УДК 004.946

© 2020 г. А.В. Федцов

(Амурский государственный университет, Благовещенск)

МОДЕЛИРОВАНИЕ ПЕРЕДАЧИ СТЕРЕОИЗОБРАЖЕНИЯ ЧЕРЕЗ ГАРНИТУРЫ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Рассматривается процесс формирования и передачи стереоскопического изображения зрительной системе человека с помощью гарнитур дополненной реальности. Исследуются метрики, влияющие на точность восприятия объема стереоизображения. Описывается численный метод процесса рендеринга стереоскопического изображения, учитывающий параметры зрительной системы человека и устройства отображения дополненной реальности.

Ключевые слова: зрительная система человека, виртуальная реальность, дополненная реальность, рендеринг, стереоскопическое зрение, стереопара.

DOI: 10.22250/isu.2020.65.51-61

Введение

В процессе формирования и передачи стереоскопического изображения зрительной системе человека с помощью HMD-устройств имеет место множество параметров, в том числе индивидуальных для каждого человека, влияющих на точность восприятия метрик виртуального пространства [1]. Имеющиеся на рынке коммерческие HMD-устройства – Oculus Rift, HTC Vive, Sony PS Move и др. – располагают базовыми настройками [2, 3]. Однако по причине привязки программного обеспечения производителя к собственному аппаратному обеспечению, т.е. закрытости исходного кода, чаще всего невозможно адаптировать решения, предложенные производителями, для тонкой настройки процесса рендеринга.

С целью преодоления данного ограничения был разработан авторский численный метод процесса формирования и передачи стереоскопического изображения зрительной системе человека. На базе представленных результатов может быть разработана интеллектуальная система, обладающая следующими возможностями: автоматизированной настройкой рендеринга под любое HDM-

устройство и особенности зрительной системы пользователя; интеграцией виртуальной и дополненной реальности на любых гарнитурах НМД; оптимизацией вычислительных затрат посредством фовеального рендеринга [4]; оригинальным интерфейсом управления компьютером с помощью движений зрачка глаза [5]; опцией трехмерной реконструкции сцены на основе стереоизображения [6].

Алгоритм передачи стереоскопического изображения

Для проведения исследования было использовано несколько устройств НМД. В данной работе описывается исследование, проведенное на гарнитуре дополненной реальности Epson Moverio BT30C. Эта гарнитура подключается к компьютеру по интерфейсу HDMI. Передача стереоизображения в гарнитуре происходит в формате горизонтальной стереопары – два изображения для левого и правого глаза передаются в одном кадре (слева направо), сжатые по вертикали на 50% (рис. 1).

Рис. 1. Горизонтальная стереопара.

В качестве программного обеспечения использовалась межплатформенная среда разработки Unity. При проведении исследования были учтены современные аспекты разработки приложений виртуальной и дополненной реальности [7 – 10], настроено окружение для разработки на базе Unity [11 – 14], создана тестовая сцена, настроен вывод стереоскопического изображения в режиме стереопары и поворот вектора камеры с использованием датчика вращения.

Сцена формировалась в трехмерной системе координат Unity таким образом, чтобы расстояние до размещенных тестовых объектов можно было сравнить с реальным расстоянием в физическом мире и, таким образом, субъективно определить точность передаваемого изображения. Было размещено три предмета (куба) с длинами сторон 20см, 40см и 50 см, на расстояниях 30 см, 1 см и 2 м соответственно, с небольшим наклоном для лучшего восприятия объема (рис. 2).

Рис. 2. Скриншот тестовой трехмерной сцены, моделируемой в Unity.

Передаваемое на гарнитуру Epson Moverio стереоизображение воспринимается зрительной системой человека как объемное изображение поверх картины физического мира, – следовательно, можно субъективно определить расстояние до виртуального объекта и сопоставить его с физическим миром. Кроме того, в гарнитуре имеются датчики, позволяющие измерить наклон по трем осям, т.е. определить вектор направления гарнитуры в трехмерном пространстве.

Критерием идеально подобранных параметров формирования объемного изображения станет полное совпадение воспринимаемых размеров генерируемого в Unity трехмерного пространства виртуальной сцены с физическим миром. При вращении головы в физическом мире виртуальное изображение должно вращаться соответственно, чтобы виртуальные объекты воспринимались привязанными к точке в физическом мире. Исследуемые параметры могут отличаться для каждого человека с учетом особенностей зрительной системы и размеров головы.

Для формирования изображения в формате стереопары использовались два компонента Camera, входящие в программную среду Unity. При этом точки области видимости для каждой камеры были смещены на расстояние, примерно равное межзрачковому расстоянию (рис. 3).

Рис. 3. Скриншот исходного изображения стереопары, моделируемой в Unity.

Таким образом, появившееся смещение позволило зрительной системе воспринимать картинку объемной.

Для считывания показаний с датчиков гарнитуры Epson Moverio был использован официальный SDK для платформы Windows. SDK обрабатывает данные со следующих датчиков – цифровой компас, гироскоп, акселерометр и возвращает трехмерный вектор положения гарнитуры в пространстве. Схема (рис. 4) показывает, что датчик положения гарнитуры смещен относительно точки вращения головы, которую принято считать центром вращения.

Рис. 4. Схема положения датчика наклона относительно центра головы, глаз и дисплеев гарнитуры.

Для точного воспроизведения наклона виртуального изображения необходимо трансформировать координаты с учетом смещения вектора наклона, положения глаз зрительной системы и дисплеев гарнитуры относительно центра вращения головы.

Полученные параметрические данные

В процессе исследования эмпирически были обнаружены следующие метрики, непосредственно влияющие на точность восприятия стереоскопического изображения зрительной системой человека:

- межзрачковое расстояние;
- области видимости виртуальной камеры;
- точка вращения и наклон виртуальной камеры по горизонтали относительно точки вращения;
- коэффициенты смещения вектора датчика наклона гарнитуры по каждой оси;

смещение точки вращения относительно вектора датчика наклона гарнитуры.

При этом неправильно подобранное значение параметра межзрачковое расстояние вызывало двоение виртуального изображения, не получалось сфокусироваться на нем. Решением стало изменение расстояния между компонентами Camera (межзрачкового расстояния) опытным путем, чтобы добиться комфортного восприятия глубины. В свою очередь неправильное значение области видимости виртуальной камеры вызывало изменение размера воспринимаемых виртуальных объектов. Например, размер виртуального изображения мог быть слишком мал по сравнению с реальным миром. Предметы с размером 40 см в системе координат Unity воспринимались меньше в 2-3 раза. Решением стало изменение параметра размера области видимости (Field of View) компонента Camera в Unity. Опытным путем были подобраны значения, позволяющие увеличить воспринимаемый размер виртуального изображения до размеров, соответствующих действительности. Для проверки использовалась линейка, размещенная на таком же расстоянии от глаз, как виртуальный объект в Unity.

Неправильный подбор точки вращения и наклона виртуальной камеры по горизонтали вызывает искажение глубинного восприятия. Например, на расстоянии 30 см от камеры виртуальный объект воспринимался с размерами, аналогичными реальному миру, однако при отдалении объекта на 1 м в системе координат Unity воспринимаемый объект отдалялся примерно на 40 см и уменьшался в размере в 2-3 раза. Опытным путем было установлено, что для изменения пространственного восприятия нужно изменить наклон (произвести вращение) компонентов Camera по оси Y для каждого глаза в Unity, изменив при этом так же точку вращения (pivot), которая по умолчанию находится в центре объекта.

Неправильно подобранное значение коэффициента смещения вектора датчика наклона гарнитуры по каждой оси вызывало отклонение угла вращения головы в физическом мире от угла вращения виртуальной сцены. Например, при повороте головы на 10 градусов вправо виртуальная сцена совершала поворот на 90 градусов по горизонтали. При наклоне головы вниз виртуальная сцена наклонялась вверх, таким образом, ось вращения была инвертирована. Опытным путем были подобраны коэффициенты для каждой оси наклона d_x , d_y и d_z так, что при вращении головы объекты виртуальной сцены воспринимались неподвижными относительно физического пространства. Так как SDK гарнитуры передает не сам вектор вращения, то во изменение вектора вращения необходимо умножить каждое измерение на разницу во времени между текущим и прошлым измерением, чтобы узнать текущее значение вектора наклона:

$$x_i = d_i d_t + x_{i-1},$$

здесь x_i – текущий угол наклона; d_i – текущее измерение датчика наклона; d_i – разница по времени между текущим и предыдущим измерением; x_{i-1} – угол наклона после предыдущего измерения. Если не учитывать время d_i при расчете вектора вращения, а складывать d_i с x_{i-1} в каждом кадре, то вектор наклона будет передаваться с потерями, а сам наклон восприниматься рывками, поскольку время подготовки каждого нового кадра отличается от предыдущего.

Неправильно подобранное значение смещения точки вращения относительно вектора датчика наклона гарнитуры вызывало движение виртуальных объектов при вращении головы в физическом мире. Установив опытным путем расстояние от датчика вращения гарнитуры до центра вращения головы согласно схеме (см. рис. 4), виртуальные объекты воспринимали неподвижными и привязанными к точке физического пространства при вращении головы в физическом мире.

Предлагаемый численный метод

Для математического описания процесса рендеринга стереоскопического изображения воспользуемся уравнениями аффинных преобразований координатных моделей, традиционно применяемых в компьютерной графике [15]:

$$X = Ax + By + Cz + D;$$

$$Y = Ex + Fy + Gz + H;$$

$$Z = Ix + Jy + Kz + L,$$

здесь x, y, z – координаты точки до преобразования; X, Y, Z – координаты точки после преобразования; $A, B, C \dots L$ – числовые коэффициенты. Для того, чтобы учесть константы D, H и F , необходимо перейти к однородным координатам и добавить строку с единицами в матрице [16, 17]. Данное преобразование может быть представлено в матричном виде:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} A & B & C & D \\ E & F & G & H \\ I & J & K & L \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

В свою очередь, при параметрическом моделировании рендеринга необходимо использовать общеизвестные типовые преобразования (рис. 5).

Вращение по оси x на угол α :

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

Рис. 5. Типовые преобразования проекционных координатных моделей:
а) поворот системы координат; б) перспективное преобразование.

Вращение по оси y на угол β , а также по оси z на угол γ :

$$\begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = \begin{vmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \\ 1 \end{vmatrix}; \quad \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = \begin{vmatrix} \cos \gamma & \sin \gamma & 0 & 0 \\ -\sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \\ 1 \end{vmatrix}.$$

Перспективное преобразование трехмерных координат в двумерные:

$$\begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = \begin{vmatrix} \frac{z_k - z_{пл}}{z_k - z} & 0 & 0 & 0 \\ 0 & \frac{z_k - z_{пл}}{z_k - z} & 0 & 0 \\ 0 & 0 & 1 & -z_{пл} \\ 0 & 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \\ 1 \end{vmatrix}.$$

Сдвиг точки:

$$\begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \\ 1 \end{vmatrix}.$$

Введем обозначения, используемые при моделировании рендеринга, преобразуемого с учетом выявленных метрик. Во-первых, это параметры, динамически изменяемые в процессе рендеринга: P – текущее положение стереокамеры в пространстве; R – вектор вращения датчика наклона. Во-вторых, параметры гарнитуры: dP – смещение датчика ее вращения относительно положения стереокамеры по осям x и z ; k – коэффициент поправки значений вектора вращения; S – размер дисплея гарнитуры. В-третьих, параметры, зависящие от особенностей зритель-

ной системы человека: m – межзрачковое расстояние (расстояние между центрами вращения камер левого и правого глаза по оси x); a – угол наклона камеры для глаза; f – область видимости камеры; e – смещение центра вращения камер левого и правого глаза относительно центра стереокамеры P по оси z .

Таким образом, если считать точки положений камеры в пространстве, датчика вращения и дисплеев гарнитуры лежащими на одной плоскости согласно схеме (см. рис. 4), то координата y данных положений будет совпадать и смещения по оси y (по вертикали) для них можно не учитывать.

Вычисление центра вращения стереокамеры с учетом смещения стереокамеры в пространстве P и смещения датчика вращения гарнитуры dP имеет вид:

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -P_x - dP_x \\ 0 & 1 & 0 & -P_y \\ 0 & 0 & 1 & -P_z - dP_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

Вращение стереокамеры по оси x на основе вектора вращения датчика наклона R и коэффициента поправки значений вектора вращения k :

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos R_x k & \sin R_x k & 0 \\ 0 & -\sin R_x k & \cos R_x k & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}.$$

Аналогично преобразование вращения стереокамеры по осям y и z . Вычисление центра вращения камеры левого глаза с учетом смещения датчика вращения гарнитуры dP относительно центра стереокамеры P , смещение центра вращения камер левого и правого глаза e относительно центра стереокамеры P по оси z и межзрачкового расстояния m :

$$\begin{pmatrix} X'_l \\ Y'_l \\ Z'_l \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & m/2 + dP_x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -e + dP_x \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}.$$

Аналогичное преобразование выполняется для камеры правого глаза. Вращение камеры левого глаза по оси x на угол α :

$$\begin{pmatrix} X'_l \\ Y'_l \\ Z'_l \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} X'_l \\ Y'_l \\ Z'_l \\ 1 \end{pmatrix}.$$

Аналогичное преобразование выполняется для камеры правого глаза. Далее проводится расчет расстояния плоскости проекции. Область видимости камеры представляет собой коэффициент отношения размеров дисплея к расстоянию до плоскости проекции:

$$Z_{пл} = f S.$$

Таким образом, при одинаковом значении параметра f на дисплеях с разными размерами будет отображено одинаковое количество информации. Поскольку мы преобразовали координаты относительно камеры, то $z_k = 0$.

Проекция камеры левого глаза:

$$\begin{pmatrix} X_l \\ Y_l \\ Z_l \\ 1 \end{pmatrix} = \begin{pmatrix} -fS_x & 0 & 0 & 0 \\ -Z'_l & -fS_y & 0 & 0 \\ 0 & -Z'_l & 1 & -fS_x \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} X'_l \\ Y'_l \\ Z'_l \\ 1 \end{pmatrix}.$$

Проекция камеры правого глаза находится аналогично.

И, наконец, финальным преобразованием станет масштабирование полученной проекции под размер дисплея:

$$\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \begin{pmatrix} 1/S_x & 0 & 0 \\ 0 & 1/S_y & 0 \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

В результате применения описанных преобразований с настройкой индивидуальных параметров зрительной системы человека и гарнитуры полученное в стереоскопическое изображение трехмерной сцены будет точно восприниматься зрительной системой человека.

Заключение

Подводя итог описываемой части проводимого исследования, отметим, что в его рамках на текущий момент эмпирически выявлены базовые метрики формирования и передачи стереоскопического изображения зрительной системе человека, реализуемого посредством HMD-устройств. Кроме того, формализованы и описаны субъективные отклонения в точности восприятия, вызываемые при неправильном выборе значений найденных метрик, рассмотрены практические нюансы считывания данных с датчика наклона гарнитуры.

Созданный численный метод, а также дальнейшие исследования позволят создать качественно новую математическую модель, описывающую процесс стереозрения, – в частности, более гибкую по сравнению с существующими [18]. Это

подчеркивает научную новизну исследования. Программная реализация рендера на основе данной модели даст практически значимые результаты: возможность проводить тонкую настройку под любое HDM-устройство и индивидуальные особенности зрительной системы; повысить качество формируемого изображения; оптимизировать вычислительные затраты; проводить трехмерную реконструкцию сцены; реализовать оригинальный интерфейс управления компьютером с помощью движений зрачка глаза.

Таким образом, следующими шагами должны стать исследования метрик процесса для гарнитур различных конфигураций, в том числе гарнитур виртуальной реальности с искривляющими линзами Френеля [19, 20]; доработка и оптимизация соответствующей вычислительной технологии; практическая реализация программного обеспечения, использующего предлагаемый численный метод.

ЛИТЕРАТУРА

1. Федцов А.В. Исследование аспектов формирования объемного изображения в устройствах дополненной реальности и их влияния на точность восприятия объема зрительной системой человека // Химия, физика, биология, математика: теоретические и прикладные исследования: сб. ст. по мат. XXXVII междунар. науч.-практ. конф. – 2020. – № 6(26). – С. 11-15.
2. Иванова А.В. Технологии виртуальной и дополненной реальности. Возможности и препятствия применения. – М.: Книга по требованию, 2018.
3. Wetzstein G. A personalized VR/AR system that adapts to the user is crucial to deliver the best possible experience // The BRIDGE – 2016. – V.46, № 4. – P. 1034-1048.
4. Rachel A., Anjul P., David L., Joohwan K. Latency Requirements for Foveated Rendering in Virtual Reality // ACM Trans. Appl. Percept. – 2017. – V.14, № 25. – P. 7-25.
5. Федцов А.В., Семичевская Н.П. Моделирование вычислительных процессов для решения задач распознавания положения зрачка глаза человека на видеопотоке с использованием технологии параллельных вычислений // Экспериментальные и теоретические исследования в современной науке: сб. ст. по мат. XXXVIII междунар. науч.-практ. конф. – 2019. – № 8(34). – С. 11-15.
6. Sisodia A., Bayer M., Townley-Smith P., Nash B. Advanced Helmet Mounted Display // Proceedings of SPIE - The International Society for Optical Engineering. – 2007 – V. 6557. – P. 799-810.
7. Feng Y. Research on Seamless Fusion of Real and Virtual Scenes in Augmented Reality System. – Shanghai University, 2007.
8. MacKenzie K.J., Dickson R.A., Watt S.J. Vergence and accommodation to multiple-image-plane stereoscopic displays: real world responses with practical image-plane separations // Journal of Electronic Imaging. – 2012. – V. 21, № 1. – P. 8288.
9. Narain R., Rachel A. A., Abdullah B., Ward G. J. Optimal Presentation of Imagery with Focus Cues on Multi-plane Displays. // ACM Trans. Graph. – 2015. – V. 34, № 4. – P. 59:1-59:12.
10. Zhdanov D.D., Potemin I.S., Kishalov A.A., Zhdanov A.D. Stochastic ray tracing methods in problems of photorealistic image synthesis for augmented reality systems // Proc. of 27th International Conference on Computer Graphics and Machine Vision. – 2018. – № 27. – P. 42-46.
11. Hocking J. Unity in Action. Multiplatform game development in C# with Unity 5. – 2015.
12. Thorn A. Mastering Unity Scripting. – 2015.

13. *Thorn A.* How to Cheat in Unity 5: Tips and Tricks for Game Development, 2015.
14. *Dickinson C.* Unity 5 Game Optimization, 2015.
15. *Пореев В.* Компьютерная графика: Учебное пособие. – СПб.: БХВ-Петербург, 2005.
16. *Мусхелишвили Н.И.* Курс аналитической геометрии – М.: Высшая школа, 1967.
17. *Смирнов В.И.* Курс высшей математики. – М. – Л.: ОГИЗ – Гостехиздат, 1948.
18. *Киреев О.С.* Математические модели стереозрения // Актуальные направления научных исследований: от теории к практике: материалы IV междунар. науч.-практ. конф. – 2015. – С. 121-123.
19. *Tapani L.* Diffractive optics for virtual displays. – Society for Image Display, 2006.
20. *Kurt A., Watt S.J., Girshick A.R., Banks M.S.* A Stereo Display Prototype with Multiple Focal Distances. // ACM Trans. Graph. – 2004. – V. 23, № 3. – P. 804-817.

Статья представлена к публикации членом редколлегии И.Е. Ереминым.

E-mail:

Федцов Алексей Владимирович – alexeyfedtsov@mail.ru.