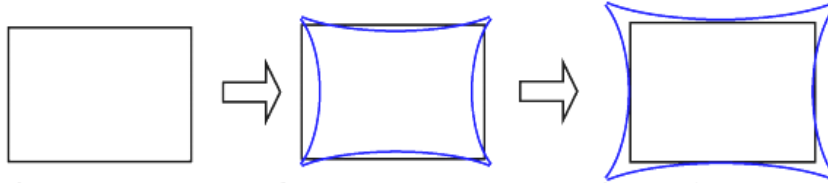


Рис. 2

Алгоритмические проблемы. Главная проблема при моделировании дополненной реальности состоит в извлечении неизвестной информации о пространстве. Проблему недостатка информации можно решить разными способами, например добавлением объекта, параметры которого заранее известны. Сопоставление фиксируемых веб- и виртуальной камерами изображений лучше всего производить по специальным областям, соответствие между которыми известно, такие области называются “feature” — особенность, черта, особая точка. Как правило, это области наибольшего перепада освещенности, углы объектов или специально приготовленные маркеры. Нахождение соответствий по маркерам дает некоторые преимущества, например: известно количество особых точек и ориентация маркера, также известно, что все особые точки компланарны (в идеале). Маркер обычно состоит из четырех особых точек, которые расположены в ярко выраженных углах (например, углы белого квадрата, изображенного на черном квадрате большего размера). Процесс обработки маркеров схематически представлен на рис. 3, где a — маркер, найденный на изображении; b — извлеченная область с необходимой информацией об ориентации маркера и символ для различения нескольких маркеров; $в$ — уменьшенная/увеличенная до установленных размеров копия для оптимального поиска соответствия изображений по базе маркеров.

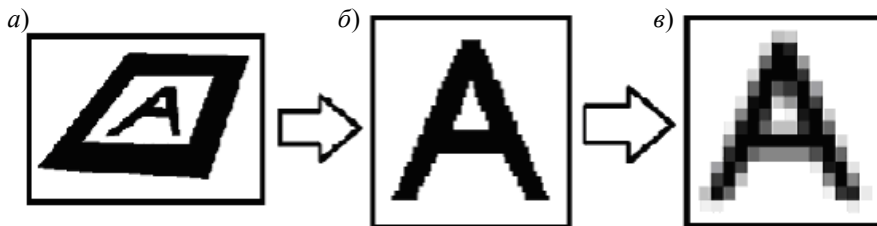


Рис. 3

Решение прикладных задач. Рассмотрим некоторые задачи, для решения которых можно использовать разработанную программу. Например, для оцифровки объектов требуется специализированная аппаратура, в отличие от рассматриваемой системы с применением мобильной веб-камеры. Кроме того, специализированные системы оцифровки, как правило, стационарны и громоздки. С помощью подходов, используемых для построения дополненной реальности, можно реализовать и оцифровку объектов, что позволит устранить недостатки стационарных систем. Систему можно использовать также для измерения длин различных объектов, т.е. как оптическую рулетку, что позволит упростить некоторые затратные по времени измерения, как, например, паспортизация зданий, или измерения, при проведении которых

присутствие человека невозможно. Высокая мобильность системы — ее главное достоинство, вероятные же недостатки зависят от программной реализации.

На примере виртуального музея Рериха [7] с использованием его работ была написана программа, посвященная обработке изображений с целью воздействия на восприятие. Программа может представлять двухмерные изображения в прямой и обратной перспективе, осуществлять нанесение изображения на цилиндрические поверхности с изменяемым радиусом кривизны и на поверхность тора со смещением текстурных координат. В программе предусмотрено изменение освещенности изображений посредством перемещения источника света и изменение общей цветовой гаммы с помощью интерактивной палитры, где можно задать любой из основных цветов и его насыщенность. Также рассматривалась возможность компенсации потери третьей координаты при проецировании на экран с помощью метода анаглифии.

Для построения дополненной реальности также была написана программа, упрощенная структурная схема которой представлена на рис. 4. Программа представляет собой виртуальную сцену со свободной камерой; в системе координат сцены, извлеченной из видеоряда с изображением маркера, отображаются главные оси и положение камеры относительно них. Информации о проекции объекта достаточно для определения направлений осей системы координат, в которой он находится, что позволяет измерять длины различных объектов.

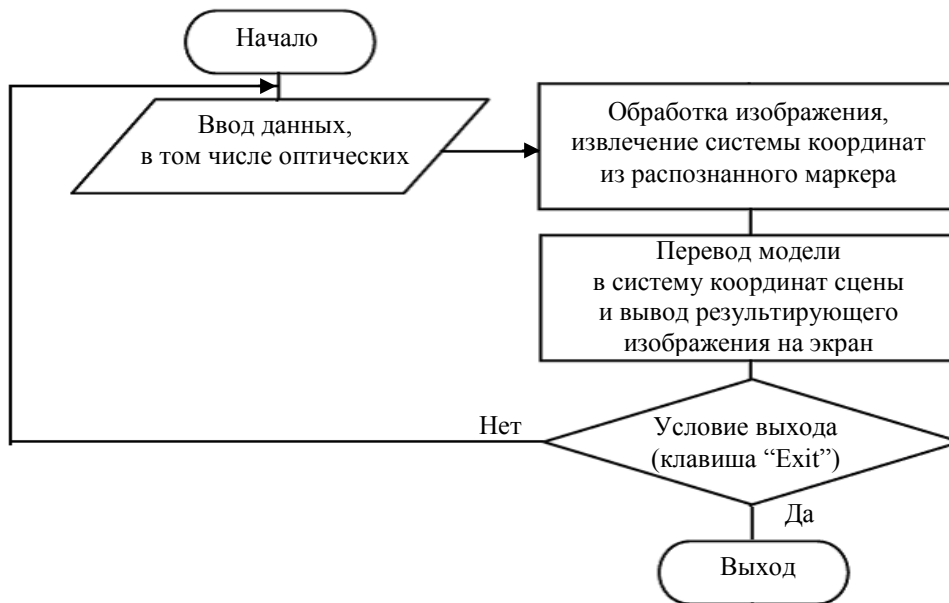


Рис. 4

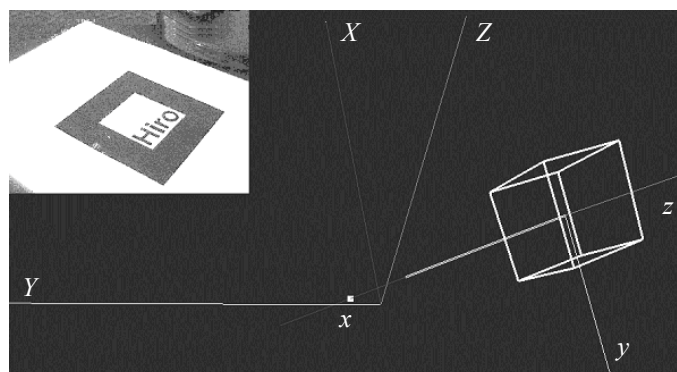


Рис. 5

На рис. 5 представлена система координат сцены, изображенная ортами X, Y, Z ; в левом верхнем углу — изображение, фиксируемое веб-камерой; белый куб с локальной системой координат x, y, z описывает рассчитанное положение и направление веб-камеры относительно сцены; белая точка — пересечение оптической оси камеры с плоскостью сцены.

Выводы. Дальнейшее развитие программы построения дополненной реальности связано с расширением возможностей приложения (и, как следствие, с повышением универсальности). Для расчета системы пространственных координат сцены по произвольному набору точек следует расчеты производить в динамической системе (а не в статике) с использованием алгоритма распознавания особых точек, где анализ ряда изображений и характера движения веб-камеры даст недостающую информацию о сцене.

СПИСОК ЛИТЕРАТУРЫ

1. Википедия, свободная энциклопедия [Электронный ресурс]: <<http://ru.wikipedia.org/wiki>>.
2. Augmented Reality Tool Kit, домашняя страница [Электронный ресурс]: <<http://www.hitl.washington.edu/artoolkit/>>.
3. Ларин М. С. Работа с пакетом программ Open Computer Vision // Науч.-техн. вестн. СПбГУ ИТМО. 2008. Вып. 48. С. 95.
4. OpenGL Projection Matrix, домашняя страница [Электронный ресурс]: <http://www.songho.ca/opengl/gl_projectionmatrix.html>.
5. Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses // IEEE J. of Robotics and Automation. 1987. Vol. RA-3, N 4. P. 323—344.
6. Roger Y. Tsai. An efficient and accurate camera calibration technique for 3D machine vision // Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. Miami Beach, FL, 1986. P. 364—37.
7. Ларин М. С. Среда виртуального пространства // Сб. докл. 8-й сессии Междунар. школы „Фундаментальные и прикладные проблемы надежности и диагностики машин и механизмов“. СПб: СПбГУ ИТМО, 2007. С. 146—149.

Сведения об авторе

Михаил Сергеевич Ларин

— студент; Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кафедра мехатроники;
E-mail: exirity@mail.ru

Рекомендована кафедрой
мехатроники

Поступила в редакцию
15.06.09 г.