

УДК 004.932

МЕТОД ФОРМИРОВАНИЯ 3D-ИЗОБРАЖЕНИЯ СЦЕНЫ ПО ОДНОЙ ФОТОГРАФИИ

Н. Н. Красильников,

доктор техн. наук, профессор

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Описан метод формирования 3D-изображения сцены, основанный на аппроксимации ее центральной проекции (2D-изображения) набором плоскостей с последующей декомпозицией сцены на эти плоскости. Для каждой из плоскостей находится аксонометрическая проекция путем использования имеющейся априорной информации об изображенной сцене и карта глубины. Завершающим шагом описываемого метода является «сборка» 3D-изображения сцены объединением аппроксимирующих ее плоскостей и проекция сцены на экран при заданных условиях наблюдения.

Ключевые слова — 3D-изображение, 3D-сканирование, карта глубины.

Введение

При наблюдении трехмерных объектов и сцен, несмотря на то, что в их проекциях на сетчатки глаз координата глубины оказывается утраченной, мы, тем не менее, воспринимаем их объемными. Объясняется это тем, что в этих проекциях, которые представляют собой 2D-изображения, содержится информация, используя которую совместно с априорной информацией, имеющейся у зрителя о наблюдаемых объектах, зрительная кора в той или иной мере восстанавливает утраченную при проецировании информацию о координате глубины. Примерами информации, используемой зрительной корой для определения (оценки) утраченной координаты глубины, являются:

— величина искажений геометрических размеров наблюдаемых объектов, возникающих в результате их центральной проекции на сетчатку глаз;

— перекрытие объектами, близко расположенными к зрителю, объектов, которые расположены от зрителя на больших расстояниях;

— распределение полутеней на криволинейных поверхностях объектов, обусловленное эффектом диффузного отражения света;

— расположение объектов относительно линии горизонта;

— наличие атмосферной дымки и пр.

Эти свойства зрения изучались и использовались в архитектуре и живописи, начиная уже с античных времен. Колонны храма Парфенона,

форма которых намеренно выбрана отличной от цилиндрической, являются примером одного из первых практических применений законов перспективы в архитектуре. Интенсивное развитие живописи в эпоху Возрождения также в значительной степени связано с использованием перечисленных выше свойств зрения.

В настоящее время возрос интерес к 3D-технологиям и, в частности, к методам получения трехмерных изображений, что, естественно, проявилось в появлении множества публикаций на эту тему, например [1–7], а затем и в разработке 3D-сканеров, 3D-кинотеатров и 3D-телевидения.

В последних двух случаях речь идет по существу о стереоскопических системах. В связи с тем, что изначально съемка 3D-фильмов является весьма дорогостоящим предприятием, появились попытки разработать методы, позволяющие преобразовывать обычные 2D-контенты в 3D-контенты, основанные на использовании перечисленных выше свойств зрения. Так, в университете Карнеги — Меллона был разработан метод преобразования 2D-изображений в 3D-сцену, подробности которого не разглашаются, но известно, что 3D-изображение строится с использованием только вертикальных и горизонтальных поверхностей, обнаруженных на 2D-изображении [8]. Были также разработаны методы, примененные в настоящее время в телевизорах, в которых использованы свойства взаимного перекрытия объектов переднего и последующего планов, а также разница в цвете и контрасте

между изображениями объектов различных планов. Однако качество получаемых при этом изображений невелико. В данной статье рассматривается метод формирования 3D-изображения сцены и составляющих ее объектов, которые можно аппроксимировать плоскостями, по одной 2D-фотографии.

Перспективные преобразования прямых линий

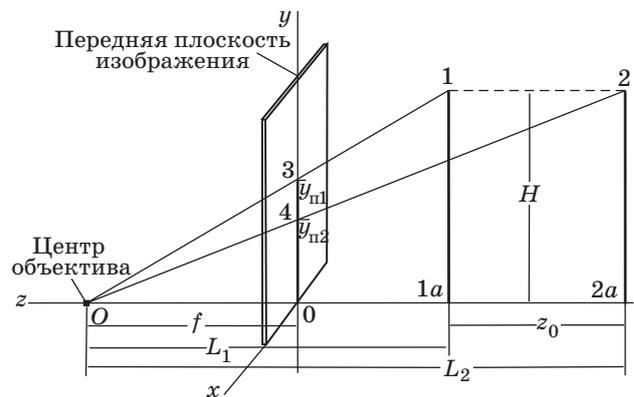
Приведем свойства перспективных преобразований для прямых линий.

1. В общем случае прямая линия на объекте проецируется в прямую линию на изображении центральной проекции, но с *неравномерной шкалой*. Например, движение точки с постоянной скоростью по прямой линии на объекте отображается ее движением на изображении тоже по прямой линии, но с неравномерной скоростью. Чем больше точка удаляется, тем меньше становится видимая скорость ее движения. Эта неравномерность шкалы приводит к тому, что при перспективном преобразовании не сохраняется соотношение длин и площадей в пространстве и в центральной проекции.

2. Параллельные прямые линии в пространстве изображаются в центральной проекции прямыми линиями, сходящимися в одной точке, так называемой точке схода. Если эти линии параллельны горизонтальной плоскости, то их точка схода располагается на линии горизонта. Если они параллельны наклонной восходящей плоскости, то точка их схода располагается выше линии горизонта. Если же они параллельны наклонной нисходящей плоскости, то их точка схода располагается ниже линии горизонта.

Исключением из этого правила является случай, когда прямые линии в пространстве параллельны плоскости, на которую они проецируются. При этом на изображении эти линии также остаются параллельными и не имеют точки схода, а шкала проекции остается равномерной.

Продолжая рассмотрение, выберем правостороннюю систему координат и обратимся к рис. 1, на котором показана так называемая модель камеры с передней плоскостью изображения. Рассмотрим центральную проекцию двух простейших объектов, изображенных на этом рисунке (линий 1-1а и 2-2а), на плоскость, при этом координатную ось z декартовой системы координат совместим с осью камеры, а оси x и y — с плоскостью изображения. На рисунке использованы следующие обозначения: H — длина линий, одинаковая для обеих линий; L_1 — расстояние между линией 1-1а и центром объектива; L_2 — расстояние между линией 2-2а и центром объектива;



■ Рис. 1. Модель камеры с передней плоскостью изображения

f — фокусное расстояние объектива; z_0 — расстояние между линиями 1-1а и 2-2а; $y_{п1}$ — длина проекции линии 1-1а на переднюю плоскость изображения (плоскость проецирования); $y_{п2}$ — длина проекции линии 2-2а.

Из подобия треугольников 1-1а-О и 3-0-О следует, что

$$y_{п1}/f = H/L_1, \quad (1)$$

а из подобия треугольников 2-2а-О и 4-0-О следует, что

$$y_{п2}/f = H/L_2. \quad (2)$$

В общем случае длина проекции $y_{п}$ линии протяженностью H , которая параллельна плоскости проецирования независимо от ее углового положения по отношению к координатным осям x , y и удалена от центра объектива на расстояние L , равна

$$y_{п} = Hf/L, \quad (3)$$

где f/L можно рассматривать как масштабирующий множитель. Отсюда следует, что любая точка объекта с координатами $x_{об}$, $y_{об}$ отобразится на центральной проекции точкой с координатами $x_{п}$, $y_{п}$:

$$x_{п} = x_{об} f/L; \quad y_{п} = y_{об} f/L. \quad (4)$$

Из формулы (3), в частности, следует, что при $L = f$ длина проекции линии равна длине самой проецируемой линии. Решая совместно уравнения (1) и (2), найдем расстояние по глубине между двумя линиями 1-1а и 2-2а

$$z_0 = L_1 \frac{y_{п1} - y_{п2}}{y_{п2}}. \quad (5)$$

Рассмотрев основные аспекты перспективного преобразования прямых линий, можно перейти к перспективным преобразованиям плоского прямоугольника.

Перспективные преобразования плоского прямоугольника

Рассмотрим теперь задачу определения истинных размеров одного из простейших объектов, которым является плоский прямоугольник, по его центральной проекции. Будем считать, что две его стороны вертикальны и, следовательно, параллельны плоскости проекции, которая перпендикулярна горизонтальной плоскости, а плоскость, в которой он лежит, повернута на некоторый угол α_0 относительно направления главного луча. В этом случае проекция прямоугольника будет иметь вид трапеции, поскольку размер проекции удаленной стороны вследствие перспективных искажений будет меньше размера ближайшей стороны. Примем также, что нам известны фокусное расстояние объектива f , использованного при фотографировании, расстояние L_1 , с которого производилось фотографирование, а также размеры проекций ближайшей и удаленной сторон, соответственно $y_{п1}$ и $y_{п2}$, которые определяются по изображению.

Переходя к определению размеров прямоугольника, найдем размер его вертикальных сторон: $H = y_{п1}L_1/f$, что следует из формулы (1). Для определения размера его горизонтальных сторон, который обозначим $L_{об}$, найдем вначале расстояние по глубине z_0 между этими сторонами по формуле (5). Поскольку плоскость прямоугольника составляет с главным лучом угол α_0 , то глубина представляет собой размер проекции его ширины на вертикальную плоскость, в которой лежит главный луч. Отсюда следует, что

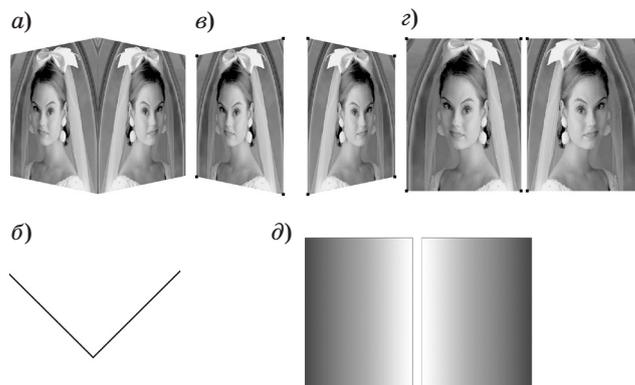
$$L_{об} = z_0 / \cos \alpha_0. \quad (6)$$

Опуская вследствие громоздкости вычислений рассмотрение случая, когда проецируемой фигурой является плоский многоугольник, стороны которого не обязательно параллельны плоскости, на которую он проецируется, отметим лишь, что методика анализа и в этом случае полностью сохраняется.

Рассмотренный случай, несмотря на свою простоту, важен при решении ряда практических задач, поскольку на практике часто встречаются объекты, оболочки которых могут быть аппроксимированы плоскостями прямоугольной формы. К таким объектам относятся наружные стены зданий, стены интерьеров, предметы мебели, а также ряд других объектов.

Метод формирования 3D-изображения объекта по его двумерной фотографии

Пояснение метода проведем на примере несложного объекта, представляющего собой два



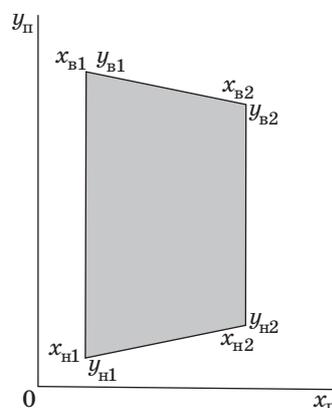
■ **Рис. 2.** К пояснению принципа формирования 3D-изображения объекта по его двумерной фотографии: *a* — вид на объект спереди (его фотография); *б* — вид сверху; *в* — две аппроксимирующие плоскости; *г* — аксонометрические проекции каждой из аппроксимирующих плоскостей; *д* — карты глубины каждой из двух аппроксимирующих плоскостей

соединенных прямоугольных рекламных щита (рис. 2, *a*, *б*). Будем считать, что вертикальные линии прямоугольных щитов параллельны плоскости проецирования, а расстояние между ближайшей точкой объекта, т. е. местом, где щиты соединяются, и центром объектива с фокусным расстоянием f равно L_1 .

Суть метода заключается в следующем.

1. Вначале выполняется декомпозиция изображения объекта на две аппроксимирующие плоскости (рис. 2, *в*).

2. Далее для каждой из аппроксимирующих плоскостей производится пересчет от перспективной проекции к аксонометрической проекции. Для пояснения того, как это делается, обратимся к рис. 3, на котором показана центральная проекция правой плоскости щита, изображенного на рис. 2, *в*. Из рис. 3 видно, что размер центральной проекции дальней вертикальной гра-



■ **Рис. 3.** Центральная проекция прямоугольной плоскости, линии вертикальных границ которой параллельны плоскости изображения

ницы прямоугольного щита $y_{в2} - y_{н2}$ меньше размера центральной проекции его ближней вертикальной границы $y_{в1} - y_{н1}$. Величина этого изменения составляет $K_{\min} = (y_{в2} - y_{н2}) / (y_{в1} - y_{н1})$.

В такой же мере имеет место уменьшение и горизонтального размера центральной проекции прямоугольного щита, хотя это сразу и не бросается в глаза. Из рис. 3 также следует, что коэффициент изменения координат K линейно изменяется с изменением координаты проекции $x_{п}$, поскольку уменьшению $x_{п}$ соответствует уменьшение расстояния между рассматриваемой точкой и центром объектива, т. е.

$$K = K_{\min} - (K_{\min} - 1) \frac{x_{в2} - x_{п}}{x_{в2} - x_{в1}}. \quad (7)$$

Поэтому при определении координат аксонометрической проекции каждой точки объекта на переднюю плоскость изображения x_a и y_a необходимо это учесть следующим образом:

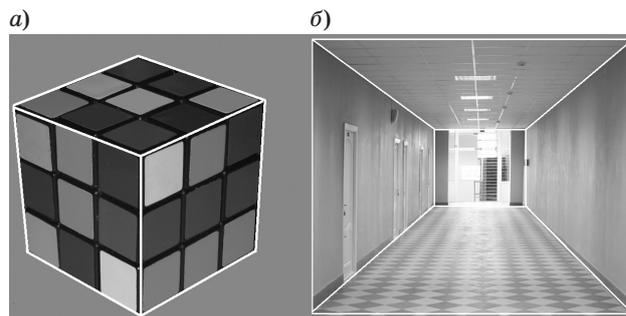
$$\begin{aligned} y_0 &= (y_{в1} + y_{н1}) / 2; \quad x_a = x_{в1} + (x_{п} - x_{в1}) / K; \\ y_a &= y_0 + (y_{п} - y_0) / K. \end{aligned} \quad (8)$$

В результате этих расчетов находится аксонометрическая проекция объекта в масштабе, который определяется масштабирующим множителем f/L_1 . Координаты $x_{об0}$ и $y_{об0}$ произвольных точек объекта находятся путем деления координат x_a и y_a на масштабирующий множитель f/L_1 . Полученные таким образом аксонометрические проекции каждой из аппроксимирующих плоскостей представлены на рис. 2, з.

3. Затем определяется расстояние по глубине z_0 между вертикальными линиями, ограничивающими прямоугольные аппроксимирующие плоскости, по формуле (5). Другими словами, определяется расстояние по глубине между двумя плоскостями, параллельными передней плоскости изображения, в которых лежат эти линии.

4. После этого необходимо определить координату глубины для каждой точки аппроксимирующей плоскости, т. е. рассчитать карту глубины. С этой целью вначале методом линейной интерполяции находятся координаты глубины вдоль прямых линий, соединяющих вершины аппроксимирующих плоскостей, для которых координаты глубины известны (в рассматриваемом примере, поскольку плоскости вертикальны, для всех точек вертикальных линий координата глубины будет неизменной). Линейная интерполяция оказывается возможной, поскольку в аксонометрической проекции шкала «дальности», в отличие от шкалы «дальности» центральной проекции, равномерная.

5. Далее методом линейной интерполяции вычисляются координаты глубины для всех точек (пикселей) аппроксимирующей плоскости и та-



■ Рис. 4. Изображения объектов, для аппроксимации которых необходимо использовать 3 (а) и 5 (б) плоскостей

ким образом формируются карты глубины каждой из двух аппроксимирующих плоскостей (рис. 2, д). На картах глубины более светлый тон соответствует точкам, находящимся ближе к центру объектива, а более темный — точкам, расположенным дальше.

6. И, наконец, завершающим шагом описываемого метода является «сборка» объекта путем объединения аппроксимирующих его плоскостей.

В зависимости от сложности и количества объектов, изображенных на фотографии, может потребоваться большее число аппроксимирующих плоскостей, чем в рассмотренном примере (рис. 4, а, б).

Найденные аксонометрическая проекция объекта и карта глубины содержат всю информацию о *видимой части* 3D-объекта и позволяют вывести его изображение на экран дисплея для рассматривания при различных ракурсах и с различных расстояний. Для получения полного 3D-изображения объекта, а не только видимой его части, как и в случае традиционных 3D-сканеров, возникает необходимость в склейках частей 3D-изображений, полученных при различных ракурсах съемки. Обусловлено это тем, что ни один непрозрачный физический объект невозможно одновременно видеть со всех сторон.

Представление 3D-изображения в виде аксонометрической проекции и карты глубины, так называемый 2D+Z-формат, применяют в 3D-телевидении (точнее, в стереоскопическом телевидении). Этот формат может быть преобразован в другие форматы 3D-изображений.

Отображение 3D-изображения на экране дисплея

Покажем теперь, как, располагая аксонометрической проекцией объекта и картой глубины, вывести на экран дисплея центральную проекцию объекта при заданном ракурсе и расстоянии наблюдения. Задача заключается в определении

координат каждой точки формируемого изображения 3D-объекта, соответствующего новому ракурсу наблюдения этого объекта и новому расстоянию, с которого он наблюдается. Будем считать известными расстояние, с которого была произведена съемка L_1 , и фокусное расстояние объектива f . Решая задачу, вначале находим координаты объекта, затем осуществляем поворот объекта на заданные углы и, наконец, находим центральную проекцию объекта, видимого с заданного расстояния наблюдения.

Определение декартовых координат x и y точек объекта по его аксонометрической проекции. Как уже было отмечено, декартовы координаты произвольной точки объекта x_{o60}, y_{o60} , соответствующие его исходному положению, находятся путем деления координат x_a и y_a на масштабирующий множитель f/L_1 , т. е.

$$x_{o60} = x_a L_1/f; y_{o60} = y_a L_1/f. \quad (9)$$

Что касается координаты глубины произвольной точки объекта z_{o60} , то она полностью определяется его картой глубины.

Определение декартовых координат точек объекта, изменившихся в результате его поворота вокруг координатных осей x и y на заданные углы. Поворот объекта вокруг оси z в рассматриваемом случае интереса не представляет, так как его вращение в плоскости проекции не приводит к изменению ракурса наблюдения. Возьмем точку на поверхности объекта с координатами $x_{o60}, y_{o60}, z_{o60}$, повернем объект сначала на угол α вокруг оси y и найдем его новые координаты $x_{o61}, y_{o61}, z_{o61}$:

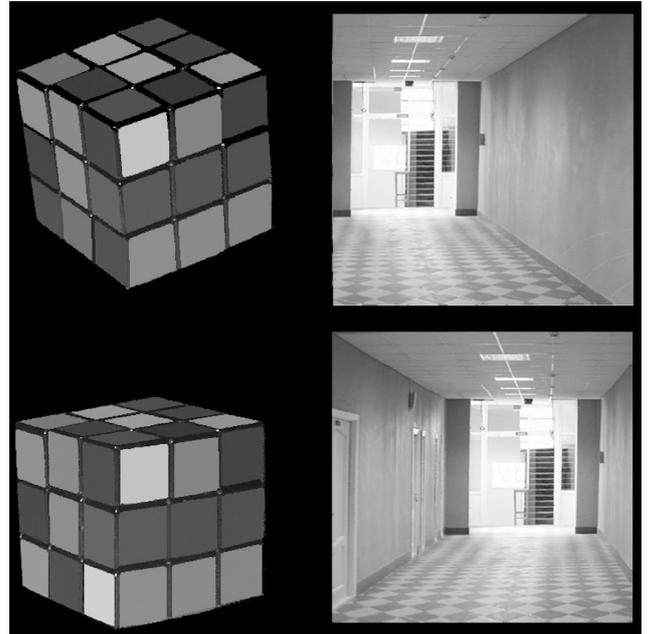
$$\begin{aligned} x_{o61} &= x_{o60} \cos \alpha - z_{o60} \sin \alpha; y_{o61} = y_{o60}; \\ z_{o61} &= z_{o60} \cos \alpha + x_{o60} \sin \alpha. \end{aligned} \quad (10)$$

Допустимое значение угла α ограничивается значениями, при которых изображение повернутого 3D-объекта еще не включает в себя тех фрагментов, которые при его фотографировании были заслонены самим объектом. Далее, возьмем эту же точку на поверхности объекта после его поворота по азимуту на угол α вокруг оси y и найдем ее новые координаты $x_{o62}, y_{o62}, z_{o62}$ после поворота объекта вокруг оси x на угол места β :

$$\begin{aligned} x_{o62} &= x_{o61}; y_{o62} = y_{o61} \cos \beta - z_{o61} \sin \beta; \\ z_{o62} &= y_{o61} \sin \beta + z_{o61} \cos \beta. \end{aligned} \quad (11)$$

Подставив значения $x_{o61}, y_{o61}, z_{o61}$ из формулы (10) в формулу (11), получим

$$\begin{aligned} x_{o62} &= x_{o60} \cos \alpha - z_{o60} \sin \alpha; \\ y_{o62} &= y_{o60} \cos \beta - z_{o60} \cos \alpha \sin \beta - x_{o60} \sin \alpha \sin \beta; \\ z_{o62} &= y_{o60} \sin \beta + z_{o60} \cos \alpha \cos \beta + x_{o60} \sin \alpha \cos \beta. \end{aligned} \quad (12)$$



■ Рис. 5. Примеры проекций 3D-изображений, полученных описанным методом

Переходя к заключительной операции вычислений, найдем декартовы координаты $x_{пр}$ и $y_{пр}$ центральной проекции объекта на экран дисплея. Согласно формуле (4), между искомыми координатами центральной проекции объекта и соответствующими координатами объекта имеют место соотношения

$$x_{пр} = x_{o62} f/L; y_{пр} = y_{o62} f/L, \quad (13)$$

где L — расстояние от центра объектива до точки объекта, координаты которой рассчитываются. Это расстояние равно сумме расстояний: расстояния L_n между центром объектива и ближайшей к камере точкой объекта и расстояния $z_{гл}$ от ближайшей точки к центру объектива до рассматриваемой точки, которое определяется из карты глубины, т. е. $L = L_n + z_{гл}$.

С учетом изложенного имеем

$$x_{пр} = x_{o62} \frac{f}{L_n + z_{гл}}; y_{пр} = y_{o62} \frac{f}{L_n + z_{гл}}. \quad (14)$$

Проекции 3D-изображений (см. рис. 4), полученные описанной выше обработкой 2D-изображений, показаны на рис. 5.

Заключение

Описанный метод может найти свое применение в тех случаях, когда требуется, используя цифровую камеру, получать 3D-изображения отдельных объектов и целых сцен, которые допускают аппроксимацию плоскостями. Это может

быть актуальным, если принять во внимание, что стоимость лазерного 3D-сканера, используемого для этой цели, в сотни раз превышает стоимость недорогой цифровой камеры.

Поскольку в результате применения описанного выше метода из 2D-изображений формируются 3D-изображения, то не составляет труда сделать следующий шаг и из 3D-изображений получать их стереопары. Опуская вследствие ограниченности объема статьи рассмотрение метода формирования стереопары из 3D-изображения, приведем в качестве иллюстрации лишь сами полученные в процессе выполнения данной работы стереоскопические изображения. На сайте [9] показаны анаглифные изображения, сформирован-

ные в конечном итоге из 2D-изображений, которые приведены в тексте настоящей статьи. Анаглифный метод представления стереоскопических изображений для данной демонстрации выбран по причине его доступности. Напомним, что для рассматривания такого рода стереоскопических изображений требуются анаглифные очки, которые в настоящее время чрезвычайно широко распространены и недороги.

Описанный в статье метод может быть также использован как элемент алгоритма, предназначенного для реализации функции преобразования 2D-изображений в стереоскопические изображения, которую начинают применять в современных телевизорах [10].

Литература

1. Horry Y., Anjo K., Arai K. Tour into the picture: using a spidery mesh interface to make animation from a single image // ACM SIGGRAPH Proc. 1997. P. 225–232.
2. Shum H.-Y., Han M., Szeliski R. Interactive construction of 3D models from panoramic mosaics // IEEE Conf. on CVPR, June 1998. P. 427–433.
3. Horn B. K. P. Height and gradient from shading // Int'l J. of Computer Vision. 1990. Vol. 5. N 1. P. 37–75.
4. Красильников Н. Н. Метод получения 3D-изображений, основанный на диффузном отражении света сканируемыми объектами // Информационно-управляющие системы. 2009. № 6(43). С. 7–11.
5. Красильников Н. Н., Красильникова О. И. Получение трехмерного изображения объекта путем изменения интенсивности диффузного отражения света различными точками его поверхности // Оптический журнал. 2010. Т. 77. № 6. С. 19–24.
6. Красильников Н. Н., Красильникова О. И. Исследование погрешностей определения координаты глубины при 3D-сканировании методом, основанном на диффузном отражении света // Информационно-управляющие системы. 2012. № 3(58). С. 2–8.
7. Красильников Н. Н. Цифровая обработка 2D- и 3D-изображений: учеб. пособие для вузов. — СПб.: БХВ-Петербург, 2011. — 608 с.
8. <http://www.techeblog.com/index.php/tech-gadget/feature-university-researchers-develop-method-to-convert-2d-images-into-3d-scenes-video> (дата обращения: 02.04.2013).
9. http://www.i-us.ru/authors/krasilnikov_nn.
10. http://3dtv-obzor.ru/2d_3d_convertacia (дата обращения: 02.04.2013).