

ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ КАРТ ГЛУБИНЫ ДЛЯ НЕПОДВИЖНЫХ ИЗОБРАЖЕНИЙ

Н. Н. Красильников^а, доктор техн. наук, профессор

О. И. Красильникова^а, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург, РФ

Постановка проблемы: создание карты глубины является центральным элементом технологии конвертации двумерного изображения в стереоскопическое изображение, однако способы ее изготовления фирмами, специализирующимися в области конвертации изображений, не разглашаются. **Цель исследования:** разработка метода получения пространственных координат объектов, представленных на двумерных изображениях, и разработка на их основе технологии построения карты глубины. **Методы:** для решения поставленной проблемы были использованы законы бинокулярного зрения, методы и алгоритмы трехмерной графики, а также экспериментальные исследования. **Результаты:** разработана технология изготовления карт глубины, которая состоит из следующих последовательно выполняемых операций. Сначала двумерное изображение сегментируется на фрагменты простых поверхностей, аппроксимирующих объекты изображенной сцены с их последующим разделением на две группы. Затем определяются пространственные координаты каждой из аппроксимирующих поверхностей путем анализа исходного двумерного изображения. Далее, исходя из критического значения экранного паралакса, определяется оптимальное значение координаты глубины экрана, на который проецируется стереоскопическое изображение. После этого на карте глубины выполняется заливка аппроксимирующих поверхностей оттенками серого, которые определяются координатам глубины объектов. Показано, что для создания эффекта реальности изображаемой сцены необходимо, чтобы размеры проекций объектов, координаты глубины которых совпадают с координатой глубины экрана, равнялись размерам самих объектов, а это, как правило, требует применения больших экранов, расположенных на значительном удалении от зрителя. В кинотеатрах это условие выполняется, а при просмотре 3D-фильмов на экранах 3D-телевизоров нет, что приводит к возникновению эффекта театра марионеток. **Практическая значимость:** разработанная технология позволяет при конвертации двумерных изображений и кинофильмов в стереоскопический (3D) формат создавать карты глубины на основании расчетов, избегая трудоемкого подбора параметров.

Ключевые слова – стереоскопическое изображение, карта глубины.

Введение

Проблема определения координаты глубины реального объекта, представленного на единственном двумерном изображении, возникает, в частности, при конвертации этого изображения в стереоскопический формат. Как правило, эта проблема имеет решение, что следует из повседневного жизненного опыта, в чем нетрудно убедиться. Действительно, при монокулярном наблюдении окружающего пространства легко определяются расстояния до окружающих нас объектов. Объясняется это тем, что центральная проекция наблюдаемой сцены на сетчатку глаза, как правило, содержит необходимую информацию для восстановления утраченной при проекции координаты глубины. Эта информация используется специальными зрительными механизмами, к числу которых, в частности, относятся: механизм, основанный на том, что удаленные объекты кажутся меньших размеров, чем объекты, расположенные близко; механизм, использующий свойство центральной проекции, заключающееся в том, что удаленные объекты расположены на изображении ближе к линии горизонта, чем объекты, расположенные ближе к зрителю, а также ряд других [1–3]. Использование этой особенности центральной

проекции позволяет восстанавливать утраченную при проецировании информацию о глубине сцены, представленной на 2D-изображении, создавая так называемую карту глубины.

Сегментация 2D-изображения на аппроксимирующие поверхности

Карта глубины представляет собой черно-белое (grayscale) двумерное изображение силуэтов объектов исходного изображения, на котором яркость каждого из силуэтов определяет его координату глубины. Карту глубины можно рассматривать как предельный случай векторной полигональной модели 3D-изображения обращенных к камере оболочек объектов сцены, когда размер полигона вырождается в один пиксель. Будучи спроецированным на эти оболочки, 2D-изображение создает виртуальную 3D-сцену. Таким образом, 2D-изображение совместно с картой глубины описывают 3D-сцену, и хотя часто карта глубины применяется при использовании стереоскопического метода отображения трехмерного пространства, ее роль значительно шире и она имеет самостоятельное значение.

Обратим внимание также на то, что при использовании карты глубины совместно с 2D-изображением для создания стереоскопического

(3D) изображения посредством яркости L каждого пикселя карты глубины представляют значение экранного параллакса p , а не самой координаты глубины:

$$L = A_0 + A_1 p/b, \quad (1)$$

где A_0 и A_1 — постоянные коэффициенты; b — расстояние между зрачками глаз зрителя (база). Экранный параллакс, однозначно связанный с координатой глубины виртуального объекта d_{ob} на 3D-изображении известным соотношением

$$p = \frac{b(d_{ob} - d_{sc})}{d_{ob}}, \quad (2)$$

где d_{sc} — расстояние от зрителя до экрана, обладает тем преимуществом, что имеет более компактную шкалу значений, чем шкала координаты глубины d_{ob} [4]. При этом значения коэффициентов A_0 и A_1 выбирают равными $A_0 = 127$ и $A_1 = -127$. Таким образом:

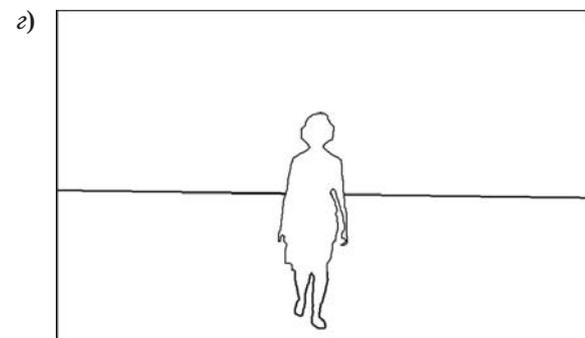
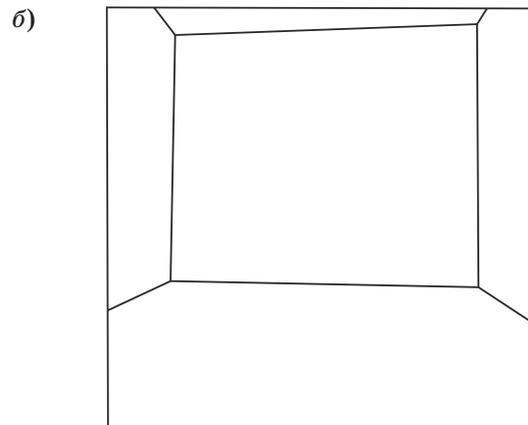
$$L = 127(1 - p/b) \quad (3)$$

и, соответственно:

$$p/b = 1 - L/127, \quad (4)$$

где L — яркость, которая представляется номером квантового уровня. При таком выборе коэффициентов A_0 и A_1 максимально используется диапазон яркостей карты глубины. При создании карты глубины наиболее удаленные объекты обычно относят к фону, а их экранный параллакс приравнивается значению экранного параллакса бесконечно удаленных объектов, т. е. равным базе b , что соответствует $L = 0$. Для наиболее близко расположенных объектов яркость на карте глубины обычно устанавливается равной уровню квантования 255.

На практике, как правило, используют приближенные карты глубины, при создании которых поверхности объектов сцены на 2D-изображении аппроксимируют небольшим количеством простых поверхностей, например, плоскостей, гауссовых поверхностей, фрагментов цилиндрических поверхностей или фрагментов поверхностей эллипсоидов [5]. С этой целью исходное 2D-изображение сегментируется на фрагменты, каждый из которых представляется одной из



■ **Рис. 1.** Исходное 2D-изображение (а, в) и схема, поясняющая его разделение на фрагменты аппроксимирующих поверхностей (б, г)
 ■ **Fig. 1.** The original 2D image (a, в) and the diagram explaining its division into fragments of approximating surfaces (б, г)

таких поверхностей. В случае неподвижных изображений процесс этот, как правило, выполняется вручную посредством так называемого метода ротоскопирования, при котором последовательно обрисовывается каждый объект. Выделенные при этом фрагменты могут представлять собой изображения как отдельных объектов сцены, так и их групп.

Объекты сцены условно можно разделить на две группы. К первой группе относятся объекты, формирующие пространство сцены, например, стены, потолок и пол интерьера, поверхность дорожного покрытия улицы, стены стоящих по бокам домов и т. п. Особенность этих объектов заключается в том, что они имеют большую протяженность по глубине и если они аппроксимируются плоскостями, то нанесенные на них параллельные линии сходятся в точку схода при бесконечном удалении. Ко второй группе относятся объекты, заполняющие сцену, например, предметы мебели, люди, деревья, автомобили, фонарные столбы и т. д. Особенность этих объектов состоит в том, что, во-первых, они имеют небольшую протяженность по глубине и могут быть аппроксимированы фрагментами плоскостей, параллельными плоскости исходного 2D-изображения, или фрагментами плоскостей, расположенными под небольшим углом к этой плоскости, это так называемые «плоские объекты». Во-вторых, эти аппроксимирующие фрагменты плоскостей опираются на горизонтальную или близкую к ней поверхность, например, на пол помещения, на поверхность земли и т. д., а не парят в воздухе.

Отмеченная особенность объектов исходного 2D-изображения проявляется при определении координат глубины различных частей этих объектов. В первом случае всегда имеется направление, перемещение по которому, не выходя за пределы фрагмента, соответствует изменению координаты глубины представленного на 2D-изображении фрагмента сцены. Во втором случае такое направление отсутствует.

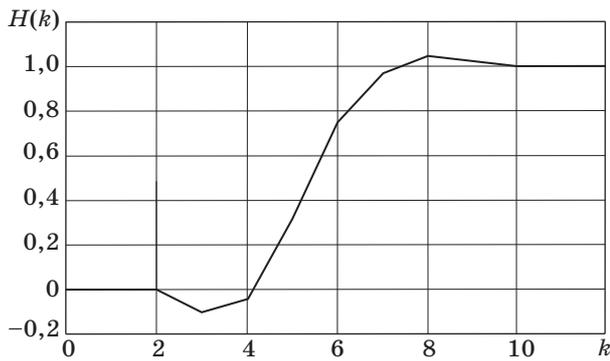
В качестве примера на рис. 1, *a–г* приведены исходные 2D-изображения и схемы, поясняющие их разделение на фрагменты аппроксимирующих поверхностей. На рис. 1, *a* объекты: пол, потолок и стены интерьера — аппроксимируются фрагментами плоскостей, которые соответствуют двум горизонтальным и трем вертикальным поверхностям сцены. На рис. 1, *в* объект — небосвод — аппроксимируется фрагментом сферы, а объекты поле и фигура девушки, идущей по полю, аппроксимируются фрагментами плоскостей. При этом первые два объекта являются протяженными по глубине, а третий имеет малую протяженность, это так называемый «плоский» объект.

Допустимая величина геометрических погрешностей при сегментации

Геометрические погрешности, допущенные при изготовлении карт глубины, заключающиеся в расхождении границ силуэтов с границами отображаемых ими объектов, проявляются на стереоскопических изображениях двойко. Если размеры силуэтов объектов на карте глубины превышают размеры соответствующих им объектов на 2D-изображении, то на стереоскопическом изображении узкие области пространства, окружающего такие объекты, смещаются по глубине к этим объектам, и появляются ореолы, заполненные деталями фона, на котором они наблюдаются. В противном случае если размеры силуэтов объектов на карте глубины оказываются меньше размеров соответствующих им объектов на 2D-изображении, то на стереоскопическом изображении узкие области объектов, примыкающие к их контурам изнутри, отсоединяются от них и смещаются по глубине к областям окружающего их пространства, например к фону.

Для оценки допустимого уровня погрешностей этого типа были выбраны шесть 2D-изображений различного сюжетного содержания и, соответственно, различной структуры, а также соответствующие им карты глубины. Далее карты глубины были подвергнуты следующим преобразованиям. Вначале из каждой карты глубины посредством графического редактора PhotoShop были изготовлены версии, геометрические размеры которых были уменьшены в n раз, где $n = 2, 4, 8, 16, 32$ и 64 раза соответственно, а затем сохранены в памяти. При уменьшении линейного размера карты глубины количество отсчетов (пикселей), которым она представляется, также сокращается пропорционально величине этого уменьшения. После этого каждая из уменьшенных версий карт глубины с уменьшенным количеством отсчетов посредством этого же редактора была восстановлена до своего первоначального размера, при этом недостающие отсчеты восстанавливались редактором путем интерполяции. Описанное преобразование вносило в карты глубины погрешности, которые выражались в размытии резких световых границ. На рис. 2 приведена диаграмма, иллюстрирующая это размытие для случая $n = 2$. Из графика, на котором по оси ординат отложена относительная величина яркостного перепада, а по оси абсцисс — номера отсчетов, следует, что зона размытия составляет 3 отсчета.

При $n = 4$ зона размытия возрастает до шести отсчетов. Далее для каждой пары «2D-изображение — преобразованная карта глубины» были получены соответствующие им стереоскопические изображения в анаглифном формате.



■ **Рис. 2.** Иллюстрация размытия световой границы, $n = 2$
 ■ **Fig. 2.** Illustration of the blurring of the light edge for the case, when $n = 2$

Естественно, что чем в большее количество раз при описанных выше преобразованиях уменьшаются размеры карты глубины, тем большие геометрические погрешности она содержит и, соответственно, тем больший уровень искажений возникает в стереоскопическом изображении, которое получается с использованием этой карты глубины.

Для оценки величины геометрических погрешностей, допущенных при изготовлении карты глубины, приводящих к заметным зрительным искажениям на стереоскопическом изображении, было разработано специальное приложение, посредством которого на экран монитора для просмотра выводились одновременно два стереоскопических изображения в анаглифном формате: эталонное, полученное с использованием неискаженной карты глубины, и искаженное, полученное с использованием карты глубины, содержащей геометрические искажения. Зритель при выполнении эксперимента рассматривал и сравнивал между собой оба стереоскопических изображения через анаглифные очки. Перед выводом стереоскопических изображений на экран зрителю сначала предъявлялись «стирающие», шумовые изображения. При выводе эталонного и искаженного изображений на экран их взаимное расположение (справа, слева) было случайным, неизвестным зрителю. Время предъявления изображений для их сравнения при выполнении эксперимента устанавливалось на один из интервалов, которые равнялись 10 и 20 с соответственно, после чего оба изображения с экрана удалялись.

Задача зрителя состояла в определении, какое из двух предъявленных изображений, правое или левое, имеет более высокое качество, т. е. является эталонным. В зависимости от принятого решения зритель должен был нажать одну из кнопок, расположенных в нижней части окна приложения, соответственно, справа или слева.

- Зависимость относительного количества P правильных решений зрителя для разных значений кратности изменения размеров карты глубины n
- Dependence of the relative number P of correct decisions of the viewer for different values of the multiplicity of the change in the size of the depth map n

n	P при времени наблюдения	
	10 с	20 с
2	0,5108	0,538833
4	0,5922	0,6075
8	0,6443	0,685167
16	0,7543	0,751833
32	0,8105	0,825167
64	0,7933	0,819667

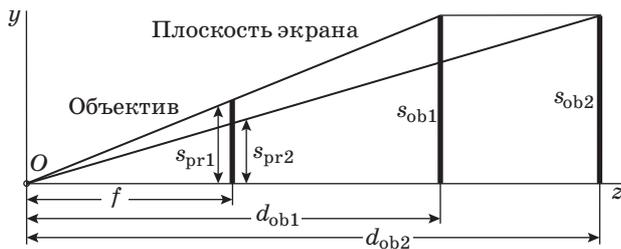
С каждой парой стереоскопических изображений (эталонное — искаженное) зритель выполнял по 20 сравнений, при этом приложение автоматически вычисляло относительное количество P правильных решений зрителя и выводило эту величину на экран. Полученные данные записывались экспериментатором. В экспериментах приняли участие 11 человек. В таблице приведены результаты измерений, усредненные по шести изображениям и 11 зрителям.

Из таблицы видно, что при $n = 2$ зрители не отличают стереоскопические изображения, полученные при использовании искаженной карты глубины, у которой зона размытости световых границ составляет 3 отсчета (пикселя), от стереоскопических изображений, которые получены с использованием неискаженной карты глубины. Из этого следует, что в качестве максимально приемлемой геометрической погрешности при изготовлении карт глубины можно рекомендовать погрешность ± 2 пикселя. Анализ карт глубины, опубликованных в Интернете, дает примерно такое же значение геометрической погрешности.

Определение пространственных координат аппроксимирующих поверхностей и выбор координаты глубины экрана

Для определения пространственных координат и размеров объектов, представленных на исходном 2D-изображении, необходимо знание фокусного расстояния объектива камеры, которой сделано это изображение.

Фокусное расстояние объектива определяет его поле зрения и в простейшем случае может быть прочитано на его корпусе. При этом следует иметь в виду, что фокусное расстояние, обозначенное на корпусе объектива камеры, приведено к изображению размером 36×24 мм. При использовании изображений, предварительно



- **Рис. 3.** Модель камеры с передней плоскостью изображения и два объекта в виде отрезков прямых, расположенные в координатной плоскости yz , расстояние между которыми известно
- **Fig. 3.** The camera model with front image plane and two objects in the form of straight lines located in a coordinate plane yz , the distance between which is known

увеличенных в K раз, необходимо использовать эффективное фокусное расстояние, увеличенное в это же число раз.

Если фокусное расстояние камеры, посредством которой получено 2D-изображение, не известно, его следует определить по самому изображению. Это можно сделать, если на изображении имеются по крайней мере два объекта, опирающихся на горизонтальную поверхность, например на дорожное покрытие, для которых известны размеры и расстояние между ними по глубине Δd . Примером могут служить фотографии, на которых представлено 2D-изображение городской улицы с автомобилями, трамваями, зданиями и другими объектами, имеющими заметную протяженность в направлении всех трех координат, которую можно приближенно оценить. Поясним сказанное рис. 3, на котором представлена модель камеры с передней плоскостью изображения и два объекта в виде отрезков прямых длиной s_{ob1} и s_{ob2} , расположенных в координатной плоскости yz , расстояние между которыми известно и равно Δd :

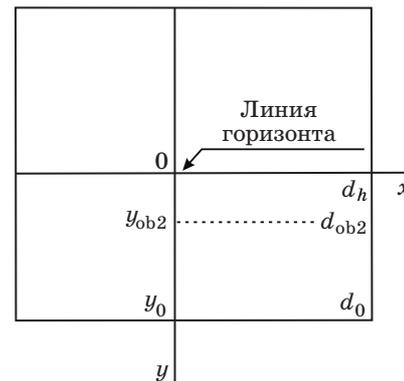
$$\Delta d = d_{ob2} - d_{ob1}, \quad (5)$$

где d_{ob1}, d_{ob2} — координаты глубины этих объектов. Для определения координат глубины в рассматриваемом случае составим очевидные пропорции

$$\frac{s_{ob1}}{s_{pr1}} = \frac{d_{ob1}}{f}, \quad (6)$$

$$\frac{s_{ob2}}{s_{pr2}} = \frac{d_{ob2}}{f}, \quad (7)$$

где s_{pr1} и s_{pr2} — проекции объектов. Решая совместно уравнения (5)–(7), находим значение фокусного расстояния камеры



- **Рис. 4.** Исходное 2D-изображение, на котором показана система координат x, y и положение линии горизонта
- **Fig. 4.** The original 2D image, which shows the coordinate system x, y and the position of the horizon line

$$f = \frac{s_{pr2} \Delta d}{s_{ob2} \left(1 - \frac{s_{ob1} s_{pr2}}{s_{ob2} s_{pr1}} \right)}. \quad (8)$$

Обратим внимание, что в данном случае найдено эффективное фокусное расстояние камеры, при этом координаты глубины объектов находятся как

$$d_{ob1} = f \frac{s_{ob1}}{s_{pr1}}; \quad (9)$$

$$d_{ob2} = f \frac{s_{ob2}}{s_{pr2}}. \quad (10)$$

Границы исходного 2D-изображения, его система координат x, y и положение на нем линии горизонта показаны на рис. 4. Кроме того, на рисунке использованы следующие обозначения: координата глубины линии горизонта d_h , координата глубины ближайшей к камере части горизонтальной поверхности d_0 , на которой расположены объекты, а также координата глубины одного из объектов, в данном случае d_{ob2} .

Нетрудно видеть, что координата y любого точечного объекта на 2D-изображении и его координата глубины d в отображаемом пространстве связаны соотношением

$$d = \frac{y_0 d_0}{y}. \quad (11)$$

Другими словами, координата глубины изменяется обратно пропорционально координате y , благодаря чему на фотографии железнодорожного пути рельсы, оставаясь прямыми, сходятся на

горизонте в точку. Величина d_0 находится путем подстановки в формулу (11) значений d_{ob2} и y_{ob2} и решения получившегося уравнения относительно d_0 :

$$d_0 = \frac{y_{ob2} d_{ob2}}{y_0}.$$

Таким образом, полученная формула позволяет определить координату глубины любого объекта путем подстановки в нее координаты точки опоры его на горизонтальную поверхность y . Размеры объектов находятся по формуле

$$s_{ob} = \frac{s_{pr} d_{ob}}{f}.$$

После того как определены координата глубины, соответствующая нижней границе изображения d_0 , и положение линии горизонта, координата глубины которой $d_h \rightarrow \infty$, следует перейти к выбору расстояния до экрана (координаты глубины экрана), на который будет выводиться стереоскопическое изображение, d_{sc} . При выборе расстояния от зрителя до экрана d_{sc} потребуем, чтобы при воспроизведении стереоскопического изображения отсутствовали бы такие неприятные артефакты, как ощущение дискомфорта,

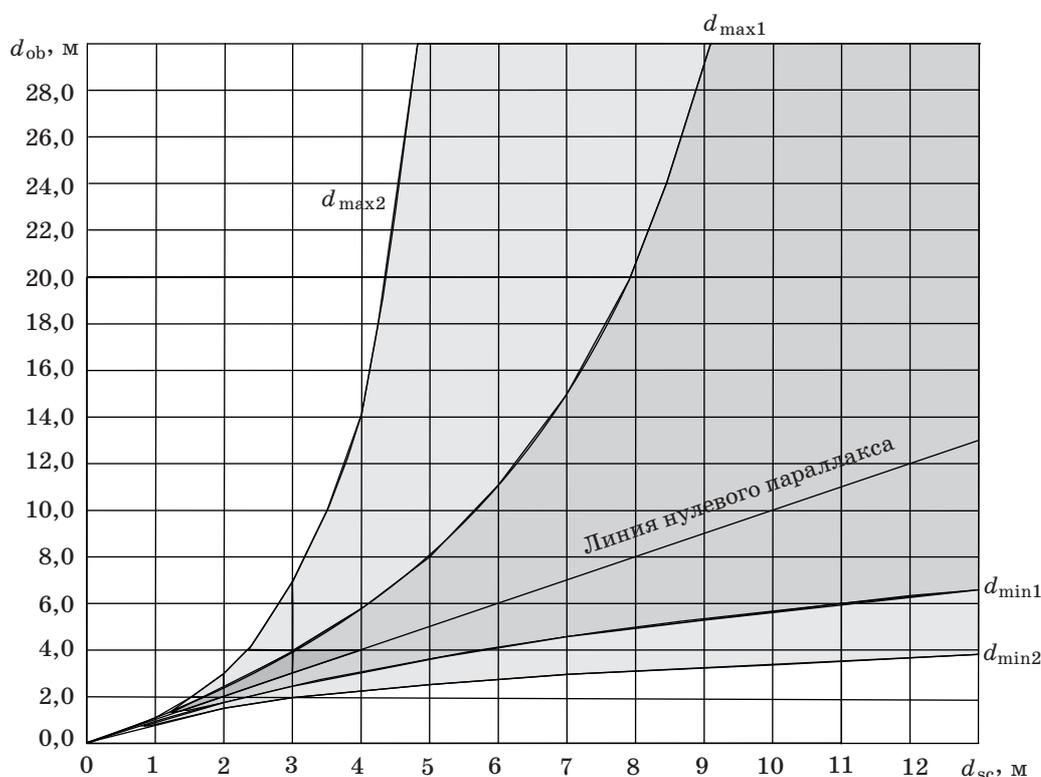
двоение изображения и пр. [6]. Для этого, как было показано в работе [7], необходимо, чтобы координаты глубин объектов, воспроизводимых в виртуальном пространстве, не выходили за пределы интервала $d_{min} - d_{max}$:

$$d_{min} = \frac{d_{sc}}{1 + \frac{d_{sc} \operatorname{tg} |\alpha_{cr}|}{b}}; \quad (12)$$

$$d_{max} = \frac{d_{sc}}{1 - \frac{d_{sc} \operatorname{tg} |\alpha_{cr}|}{b}}, \quad (13)$$

где $|\alpha_{cr}|$ — абсолютная величина критического значения углового размера экранного параллакса [7, 8]. На рис. 5 в виде кривых приведены зависимости граничных значений d_{min} и d_{max} от d_{sc} для двух значений $|\alpha_{cr}|$, равных 17 и 40 соответственно.

Из рассмотрения графика на рис. 5 следует, что чем более жесткие требования предъявляются к стереоскопическому изображению в отношении возникновения артефактов в виде ощущений дискомфорта и двоения изображения, выражающиеся в уменьшении критического значения экранного параллакса, тем большее следует выбирать расстояние до экрана d_{sc} . Беглый анализ



■ **Рис. 5.** Зависимости граничных значений d_{min} и d_{max} от d_{sc} для двух значений $|\alpha_{cr}|$, равных 17 и 40' соответственно
 ■ **Fig. 5.** Dependences of the boundary values d_{min} and d_{max} on d_{sc} for two values of $|\alpha_{cr}|$ equal to 17 and 40' respectively

3D-фильмов показал, что в большинстве случаев их создатели ориентируются на значения $|\alpha_{cr}| < 40$ угл. мин, хотя встречаются значения, достигающие до 120 угл. мин и более [7, 8].

При выборе расстояния до экрана необходимо обеспечить представление на стереоскопическом изображении заднего плана с минимальными искажениями по глубине. Учитывая изложенное, потребуем, чтобы координата глубины d_{max} равнялась бы координате глубины линии горизонта d_h , т. е. чтобы

$$d_{max} = d_h, \quad (14)$$

при этом по определению $d_h \rightarrow \infty$. Решая совместно (13) и (14) относительно d_{sc} с учетом того, что $d_h \rightarrow \infty$, находим

$$d_{sc} = b / \operatorname{tg}|\alpha_{cr}|. \quad (15)$$

Значение d_{min} получим путем подстановки найденного значения d_{sc} в формулу (12), что дает

$$d_{min} = d_{sc} / 2. \quad (16)$$

После этого необходимо проверить, попадает ли в интервал допустимых значений координата глубины нижней части изображения d_0 путем ее сравнения с координатой глубины d_{min} . В случае, если окажется, что $d_0 < d_{min}$, необходимо либо отказаться от конвертации всего изображения в 3D-формат, «отрезав» его нижнюю часть, либо сознательно пойти на искажения правильной передачи его глубины.

Перейдем теперь к определению экранного параллакса, для чего обратимся к формуле (11). Учитывая, что $d_0 = d_{min}$, а d_{min} согласно (16), равно $d_{sc} / 2$, запишем

$$d_{ob} = \frac{y_0 d_{sc}}{2y}. \quad (17)$$

Делая подстановку d_{ob} в формулу (2) и решая получившееся уравнение относительно экранного параллакса, запишем

$$p = b \left(1 - \frac{2y}{y_0} \right). \quad (18)$$

Таким образом, величина экранного параллакса по мере перемещения от линии горизонта в сторону нижней границы изображения линейно изменяется от максимального положительного значения b до максимального по абсолютной величине отрицательного значения b . Из этого следует, что плоскость, на которой расположены объекты исходного изображения, на карте глубины может быть залита посредством ли-

нейной градиентной заливки, при этом яркость линии горизонта должна быть представлена нулевым (или близким к нулю) уровнем квантования, а яркость нижней границы карты глубины — уровнем квантования 255 (или близким к нему).

Заливка аппроксимирующих поверхностей

После разделения 2D-изображения на аппроксимирующие поверхности следует приступить к заливке выделенных контурами областей. Начинать надо с протяженных объектов, для которых делается градиентная заливка. Например, в случае изображения, показанного на рис. 1, в, следует сделать линейную градиентную заливку фрагмента плоскости, которым аппроксимируется поле, и радиальную градиентную заливку фрагмента сферы, которым аппроксимируется небосвод. При этом целесообразно минимальную яркость взять равной или близкой нулевому уровню квантования по яркости, а максимальную — равной или близкой 255-му уровню квантования по яркости. Такой выбор обеспечит максимальное количество уровней квантования экранных параллаксов. После заливки выделенных фрагментов поверхностей, аппроксимирующих протяженные объекты, следует перейти к заливке фрагментов поверхностей, аппроксимирующих «плоские» объекты. Как правило, это вертикально ориентированные объекты, например, люди, столбы, деревья и т. д. Особенность их заливки заключается в том, что яркость силуэта заливаемой фигуры должна совпадать с яркостью поверхности, на которой эта фигура стоит. На рис. 1, в таким фрагментом является выделенная контуром фигура девушки, идущей по полю, которая аппроксимируется плоскостью. На рис. 6 показана



■ **Рис. 6.** Карта глубины изображения, приведенного на рис. 1, в, после заливки
 ■ **Fig. 6.** The depth map of the image shown in Fig. 1, в, after the filling

карта глубины изображения, представленного на рис. 1, в, после заливки.

Конвертация 2D-изображения в 3D-формат и представление стереоскопических изображений на малом экране

Заключительным этапом подготовки к конвертации 2D-изображения в 3D-формат является перевод всех найденных ранее размеров из единиц длины в количество пикселей. В результате предыдущих расчетов была получена формула (18), представляющая зависимость величины экранного параллакса p от координаты y для горизонтальной поверхности, на которой стоят объекты сцены, например поверхности земли:

$$p = b \left(1 - \frac{2y}{y_0} \right).$$

В этой формуле значения переменных в ее правой части известны и представлены в единицах длины, а следовательно, и рассчитанное по этой формуле значение экранного параллакса также выражается в этих же единицах, например в метрах, где $b = 0,065$ м. Однако реально на экране компьютера имеется изображение, представленное в виде матрицы пикселей, размеры которого определены числом строк и числом пикселей в строке. Поэтому всю дальнейшую обработку изображения и, в частности, установку экранных параллаксов необходимо выполнять, используя в качестве единицы измерения количество пикселей. При этом величина экранного параллакса p и базы b также должны быть представлены количеством пикселей. Для этого необходимо определить значение базы также в пикселях $b_{\text{пик}}$. С этой целью вначале определяется ширина 2D-изображения в пикселях на экране монитора $n_{\text{пик}}$. Затем определяется ширина w_0 ближайшей к камере части сцены, отображаемой на 2D-изображении, в метрах, что, как правило, несложно сделать, ориентируясь на размеры знакомых объектов. Принимая во внимание, что, согласно (18), при $y = y_0$ экранный параллакс $p = -b$, составим очевидную пропорцию

$$\frac{b_{\text{пик}}}{b} = \frac{n_{\text{пик}}}{w_0},$$

откуда

$$b_{\text{пик}} = b \frac{n_{\text{пик}}}{w_0}. \quad (19)$$

Полученный результат следует округлить.

И, наконец, необходимо определить геометрические размеры проекции на экран конвертируе-

мого 2D-изображения, а также его карты глубины, полагая, что $d_0 = d_{\text{min}}$.

Размеры объектов, приведенные к экрану, определяются путем изменения масштаба исходного 2D-изображения в K_0 раз:

$$K_0 = d_{\text{sc}} / f, \quad (20)$$

где f — эффективное фокусное расстояние.

Для того чтобы для зрителя виртуальное 3D-пространство, создаваемое стереоскопическим изображением, было бы не отличимо от реального 3D-пространства, необходимо отсутствия конфликта между зрительными механизмами, посредством которых зрительная кора оценивает его глубину. Если это условие не обеспечивается, возникают различного рода артефакты. Один из таких артефактов возникает, когда размер изображения, выводимого на экран, не соответствует тому, который бы должен быть. Проиллюстрируем это следующим примером.

Известно, что при наблюдении сцены, содержащей объекты, размеры $s_{\text{об}}$ которых знакомы зрителю, одним глазом, зрительная система без труда оценивает их удаленность $d_{\text{об}}$, основываясь на размере их проекций $s_{\text{пр}}$ на сетчатку глаза и фокусном расстоянии оптики глаза f :

$$d_{\text{об}} = \frac{s_{\text{об}} f}{s_{\text{пр}}}.$$

В случае наблюдения стереоскопического изображения двумя глазами зрительной системе по-прежнему известны $s_{\text{пр}}$, f . Однако в этом случае зрительная система приоритет отдает данным $d_{\text{об}}$, полученным посредством бинокулярного зрения на основании определения параллакса. Поскольку при этом оценка координаты глубины механизмом зрения, основанным на оценке масштаба проекции, и механизмом стереоскопического зрения различаются, то зрительная кора устраняет это расхождение, принимая результаты оценки первым механизмом недостоверными и «назначая» объекту другой размер, равный

$$s_{\text{об}} = \frac{s_{\text{пр}} d_{\text{об}}}{f}.$$

Применительно к наблюдению стереоскопического изображения на малом экране, соответственно в уменьшенном масштабе, зритель увидит объекты в уменьшенном размере, а глубину сцены сокращенной, т. е. возникает эффект театра марионеток. С этим случаем мы встречаемся при наблюдении стереоскопических изображений на экранах 3D-телевизоров. Выходом из создавшегося положения может служить применение 3D-очков (по существу, стереоскопов), в ко-

торых посредством специальной оптики достигается эффект демонстрации стереоскопических изображений на большом экране в необходимом масштабе [9].

Заключение

Методы получения и демонстрации 3D-изображений являются одним из направлений интенсивного развития информационных технологий. Следует отметить, что стереоскопический формат демонстрации 3D-изображений хотя и обладает рядом существенных недостатков, тем не менее применяется в настоящее время, поскольку для его реализации созданы сравнительно несложные устройства отображения. Можно предположить, что следующим этапом развития технологии демонстрации 3D-изображений будет переход от стереоскопического метода демонстрации 3D-изображений к голографическому методу, свободному от недостатков стереоскопического метода, однако в настоящее время трудно

оценить время появления голографических дисплеев, поскольку их разработка связана с большими технологическими трудностями.

Обращаясь к области применения 3D-изображений и видео, заметим, что она не ограничивается развлекательной индустрией. Эти технологии уже сейчас используются для 3D-визуализации и реконструкции предметов культурного наследия. Они могут быть также применены для релаксации экипажей, работающих в замкнутых пространствах, например, экипажей космических кораблей, атомных подводных лодок и т. п., где клаустрофобия является одним из факторов, ограничивающих допустимое время работы [10–12].

Естественно также ожидать, что появление технологии 3D-изображений в конечном итоге приведет к возникновению нового вида искусства со своими приемами подачи идей, подобно тому, как изобретение братьев Люмьер привело к появлению искусства кинематографии, которое отличается от театрального искусства.

Литература

1. Красильников Н. Н. Метод формирования 3D-изображения сцены по одной фотографии // Информационно-управляющие системы. 2013. № 3. С. 2–7.
2. Красильников Н. Н., Красильникова О. И. Метод конвертации 2D-изображения в стереоскопическое 3D-изображение // Оптический журнал. 2014. № 2(82). С. 20–28.
3. Красильников Н. Н., Красильникова О. И. Определение координаты глубины по 2D-изображению // Оптический журнал. 2011. № 12(78). С. 30–33.
4. Красильников Н. Н., Красильникова О. И. Методы конвертации 2D-изображений и видео в стереоскопический формат // Информационно-управляющие системы. 2015. № 5. С. 18–25. doi:10.15217/issn1684-8853.2015.5.18
5. Пат. 2600524 РФ. Способ конвертации 2D-изображения в квазистереоскопическое 3D-изображение / Красильников Н. Н., Красильникова О. И. (РФ). — № 2014129256; заявл. 15.07.2014; зарегистр. 29.09.2016.
6. Почему от 3D болит голова. Ч. 4. Параллакс // <https://geektimes.ru/post/248970/> (дата обращения: 05.01.2017).
7. Красильников Н. Н. Влияние расстояния наблюдения на глубину пространства, воспроизводимую стереоскопическим изображением // Оптический журнал. 2016. № 8(83). С. 68–76.
8. Красильников Н. Н., Красильникова О. И. Исследование критического значения экранного параллакса // Информационно-управляющие системы. 2016. № 2. С. 19–25. doi:10.15217/issn1684-8853.2016.2.19
9. Портал для избранных. Oculus Rift Development Kit 2. <http://www.hi-fi.ru/magazine/video/portal-dlya-izbrannykh-oculus-rift-development-kit-2/> (дата обращения: 05.01.2017).
10. Клаустрофобия. Причины, симптомы и признаки, лечение, профилактика патологии. <http://www.polismed.com/articles-klaustrofobija-prichiny-simptomu-i-priznaki-lechenie-profilaktika-patologii.html> (дата обращения: 10.04.2017).
11. Константинов А. Как подводники выдерживают трехмесячное пребывание под водой. <http://www.online812.ru/2010/07/29/007/> (дата обращения: 10.04.2017).
12. Приложение к статье «Проблемы редактирования 3D-изображений и видео». http://cv.guap.ru/krasilnikov_nn/files/pril007 (дата обращения: 09.04.2017).

UDC 612.843.721

doi:10.15217/issn1684-8853.2017.5.52

Technology of Making Depth Maps for Still ImagesKrasilnikov N. N.^a, Dr. Sc., Tech., Professor, NNKrasilnikov@yandex.ruKrasilnikova O. I.^a, PhD, Tech., Associate Professor, OIKrasilnikova@yandex.ru^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation

Introduction: The creation of a depth map is a central element in the technology of converting a two-dimensional image into a stereoscopic one, but the companies specializing in image conversion never disclose their depth map manufacturing methods. **Purpose:** We need to develop a method of obtaining spatial coordinates of objects presented on two-dimensional images in order to use them for the development of a depth map construction technology. **Methods:** To solve the problem, we used the binocular vision laws, methods and algorithms of three-dimensional graphics, and experimental studies. **Results:** We have developed a technology of depth map production which consists of the following sequentially executed operations. First, a two-dimensional image is segmented into fragments of simple surfaces which approximate the objects of the scene and subsequently divide them into two groups. Then, the spatial coordinates of each of the approximating surfaces are determined by analyzing the original two-dimensional image. Further, proceeding from the critical value of the on-screen parallax, the optimal value is determined for the depth coordinate of the screen onto which a stereoscopic image is projected. After this, on the depth map the approximating surfaces are filled with shades of gray which are determined by the depth coordinates of the objects. It is shown that to create the effect that the depicted scene is real, it is necessary that the sizes of the projections of the objects whose depth coordinates coincide with the depth coordinate of the screen are equal to the sizes of the objects themselves, and this, as a rule, requires the use of a large screen located at a considerable distance from the viewer. This condition is satisfied in cinemas, but not when watching a 3D movie on a 3D-TV screen, which leads to the puppet-theater effect. **Practical relevance:** In the conversion of two-dimensional images and film files into a stereoscopic (3D) format, the developed technology allows you to create depth maps on the base of calculations, avoiding laborious selection of parameters.

Keywords — Stereoscopic Image, Depth Map.

References

1. Krasilnikov N. N. The Method of Generating a 3D-Image Scene Based on a Single Photograph. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 3, pp. 2–7 (In Russian).
2. Krasilnikov N. N., Krasilnikova O. I. The Method of Converting a 2D Image into a Stereoscopic 3D Image. *Opticheskii zhurnal* [Journal of Optical Technology], 2014, no. 2(81), pp. 68–74 (In Russian).
3. Krasilnikov N. N., Krasilnikova O. I. Determining the Depth Coordinate from a 2D Image. *Opticheskii zhurnal* [Journal of Optical Technology], 2011, no. 12(78), pp. 787–789 (In Russian).
4. Krasilnikov N. N., Krasilnikova O. I. Methods of Converting 2D Images and Videos into Stereoscopic Format. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 5(78), pp. 18–25 (In Russian). doi:10.15217/issn1684-8853.2015.5.18.
5. Krasilnikov N. N., Krasilnikova O. I. *Sposob konvertatsii 2D-izobrazhenija v kvazistereoskopicheskoe 3D-izobrazhenie* [The Method of Converting a 2D Image into a Quasi-stereoscopic 3D Image]. Patent RF, no. 2600524, 2016.
6. *Pochemu ot 3D bolit golova*. Ch. 4. *Parallaks* [Why Headache from 3D. Part 4. Parallax]. Available at: <https://geektimes.ru/post/248970/> (accessed 05 January 2017).
7. Krasilnikov N. N. Effect of Viewing Distance on Spatial Depth Reproduced by a Stereoscopic Image. *Opticheskii zhurnal* [Journal of Optical Technology], 2016, no. 8(83), pp. 502–507 (In Russian).
8. Krasilnikov N. N., Krasilnikova O. I. Study of Critical Values of On-Screen Parallax. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 2, pp. 19–25 (In Russian). doi:10.15217/issn1684-8853.2016.2.19
9. *Portal dlia izbrannykh. Oculus Rift Development Kit 2* [The Portal for the Elite. Oculus Rift Development Kit 2]. Available at: <http://www.hi-fi.ru/magazine/video/portal-dlya-izbrannykh-oculus-rift-development-kit-2/> (accessed 05 January 2017).
10. *Klaustrofobiia. Prichiny, simptomy i priznaki, lechenie, profilaktika patologii* [Claustrophobic. Causes, Symptoms and Signs, Treatment, Prevention of Disease]. Available at: <http://www.polismed.com/articles-klaustrofobiya-prichiny-simptomy-i-priznaki-lechenie-profilaktika-patologii.html> (accessed 10 April 2017).
11. Konstantinov A. *Kak podvodniki vyderzhivaiut trekhmesiachnoe prebyvanie pod vodoi* [How do Submariners Survive a Three-Month Stay under Water]. Available at: <http://www.online812.ru/2010/07/29/007/> (accessed 10 April 2017).
12. *Prilozhenie k stat'e «Problemy redaktirovaniia 3D-izobrazhenii i video»* [The Appendix to the Article «Problems of Editing 3D Images and Video»]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril007 (accessed 09 April 2017).