

О СИСТЕМАХ ПЕРСПЕКТИВ ДЛЯ КОМПЬЮТЕРНОЙ ГРАФИКИ

А. М. Ковалев¹, В. Е. Лищенко², М. В. Степанов²¹*Конструкторско-технологический институт научного приборостроения СО РАН,
г. Новосибирск*²*Институт автоматизации и электрометрии СО РАН, г. Новосибирск
E-mail: vlish@ngs.ru*

Рассматриваются концепции использования в компьютерной графике новых систем перспектив, открывающих возможности отображения трехмерных визуальных сцен, приближенных к естественному зрительному восприятию.

Введение. Компьютерная графика достигла больших успехов в виртуализации моделируемой среды. Современные графические системы способны в реальном времени визуализировать большие пространства с высокой степенью детализации. Стало возможным правдоподобное изображение различных природных феноменов и сложных динамических объектов: облачного покрова, деревьев, травы, человека с его движениями и мимикой. Успешно имитируются оптические эффекты – отражение и преломление света. Этому прогрессу содействовали достижения в микроэлектронике: высокая производительность современных процессоров, большой объем памяти, вычислительные конвейеры в графических акселераторах, реализующие матричные операции и обработку элементов изображения.

В компьютерной графике до сих пор применяется проективное преобразование, свойственное геометрической оптике, т. е. используется ренессансная перспектива, дошедшая до нас от Леонардо да Винчи (1452–1519 годы). Такая система перспективы подразумевается во всех широко распространенных программных интерфейсах приложений (API – Application Program Interface) для трехмерной графики, например OpenGL и Direct3D, а также в аппаратном обеспечении компьютеров – видеоадаптерах.

Строго говоря, ренессансная перспектива не соответствует естественному зрительному восприятию человека. Более того, как показано в [1], безошибочно изобразить пространство на плоскости принципиально невозможно. Для решения проблемы Раушенбах предлагает использовать набор систем перспектив, в котором ренессансная рассматривается как частный случай. Достоинством системы перспектив Раушенбаха является приближенность к естественному зрительному восприятию, однако у нее есть ряд недостатков. Полученные зависимости эмпиричны, их трудно реализовать на практике в виде программного обеспечения трехмерной графики. Кроме того, в его системах есть ограничения, которые в итоге приводят к отсутствию общности.

В настоящее время появились дополнительные знания о процессах зрительного восприятия, о геометрической структуре визуального пространства в сознании человека [2–5]. Компьютерная графика оснащается трехмерными (3М) дисплеями, в том числе стереоскопическими многокурными [6–8]. Разрабатываются новые дисплеи на основе объемных сред и голографического метода [9–11]. Таким образом, представляется актуальной и возможной разработка более общей теории перспектив, доступной для реализации в современной компьютерной графике.

Цель данной работы – обобщение проективных преобразований для трехмерной компьютерной графики, учитывающее закономерности зрительного восприятия и возможности существующих и создаваемых дисплеев.

Визуальное пространство. Современные представления о зрительном восприятии позволяют утверждать, что формирование визуального пространства (ВП) в сознании человека может быть описано как результат ряда процессов: физических, физиологических, психофизических и когнитивных [2]. Процесс видения можно условно разделить на следующие этапы:

- 1) сбор информации сенсорной подсистемой зрения;
- 2) построение визуального пространственного образа;
- 3) формирование собственно ВП.

На первом этапе путем сканирования пространства предметов остронаправленным взглядом [3], во-первых, строится двумерный (2М) сетчаточный образ в ренессансной перспективе. Угловой размер предмета на сетчатке в пределах малого рабочего поля определяется как $\beta = S/z$, где S – натуральный размер предмета, z – расстояние до предмета. Во-вторых, оцениваются расстояния z по признакам глубины предметов – прямым и косвенным. Прямые признаки глубины стимулируют аккомодацию глаз, их конвергенцию и стереоскопическую диспаратность. К косвенным признакам относят: окклюзии и автоокклюзии; уменьшение размеров и смещение к горизонту предметов при их удалении; уменьшение контраста и градиента текстуры при увеличении расстояния; двигательный параллакс. Практика показывает, что сенсорная подсистема дает верную информацию о расстояниях до предметов, и притом с необходимой для зрительной системы точностью. Функции воспринимаемых расстояний при оценке дальности и размеров предметов даны в [4].

На втором этапе построения ВП сетчаточный образ согласно психофизическому закону Эммерта испытывает «растяжение» по размеру и глубине пропорционально воспринимаемым расстояниям до предметов. Как показано в [3], эти расстояния определяются оптической силой редуцированного аккомодирующего глаза, представленного в виде тонкой линзы

$$1/z' = (1/z) + (1/d), \text{ или } A' = A + \delta. \quad (1)$$

Здесь z' – воспринимаемое, или фокусное, расстояние; d – гиперфокальное расстояние; A' – оптическая сила; A – расстояние аккомодации; δ – аккомодация покоя. В первом выражении расстояния заданы в метрах, во втором – в диоптриях. Из (1) $z' = z/(1 + \delta z)$, и поэтому «растянутый» размер предмета $S' = \beta z' = S/(1 + \delta z)$. Если учесть вращения глаз при фиксации предметов в разных точках пространства, то можно получить проективное преобразование вида

$$\mathbf{P}' = \mathbf{P}/(1 + \delta r), \quad (2)$$

где $\mathbf{P}' = (x' y' z')$ – радиус-вектор точки изображения; $\mathbf{P} = (x y z)$ – радиус-вектор точки физического пространства; $r = \sqrt{x^2 + y^2 + z^2}$ – расстояние до точки \mathbf{P} от центра проекции – начала системы координат.

После растяжения сетчаточного образа 2М-изображение становится трехмерным. При бинокулярном зрении возникает стереопсис, который фузирует (сплавляет) два 3М-изображения в одно объемное стереоскопическое изображение (рис. 1). И так, на втором этапе при помощи (2) строится пространственный образ окружающей среды, или проективная сенсорная модель ВП [3]. На третьем этапе формируется собственно ВП, внутренняя геометрия которого соответствует гиперболической геометрии Лобачевского – Гильберта [4, 5].

Дисплеи. Из рассмотренного процесса видения следует, что для отображения 3М-пространства необходимы 3М-дисплеи. Они должны воспроизводить не только точки с заданными угловыми координатами, но и обеспечивать стимулы аккомодации и конвергенции глаз для воссоздания глубины точек и инициации бинокулярного стереосинтеза, или стереопсиса.

Зрительная система как средство восприятия визуальной информации обладает уникальными «техническими» характеристиками. Поле зрения с учетом движения глаз охватывает телесный угол, близкий к 2π . Разрешающая способность 1 угл. мин. Объем аккомодации, или фокусировки, равен 4–5 дптр, что соответствует диапазону расстояний от 20–25 см до ∞ . Глубина аккомодации, т. е. аксиальная протяженность фокуса, составляет $\pm(0,15–0,30)$ дптр в зависимости от диаметра зрачка глаза. Объем конвергенции, или максимальная диспаратность, достигает 14–18°. Разрешение стереоскопического зрения от 2 до 10 угл. с.

К сожалению, в мировой практике не существует средств отображения визуальной информации, отвечающих в полной мере возможностям естественного зрительного восприятия. Все известные дисплеи можно отнести к трем группам:

- 1) однофокусные дисплеи на базе двумерного экрана;
- 2) многофокусные дисплеи на основе трехмерной отображающей среды;
- 3) голографические дисплеи.

Первая группа, широко распространенная на практике, содержит дисплеи трех типов: двумерные, стереоскопические и автостереоскопические.

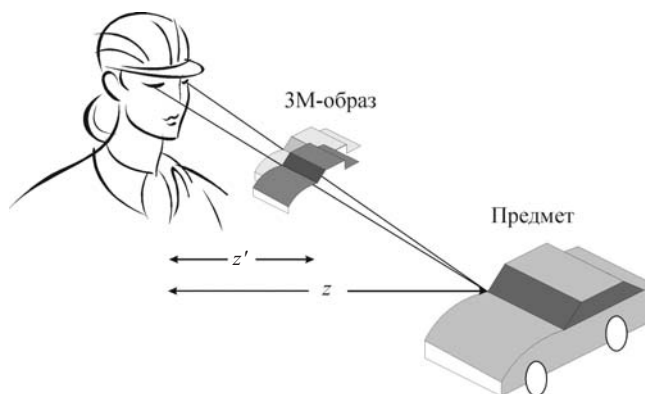


Рис. 1

К 2М-дисплеям относятся: мониторы компьютеров, телевизоры, проекционные системы с отражающими и просветными экранами (плоскими, сферическими или цилиндрическими), наголовные монокулярные дисплеи и т. д. Такие дисплеи способны обеспечить объем аккомодации, равный глубине фокуса, но не создают условий для конвергенции глаз.

Стереоскопические 3М-дисплеи используют два двумерных изображения, левое и правое (стереопару), и очковый принцип сепарации изображений. Стереопара может быть получена или на одном экране, или на двух [6]. В первом случае применяется временная сепарация, во втором – поляризационная. Деление монохромных стереопар выполняется методом цветных анаглифов (фильтров). Наголовные стереодисплеи оснащаются двумя позиционно разделенными экранами и окулярами.

Автостереоскопические 3М-дисплеи в идейном смысле восходят к интегральной фотографии Липпмана (1908 г.). На одном плоском экране может размещаться N изображений – ракурсов сцены (например, восемь ракурсов в работе [7]). Применяется безочковая растровая сепарация, обслуживающая несколько наблюдателей. Примеры таких дисплеев даны в [8].

Общим недостатком однофокусных дисплеев является малая глубина отображаемого пространства. Объем аккомодации равен глубине фокуса $\pm(0,15-0,30)$ дптр. В стереодисплеях объем конвергенции, соответствующий объему аккомодации, определяется диспаратностью $\pm(35-70)$ угл. мин. Превышение этого объема приводит к визуальному дискомфорту, что подтверждается экспериментально [12]. Глубина пространства, воспринимаемая без бинокулярного стресса, зависит от фокусного расстояния, или удаления наблюдателя от экрана, что отражено в таблице. Размер экрана соответствует полю зрения $40 \times 30^\circ$. Глубина фокуса $\pm 0,25$ дптр. Относительная глубина равна произведению глубины в диоптриях и минимального расстояния до изображения в метрах. Из таблицы видно, что имеет смысл увеличивать размеры экранов и наблюдать изображения с большого расстояния. Однако при этом теряется ближняя зона пространства.

Вторая группа объединяет дисплеи с 3М-отображающей средой, в которой формируются элементы изображения – воксели. Объемный «экран» может быть свободным от вещества или заполнен газом, жидкостью или твердым телом [9]. В 2004 г. фирма “LightSpace Technologies Inc.” впервые продемонстрировала твердотельный объемный дисплей DepthCube Z1024, содержащий пакет из 20 параллельных светорассеивающих пластин-модуляторов, управляемых оптоэлектронным способом [10]. Отображающая среда представляет собой параллелепипед размером $40 \times 30 \times 10$ см (ширина, высота, глубина соответственно) и содержит $15,3 \cdot 10^6$ цветных вокселов. В [9] предлагается объемный экран в виде куба с ребром 30 см, который вмещает

Расстояние до экрана, м	Размер экрана, м	Глубина, м	Относительная глубина
0,5	0,36 × 0,27	0,44 ... 0,57	0,22
1,0	0,72 × 0,54	0,80 ... 1,33	0,40
2,0	1,44 × 1,08	1,33 ... 4,00	0,67
4,0	2,88 × 2,16	2,00 ... ∞	1,00

пакет из 100 пластин. Преимущество дисплеев состоит в том, что воксели являются одновременно стимулами аккомодации и конвергенции для множества наблюдателей. Преимущество для одних задач и недостаток для других – изображаемые предметы становятся прозрачными. К недостаткам следует отнести малый объем глубины, определяемый физическими размерами отображающей среды (10–30 см).

Голографические дисплеи на базе акустооптических пластин [11] находятся на той стадии разработки, когда обсуждать конкретные характеристики рано. Достоинством голографического метода является регистрация множества ракурсов сцены для множества наблюдателей. Переход от одного ракурса к другому происходит непрерывно. Однако вряд ли такие дисплеи способны отобразить глубокое пространство. Голограммы, генерируемые компьютером, обеспечивали глубину 20 см [8].

Мы не рассматривали все параметры дисплеев: поле зрения, разрешающую способность, цвет, яркость и т. д. Нас интересовало главное – возможность дисплеев создавать прямые признаки глубины, необходимые для построения сенсорной модели, а затем и визуального пространства в сознании человека. К сожалению, в отношении этих параметров возможности известных дисплеев оказались ограниченными.

Системы перспектив. Пусть в декартовой системе координат, начало которой совпадает с оптическим центром глаза O , а ось OZ – с главной зрительной осью, задан дисплей с полем зрения $w \times h^\circ$ и глубиной $\zeta_D \geq 0$ дптр (рис. 2). На дисплее отображается область евклидова пространства глубиной $\zeta_E = z_{\text{NEAR}}^{-1} - z_{\text{FAR}}^{-1}$, причем $0 < z_{\text{NEAR}} \leq z \leq z_{\text{FAR}} \leq \infty$. Пусть в отображаемом пространстве на расстоянии z_0 расположена виртуальная картинная плоскость K_E . Соответственная плоскость K_D в дисплейном пространстве сдвинута относительно K_E на расстояние Δ дптр. Требуется найти системы перспектив с учетом психофизических особенностей зрительного восприятия пространства предметов.

Вначале рассмотрим идеальный с геометрической точки зрения 3М-дисплей. Пусть $\mathbf{P} = (xyz)$ – радиус-вектор точки евклидова пространства. При малых полях зрения ($\cos w$ и $\cos h$ близки к 1) проективное преобразование

$$\mathbf{P}' = \mathbf{P}/(1 + \delta z) \quad (3)$$

переводит отображаемое пространство в линейную модель ВП [3], близкую к проективной сенсорной модели (2). Исключая из (3) δ – параметр зритель-

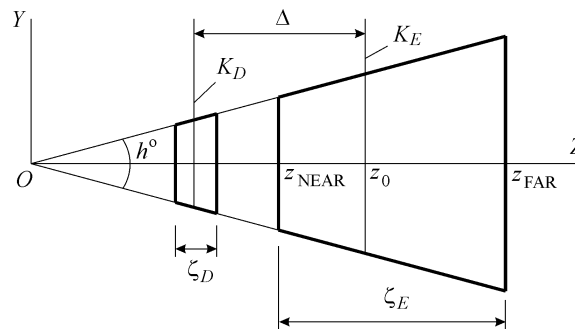


Рис. 2

ной системы, получим идеальный дисплей, у которого $\mathbf{P1}'|_{\delta=0} = \mathbf{P1} \equiv \mathbf{P}$, $\zeta_E = \zeta_D$ и $\Delta = 0$. При совмещении областей двух пространств – евклидова и дисплейного – выполняются, безусловно, все критерии соответствия изображений естественному восприятию [13], а именно: сохранение подобия предметов; сохранение интерпозиции предметов; сохранение размеров, или масштабов, разноудаленных предметов; достоверная передача удаленности предметов. Итак, любой 3М-дисплей, пространство которого физически или виртуально совмещено с отображаемым, может считаться идеальным. Устройства для отображения интерьеров и открытой протяженной местности, близкие к идеальным, – это наголовные стереодисплеи или стерео- и авто-стереодисплеи с большим экраном (см. последнюю строку таблицы).

Пусть $\zeta_E > \zeta_D \geq 0$, а $\Delta > 0$. Например, дисплей в пространстве от 40 до 50 см (2,5–2,0 дптр) глубиной $\zeta_D = 0,5$ дптр изображает область евклидова пространства глубиной $\zeta_E = 0,9$ дптр в диапазоне от 1 до 10 м (1,0–0,1 дптр), как показано на рис. 3 (расстояния даны в метрах). Наклонная прямая обозначает границу зрения. Дисплей с визуальным комфортом воспроизводит уменьшенную копию пространства, отличается от идеального, и упомянутые критерии соответствия, как правило, нарушаются.

Выворачиваем глубину пространств путем сжатия евклидова пространства к картинной плоскости K_E , расположенной на расстоянии z_0 от наблюдате-

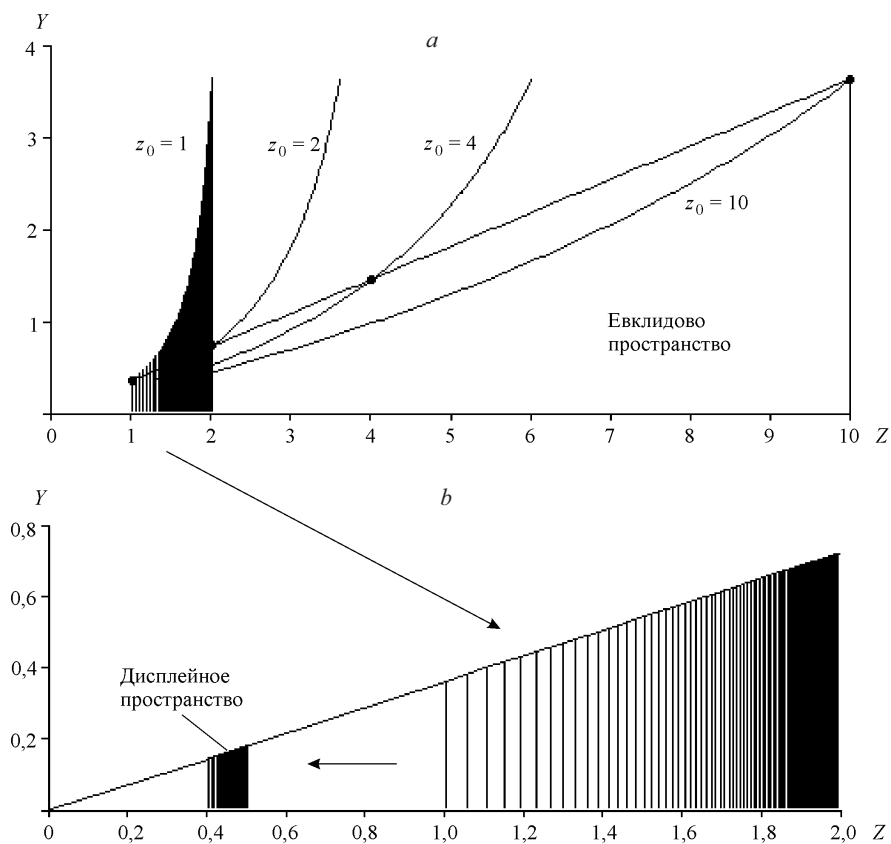


Рис. 3

ля, с коэффициентом сжатия $k = \zeta_D / \zeta_E$. Аппликата точек сжатого пространства в диоптрийном исчислении расстояний равна $z_k^{-1} = kz^{-1} + (1-k)z_0^{-1}$. При таком сжатии нарушается интерпозиция и удаленность предметов, но сохраняется их масштабное соотношение и подобие (рис. 3, а). Преобразование вида

$$\mathbf{P2} = \mathbf{P} \frac{z_0}{kz_0 + (1-k)z}, \quad (4)$$

где $z_2 = z_k$, восстанавливает интерпозицию предметов. Результат преобразования для $z_0 = 1$ м обозначен стрелкой от рис. 3, а к рис. 3, б. Чтобы далее перейти к дисплейному пространству, необходимо выполнить преобразование, подобное (3):

$$\mathbf{P3} = \mathbf{P2} / (1 + \Delta z_k) = \mathbf{P} \frac{z_0}{kz_0 + (1-k + \Delta z_0)z}. \quad (5)$$

Заметим, что для плоского дисплея $\zeta_D = 0$. Поэтому, подставляя $k = 0$ в (4) и (5), получим преобразования $\mathbf{P2} = \mathbf{P}z_0/z$ и $\mathbf{P3} = \mathbf{P}z_{D0}/z$ для ренессансной перспективы на плоскости, где $z_{D0} = z_0/(1 + \Delta z_0)$ – расстояние до дисплейной картинной плоскости K_D . Параметры зрительной системы не входят в формулы преобразований. Поэтому можно утверждать, что проективное преобразование (5) является обобщением ренессансной перспективы для 3М-дисплеев.

Искажения масштабов и дальности, характерные для ренессансной перспективы [1], можно исправить с учетом (3) проективным преобразованием вида

$$x4 = \frac{xm_X}{1 + \delta_X z}; \quad y4 = \frac{ym_Y}{1 + \delta_Y z}; \quad z4 = \frac{zz_0}{kz_0 + (1-k + \Delta z_0)z}, \quad (6)$$

где $(x4 y4 z4)$ – радиус-вектор $\mathbf{P4}$ преобразованной точки; (xyz) – радиус-вектор \mathbf{P} отображаемой точки; $m_X = (1 + \delta_X z_0)/(1 + \Delta z_0)$ – масштабный коэффициент по оси X ; $m_Y = (1 + \delta_Y z_0)/(1 + \Delta z_0)$ – масштабный коэффициент по оси Y ; δ_X, δ_Y – обратные величины гиперфокальных расстояний, разные при оценке размеров и дальности предметов [4, 13]; $z4 = z3$ из (5). Очевидно, что при $z = z_0$ преобразования $\mathbf{P4}$ и $\mathbf{P3}$ совпадают. Таким образом, предметы или их части в картинной плоскости не испытывают никаких искажений.

Пусть m – коэффициент масштаба с аккомодацией покоя δ , а t – коэффициент дальности с аккомодацией τ , причем $t > m$, а $\tau > \delta$. При $m_X = m_Y = m$ и $\delta_X = \delta_Y = \delta$ преобразование (6) сохраняет подобие и соотношение масштабов разноудаленных предметов. При $m_X = m_Y = t$ и $\delta_X = \delta_Y = \tau$ достоверно передается подобие и воспринимаемая удаленность предметов. При $m_X = m, m_Y = t$ и $\delta_X = \delta, \delta_Y = \tau$ правильно воспроизводится ширина и глубина предметов, но искажается их высота, т. е. нарушается подобие (рис. 4). Здесь наклонная линия обозначает границу поля зрения (расстояния даны в метрах). Графики функции $x4(z4)$ для боковой границы поля зрения приведены на рис. 4, а, функции $y4(z4)$ для верхней границы – на рис. 4, б. Графики даны для разных положений z_0 плоскости K_E . На всех графиках просматривается увеличение предметов в дальней зоне по сравнению с ближней. Причем коэффициент увеличения всегда больше «ренессансного», что соответствует

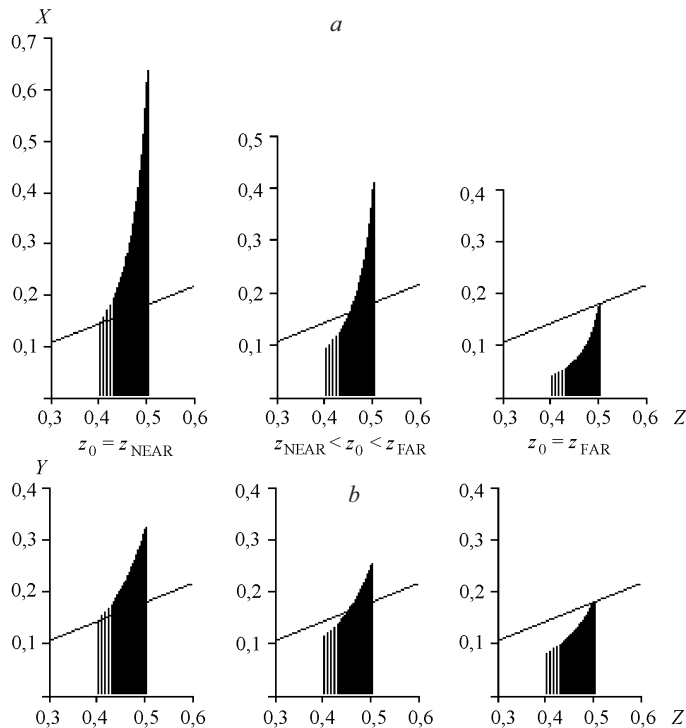


Рис. 4

естественному восприятию. При $z_0 = z_{\text{NEAR}}$ предметы подобны в ближней зоне, по мере удаления ширина становится больше высоты. При $z_0 = z_{\text{FAR}}$, наоборот, в дальней зоне предметы подобны, а по мере приближения высота предметов становится больше ширины. Если плоскость K_E находится в средней зоне, подобие наблюдается также в средней зоне. Как правило, при этом достигается минимум суммарных искажений по всем критериям соответствия изображений естественному восприятию [13].

В общем случае проективное преобразование (6) нелинейно, но при $k = 0$, т. е. для плоского дисплея, его можно записать в виде

$$x_4 = \frac{x m_X}{1 + \delta_X z}, \quad y_4 = \frac{y m_Y}{1 + \delta_Y z}, \quad z_4 = \frac{z_0}{1 + \Delta z_0} = z_{D0},$$

где координата z_4 становится константой. Для вариантов перспектив при $m_X = m_Y$ и $\delta_X = \delta_Y$ преобразование оказывается линейным в однородных координатах и соответствует системам перцептивной перспективы на плоскости [1, 13]. Поэтому можно утверждать, что преобразование (6) является обобщением перцептивной перспективы для 3М-дисплеев.

Заключение. В данной работе рассмотрены некоторые закономерности зрительного восприятия при построении проективной сенсорной модели визуального пространства в сознании человека. Оценены возможности известных дисплеев создавать прямые признаки глубины предметов, такие как конвергенция и аккомодация глаз, необходимые для построения сенсорной модели. Отмечено, что в мировой практике дисплеестроения не существует

средств отображения визуальной информации, отвечающих в полной мере возможностям естественного зрительного восприятия.

Введено понятие идеального 3М-дисплея, пространство которого физически или виртуально совмещено с отображаемым пространством предметов, что гарантирует отсутствие каких-либо искажений. Предложены проективные преобразования, которые являются обобщением ренессансной и перцептивной систем перспектив для 3М-дисплеев, отличных от идеального. Введено понятие виртуальной картинной плоскости, в окрестности которой искажения зрительного восприятия минимальны. Обобщенные системы перспектив включают 2М-дисплеи как частный случай.

Для компьютерной графики могут быть определены две группы систем перспектив: линейные и нелинейные перспективы. Свойства этих групп, а также разработка алгоритмов и программ будут рассмотрены в других работах.

СПИСОК ЛИТЕРАТУРЫ

1. **Раушенбах Б. В.** Системы перспектив в изобразительном искусстве. Общая теория перспектив. М.: Наука, 1986.
2. **Indow T.** The global structure of visual space // Advanced Series of Mathematical Psychology. Singapore: World Scientific Publishing, 2004. Vol. 1.
3. **Ковалев А. М.** О моделях визуального пространства // Оптика и спектроскопия. 2006. **100**, № 1. С. 134.
4. **Ковалев А. М.** Описание визуального пространства в моделях Клейна и Пуанкаре // Автометрия. 2006. **42**, № 4. С. 57.
5. **Ковалев А. М.** Об анизотропной модели визуального пространства // Там же. № 6. С. 53.
6. **Ferguson J. L., Robinson S. D., McLaughlin C. W. et al.** An innovative beamsplitter-based stereoscopic/3D display design // Proc. SPIE. 2005. **5664**. P. 488.
7. <http://www.newsight.com/Technology/Hardware/multiview19-45.html>
8. **Pastoor S., Wöpping M.** 3-D Displays: A review of current technologies // Displays. 1997. **17**, N 2. P. 100.
9. **Компанетс I., Gonchukov S.** 3-D medium based displays // Proc. SPIE. 2005. **5821**. P. 134.
10. <http://www.lightspacetech.com>
11. **Petrov V.** 3-D image reconstruction based on acoustically created hologram // Proc. SPIE. 2005. **5821**. P. 67.
12. **Wöpping M.** Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus // Journ. of the SID. 1995. **3/3**. P. 101.
13. **Ковалев А. М.** Оценка искажений предметов при отображении перцептивного пространства на картинную плоскость // Автометрия. 2004. **40**, № 6. С. 87.

Поступила в редакцию 12 февраля 2007 г.