

ОПРЕДЕЛЕНИЕ КООРДИНАТЫ ГЛУБИНЫ ПО 2D-ИЗОБРАЖЕНИЮ

© 2011 г. Н. Н. Красильников, доктор техн. наук; О. И. Красильникова, канд. техн. наук

Санкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург

E-mail: nnk_k23@aanet.ru

Рассмотрена проблема определения координаты глубины расположения объектов в трехмерной сцене по ее двумерному изображению путем использования дополнительной априорной информации, как правило, имеющейся у зрителя.

Ключевые слова: 2D-изображения, 3D-сканирование, координата глубины.

Коды OCIS:100.6890, 3000.30200.

Поступила в редакцию 23.03.2011.

Введение

При рассматривании трехмерной сцены и составляющих ее объектов информацию о ней зритель получает из ее проекций на сетчатки глаз, которые по своей природе являются двумерными. Однако несмотря на это, в зрительной коре головного мозга формируется 3D-представление о наблюдаемой сцене и составляющих ее объектах [1]. При этом возникает кажущийся парадокс. С одной стороны, изображения на сетчатках глаз непосредственно не содержат данных о координатах глубины наблюдаемых объектов, поскольку являются двумерными, а с другой стороны, эти координаты каким-то образом восстанавливаются при формировании их представления в зрительной коре.

Обращаясь к анализу этой проблемы, сразу же исключим из рассмотрения бинокулярный фактор наблюдения, т. е. стереоэффект, поскольку при наблюдении объектов одним глазом, а также при наблюдении 2D-фотографий и картин зрение продолжает восстанавливать третью, утраченную координату. Заметим также, что ряд животных и птиц, например, лошади, зайцы и др., вообще лишены стереоскопического зрения, поскольку поля зрения левого и правого глаз у них не перекрываются, но несмотря на это, они ориентируются в пространстве. Можно также исключить из последующего рассмотрения различие в фокусировке глаз при наблюдении объектов, расположенных

на различных расстояниях, как несущественное, тем более, что оно проявляется на сравнительно небольших расстояниях. Из этого, конечно, не следует, что перечисленные выше факторы вообще не играют никакой роли при нахождении третьей координаты, однако они не являются определяющими.

Анализ проблемы приводит к выводу, что кажущийся парадокс разрешается, если учесть наличие априорной информации, которой располагает зритель при "десифровке" 2D-изображений наблюдаемой сцены.

Для пояснения сказанного обратимся к рис. 1а, на котором изображена городская улица. При рассматривании этого изображения у зрителя не возникает каких-либо затруднений в определении ни расстояний до изображенных на нем объектов, ни их размеров. Объясняется это тем, что при "десифровке" наблюдаемого изображения зрительная система использует дополнительную информацию, заключающуюся в том, что все объекты, изображенные на этом рисунке, не парят в пространстве, а привязаны к плоскости проезжей части улицы и к плоскости тротуара. Здесь действует простое правило: объект, который изображен на большем расстоянии от линии горизонта вниз, находится ближе к наблюдателю, а объект, который изображен ближе к линии горизонта, находится дальше от наблюдателя. В общем случае поверхность, на которой находятся наблюдаемые объекты, может иметь более сложный вид, как, например,



Рис. 1. Изображения городской улицы (а) и руин античного храма, расположенного на холме (б).

это показано на рис. 1б, но это правило по-прежнему остается справедливым. Кроме того, при определении расстояний до наблюдаемых объектов зрительной системой может также использоваться и другая информация, часть которой зрителю обычно известна, например, информация о размерах наблюдаемых объектов: высоте фонарных столбов, автомобилей, среднем росте людей и т. д. Принцип определения расстояний до объектов, составляющих наблюдаемую сцену, который используется зрительной системой, может быть также использован и для получения 3D-изображения на основании 2D-изображения наблюдаемой сцены, полученного путем его простого фотографирования обычной цифровой камерой. Этот принцип позволяет отказаться от использования для этой цели дорогостоящего лазерного сканера.

Определение расстояния до объекта сцены по его 2D-изображению

Рассмотрим ряд простых, но достаточно типичных примеров определения расстояния до объекта по его 2D-изображению.

1. Вначале обратимся к простейшему, но часто встречающемуся случаю, когда направление на фотографируемый объект совпадает с оптической осью камеры, а истинная высота объекта $Z_{об}$ и фокусное расстояние объектива камеры f , посредством которой получено изображение, известны. В этом случае для

вычисления расстояния $X_{об}$ до объекта достаточно определить вертикальный размер проекции этого объекта $Z_{пр}$ на светочувствительную поверхность камеры (передней плоскости изображения) и воспользоваться известным соотношением [2]

$$Z_{пр} = \frac{Z_{об}f}{X_{об} + f},$$

решив которое относительно $X_{об}$, находим

$$X_{об} = f \left(\frac{Z_{пр}}{Z_{об}} - 1 \right).$$

2. Переходя к рассмотрению второго случая, будем считать, что 2D-изображение представляет собой изображение сцены, заполненной объектами, расстояния до которых требуется определить, при этом размеры самих объектов неизвестны. Известно лишь, что эти объекты расположены на плоской горизонтальной поверхности, а также известна высота H , с которой проведена съемка, фокусное расстояние объектива f и положение линии горизонта. Для решения этой задачи обратимся к рис. 2, на котором приведена оптическая модель камеры с передней плоскостью изображения. Для конкретности дальнейшего изложения под расстояниями до объектов будем понимать расстояния между центром объектива и одной из точек, в которых объекты соприкасаются с горизонтальной плоскостью. В рассматриваемом примере будем искать расстояние L до точки A .

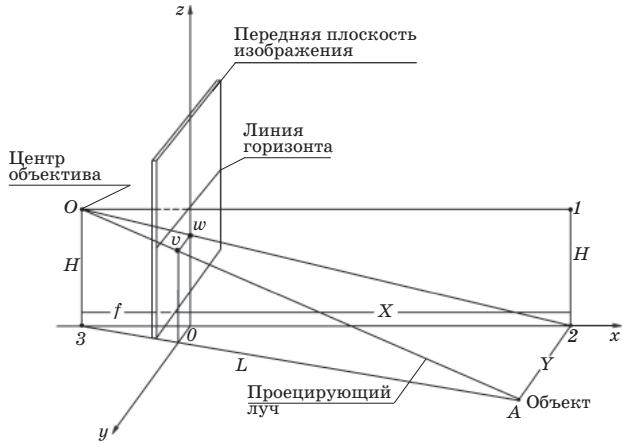


Рис. 2. Оптическая модель камеры с передней плоскостью изображения.

Из прямоугольного треугольника, образованного проекцией проецирующего луча OA на координатную плоскость xz , т. е. линии $O-2$, а также линии $O-1$ и линии $1-2$, следует, что

$$\frac{H}{w} = \frac{X+f}{f}. \quad (1)$$

Из рассмотрения прямоугольника, образованного проекцией проецирующего луча OA на координатную плоскость xy , т. е. линии $A-3$, а также линии стороны $2-3$ и стороны $A-2$, следует, что

$$\frac{Y}{v} = \frac{X+f}{f}. \quad (2)$$

Кроме того, как нетрудно видеть,

$$L^2 = Y^2 + (X+f)^2. \quad (3)$$

Решая уравнения (1)–(3) относительно L , найдем

$$L = \frac{H}{w} \sqrt{v^2 + f^2}, \quad (4)$$

где w – расстояние на передней плоскости изображения между проекцией линии горизонта и проекцией точки A , v – расстояние на передней плоскости изображения между вертикальной линией, которая делит изображение пополам, и проекцией точки A . При $L \rightarrow \infty$ имеем $w \rightarrow 0$, т. е. бесконечно удаленный объект изображается на линии горизонта. Как правило, размер изображения на фотографии, экране монитора и т. д. отличается от размера изображения в передней плоскости, приведенной на рисунке модели камеры. Кроме того, при регистрации изображения на мониторе

эти размеры обычно выражаются в пикселях. Поэтому преобразуем формулу (4) к более удобному виду

$$L = \frac{H}{aw_i} \sqrt{a^2 v_i^2 + f^2}, \quad (5)$$

где $w = aw_i$, $v = av_i$, а w_i и v_i – координаты изображения точки A на изображении, по которому проводятся измерения (например на экране монитора компьютера), a – коэффициент, согласующий размерности.

3. Рассмотрим теперь случай, когда фокусное расстояние объектива камеры неизвестно. С таким случаем мы можем встретиться при съемке камерой, имеющей объектив с переменным фокусным расстоянием (трансфокатор). Для решения задачи при данных условиях необходимо и достаточно знать расстояние L_1 до одного из изображенных объектов (более конкретно, расстояния до точки A соприкосновения одного из объектов с горизонтальной плоскостью, на которой он расположен) и тогда расстояния до других объектов могут быть найдены. Задача решается следующим образом. Вначале необходимо найти неизвестное фокусное расстояние объектива. Решая уравнение (5) относительно фокусного расстояния f и делая подстановку в найденную формулу известных значений высоты H , с которой проведена съемка, известных значений координат точки проекции w_{i1} и v_{i1} , а также значения L_1 и a , находится сначала фокусное расстояние

$$f = \sqrt{\frac{L_1^2 a^2 w_{i1}^2}{H^2} - a^2 v_{i1}^2},$$

а затем, используя формулу (5), определяются расстояния до других объектов.

В заключение отметим, что если мы располагаем фотографией, для которой вообще неизвестны условия съемки, т. е. неизвестны H и f , а также положение линии горизонта, то для ее “дешифровки” потребуется априорное значение расстояний (или размеров) не менее трех объектов, расположенных на различных расстояниях до снимающей камеры.

Аналогичным образом решаются задачи применительно к случаю, когда поверхность, на которой расположены объекты, не является плоской и горизонтальной.

Экспериментальная проверка

На рис. 3 показано изображение трехмерной сцены, полученное с помощью программы

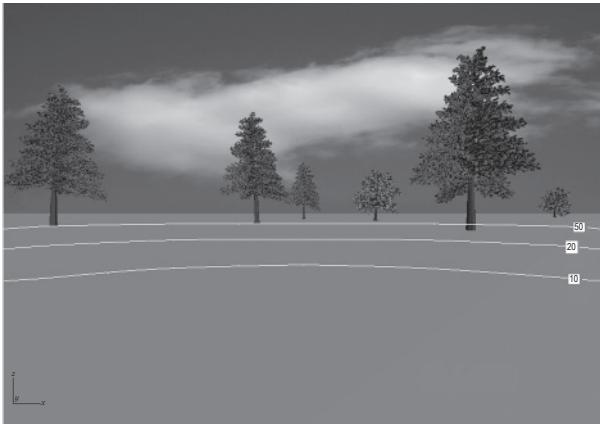


Рис. 3. Изображение трехмерной сцены, полученное с помощью программы 3D Studio MAX.

3D Studio MAX при использовании объектива с фокусным расстоянием $f = 20$ мм, при высоте съемки $H = 1,5$ м. По этому изображению описанным выше способом по формуле (5) были рассчитаны расстояния до изображенных на ней деревьев. Результаты измерения для деревьев (слева направо) составили: 51,9, 53,8, 100, 74,3, 34,1 и 123,3 м. Истинные расстояния при этом были 50, 50,5, 100, 71,6, 33,5 и 150 м соответственно. Расхождения между истинными и измеренными расстояниями объясняются недостаточным количеством строк в изображении, по которому проводилось измерение. Другими словами, погрешность измерения обусловлена дискретностью изображения, использованного при измерениях. Снизить погрешность измерения можно путем увеличения частоты пространственной дискретизации изображения при съемке.

Для удобства практических измерений на изображение, по которому они выполняются, рекомендуется наносить линии постоянной дальности. Формула для расчета линий постоянной дальности может быть получена из формулы (5) путем ее решения относительно w_n

$$w_n = \frac{H}{aL} \sqrt{a^2 v_n^2 + f^2}.$$

Для построения линий постоянной дальности необходимо, задав значения L , H и f , изменять координату v_n и для каждого его значения определять координату w_n . На рис. 3 показаны эти линии для ряда значений L .

Заключение

Таким образом, априорная информация о наблюдаемой сцене и составляющих ее объектах, которой обычно располагает зритель, позволяет зрительной системе извлекать дополнительные данные об объектах, которые непосредственно в двумерном изображении не содержатся. Помимо информации о размерах объектов и положения их изображений относительно линии горизонта, к такой полезной информации относится информация о параллельности и перпендикулярности прямых линий в изображениях объектов.

Кроме того, по распределению светотеней на изображениях объектов, как показано в [3], можно судить об их протяженности в 3-х измерениях.

Благодаря этому в ряде случаев по двумерному изображению оказывается возможным реконструировать 3D-изображение видимой части сцены.

* * * * *

ЛИТЕРАТУРА

1. *Marr D.* Зрение (Информационный подход к изучению представления и обработки зрительных образов). М.: Радио и связь, 1987. 400 с.
2. *Дуда Р., Харт Р.* Распознавание образов и анализ сцен. М.: Мир, 1976. 511 с.
3. *Красильников Н.Н., Красильникова О.И.* Получение трехмерного изображения объекта путем измерения интенсивности диффузного отражения света различными точками его поверхности // Оptический журнал. 2010. Т. 77. № 6. С. 19–24.