
ЧАСТЬ I

Формирование изображений и модели
изображений

Камеры

Существует множество устройств формирования изображения: от глаз животных до видеокамер и радиотелескопов. Не все эти устройства могут содержать линзы. Например, у первой модели камеры-обскуры (что буквально означает “темная комната”), изобретенной в XVI веке, линз не было, вместо них использовалось отверстие, которое позволяло лучам света фокусироваться на стенке или полупрозрачной пластинке и таким образом демонстрировать законы перспективы, которые за сто лет до этого открыл Брунеллеччи (Brunelleschi). Уже в 1550 году отверстия стали заменять все более и более усложняющимися линзами, а современная цифровая или фотокамера — это, по сути, камера-обскура, способная регистрировать количество света, которое попадает на каждый маленький участок ее задней панели (рис. 1.1).

Поверхность камеры, на которой формируется изображение, обычно представляет собой прямоугольник, а форма человеческой сетчатки намного ближе к сферической поверхности, поэтому панорамные фотоаппараты могут оборудоваться цилиндрическими чувствительными поверхностями. Помимо формы, датчики изображения имеют и другие характеристики. Они могут фиксировать изображение, дискретное в пространстве (как, например, в человеческом глазе с палочками и колбочками, 35-миллиметровых камерах с эмульсионными зернами и цифровых камерах с прямоугольными элементами изображений, или

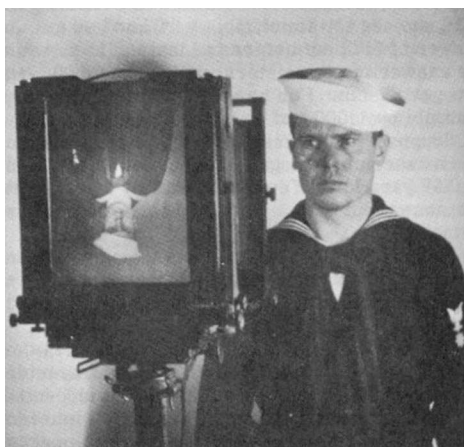


Рис. 1.1. Формирование изображения на задней панели фотокамеры. Рисунок из учебника *“US NAVY MANUAL OF BASIC OPTICS AND OPTICAL INSTRUMENTS”*, подготовленного Комитетом по управлению военно-морским персоналом США, перепечатан корпорацией *Dover Publications* (1969)

пикселями) или непрерывное (как, например, в старых телевизионных трубках). Сигнал, фиксируемый датчиком изображения в определенной точке своей чувствительной поверхности, может быть дискретным или непрерывным и состоять из одного числа (черно-белая камера), нескольких значений (например, интенсивности красного, зеленого и синего цветов для цветной камеры или реакции трех типов колбочек для человеческого глаза), большого набора чисел (например, реакции гиперспектральных сенсоров) или даже непрерывной функции длины волны (спектрометры). Изучению перечисленных выше характеристик и посвящена данная глава.

1.1. КАМЕРЫ-ОБСКУРЫ

1.1.1. Перспективная проекция

Возьмем коробку, в одной из ее стенок сделаем булавкой маленькое отверстие, а затем заменим противоположную стенку полупрозрачной пластинкой. Если держать эту коробку перед собой в комнате со слабым освещением так, чтобы отверстие было направлено на какой-то источник света (скажем, свечу), то на полупрозрачной пластинке можно увидеть перевернутое изображение свечи (рис. 1.2). Это изображение образуют лучи света, идущие от сцены, находящейся перед коробкой. Если отверстие уменьшить до точки (что, конечно же, физически невозможно), то через плоскость пластинки (или плоскость изображения), отверстие и некоторую точку сцены будет проходить ровно один луч.

В действительности, отверстие имеет конечный (хотя и небольшой) размер, и в каждой точке плоскости изображения собирается свет от целого конуса лу-

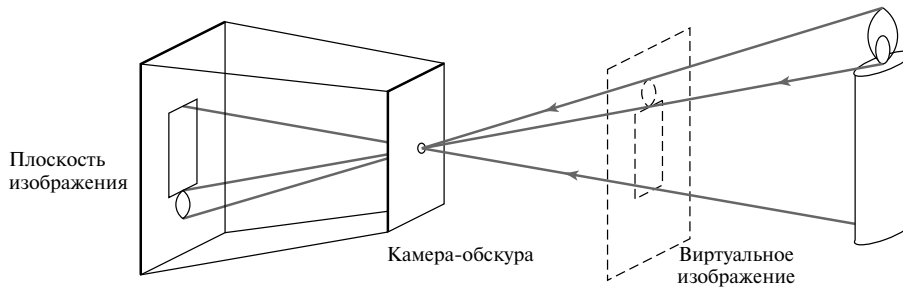


Рис. 1.2. Модель формирования изображения через отверстие

чей, образующих определенный телесный угол, так что эта идеализированная и чрезвычайно упрощенная геометрическая модель формирования изображения не совсем строго описывает действительность. Кроме того, настоящие камеры обычно оснащены линзами, что также весьма усложняет ситуацию. Тем не менее, модель точечной перспективы (ее еще называют центральной перспективой), впервые предложенная Брунеллеччи в начале XV века, удобна с математической точки зрения. Несмотря на свою простоту, она часто является приемлемым приближенным описанием процесса формирования изображения. В результате перспективной проекции возникает перевернутое изображение, поэтому иногда вместо него удобно рассматривать мнимое изображение, расположенное в плоскости, лежащей перед отверстием на таком же расстоянии, как и реальная плоскость изображения (рис. 1.2). Это мнимое изображение не перевернуто, но во всем остальном оно точно соответствует настоящему. В зависимости от ситуации, удобнее может быть одно или другое. На рис. 1.3, *a* показан очевидный эффект перспективной проекции: относительный размер объектов зависит от расстояния до них. Например, изображения B' и C' столбиков B и C кажутся одинаковыми по высоте, а в действительности столбики A и C в два раза меньше, чем B . На рис. 1.3, *б* иллюстрируется еще один общеизвестный эффект: проекции двух параллельных прямых, лежащих в некоторой плоскости Π , сходятся на горизонтальной линии H , образованной пересечением плоскости изображения с плоскостью, параллельной Π и проходящей через отверстие. Заметим, что прямая L на плоскости Π , параллельная плоскости изображения, вообще не дает никакого изображения.

Названные свойства легко доказать чисто геометрически. В то же время, часто бывает удобно (кроме того, элегантно) оперировать такими понятиями, как система отсчета, координаты и уравнения. Рассмотрим, например, систему координат $(O, \mathbf{i}, \mathbf{j}, \mathbf{k})$, связанную с камерой-обскурой, начало координат O которой совпадает с отверстием, а векторы \mathbf{i} и \mathbf{j} образуют базис векторной плоскости, параллельной к плоскости изображения Π' , которая находится на расстоянии f' от отверстия в положительном направлении вектора \mathbf{k} (рис. 1.4). Прямая, проходящая через отверстие и перпендикулярная к Π' , называется оп-

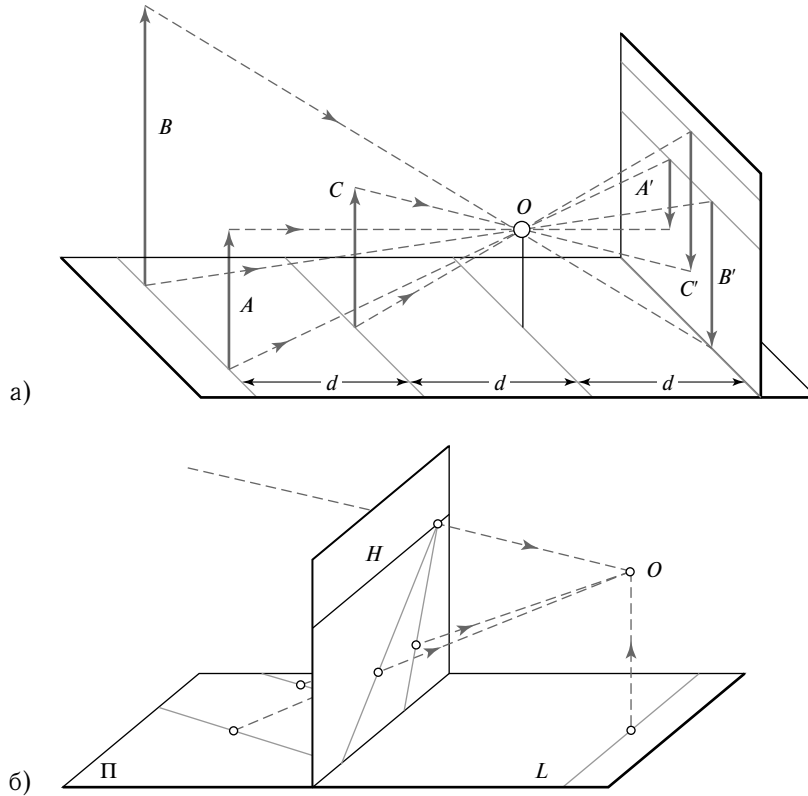


Рис. 1.3. Эффекты перспективы: а) объекты, которые находятся дальше, кажутся меньшими, чем те, которые находятся ближе: расстояние d от отверстия O до плоскости, в которой лежит точка C , в два раза меньше, чем расстояние до плоскости, в которой лежат точки A и B ; б) изображения параллельных прямых пересекаются на горизонтальной линии ([Hilbert and Cohn-Vossen, 1952], fig. 127). Заметим, что в случае *a* плоскость изображения находится за отверстием (физическая сетчатка), а в случае *б* — перед ним (мнимая плоскость изображения). На большинстве схем в этой главе и далее в книге фигурирует именно физическая плоскость изображения, хотя при необходимости будут использованы и мнимые изображения

тической осью, а точка C' , в которой она пересекается с Π' , называется центром изображения. Эту точку можно использовать в качестве начала системы координат, связанной с плоскостью изображения, и она играет важную роль в процедурах калибровки камер.

Обозначим через P точку сцены с координатами (x, y, z) , а через P' — ее изображение с координатами (x', y', z') . Так как P' лежит в плоскости изображения, то $z' = f'$. Поскольку три точки P , O и P' лежат на одной прямой, то $\vec{OP'} = \lambda \vec{OP}$ для некоторого числа λ . Отсюда следует, что

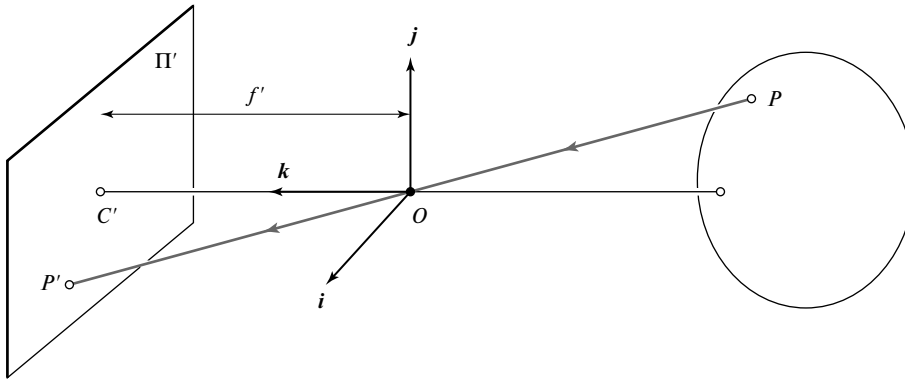


Рис. 1.4. Для вывода уравнений перспективной проекции в этом разделе использован тот факт, что точка P , ее изображение P' и отверстие O лежат на одной прямой

$$\begin{cases} x' = \lambda x \\ y' = \lambda y \\ f' = \lambda z \end{cases} \iff \lambda = \frac{x'}{x} = \frac{y'}{y} = \frac{f'}{z},$$

следовательно,

$$\begin{cases} x' = f' \frac{x}{z}, \\ y' = f' \frac{y}{z}. \end{cases} \quad (1.1)$$

1.1.2. Аффинная проекция

Как было сказано в предыдущем разделе, точечная перспективная проекция только приблизительно описывает геометрию процесса построения изображения. В этом разделе представлен ряд грубых приближений, называемых *аффинными проекционными моделями*, которые иногда также могут оказаться полезными. Рассмотрим две аффинные модели: *слабоперспективную* и *ортогональную*. Третья модель из семейства аффинных, *парперспективная*, представлена в главе 12.

Фронтально-параллельная плоскость Π_0 определяется уравнением $z = z_0$ (рис. 1.5). Для любой точки P на Π_0 уравнение перспективной проекции (1.1) можно переписать как

$$\begin{cases} x' = -mx \\ y' = -my \end{cases} \quad m = -\frac{f'}{z_0}. \quad (1.2)$$

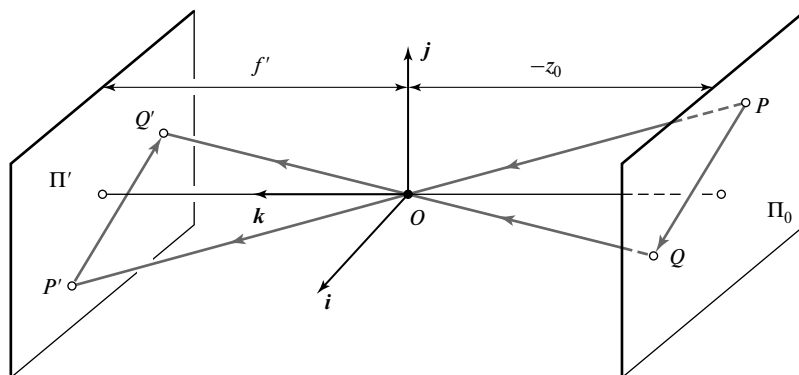


Рис. 1.5. Слабоперспективная проекция: все линейные элементы на плоскости Π_0 проектируются с одинаковым увеличением

По физическим соображениям z_0 должно быть отрицательным (плоскость должна находиться перед отверстием), так что коэффициент увеличения m , связанный с плоскостью Π_0 , больше нуля. Название параметра объясняется следующими соображениями: рассмотрим две точки P и Q на плоскости Π_0 и их изображения P' и Q' (рис. 1.5); очевидно, что векторы \overrightarrow{PQ} и $\overrightarrow{P'Q'}$ параллельны, поэтому $|\overrightarrow{P'Q'}| = m|\overrightarrow{PQ}|$. Это выражение и представляет зависимость размера изображения от расстояния до объекта, о которой упоминалось ранее.

Когда глубина сцены мала по сравнению со средним расстоянием до камеры, увеличение можно считать постоянным. Эта проекционная модель называется слабой проекцией, или *масштабной ортографией*. Если заранее известно, что камера постоянно остается на приблизительно одинаковом расстоянии от сцены, можно нормировать координаты изображения таким образом, чтобы $m = -1$. Это будет уже ортогональная проекция, которая определяется как

$$\begin{cases} x' = x, \\ y' = y. \end{cases} \quad (1.3)$$

Причем все лучи света параллельны оси k и перпендикулярны плоскости изображения Π' (рис. 1.6). Хотя слабоперспективная модель приемлема при описании многих процессов формирования изображения, применение чистой ортогональной проекции обычно дает нереалистичные результаты.

1.2. КАМЕРЫ С ЛИНЗАМИ

Большинство камер оснащены линзами. Можно назвать две основные причины такого технического решения. Первая — аккумуляция света, так как одиночный световой луч сможет достичь каждой точки на плоскости изображения

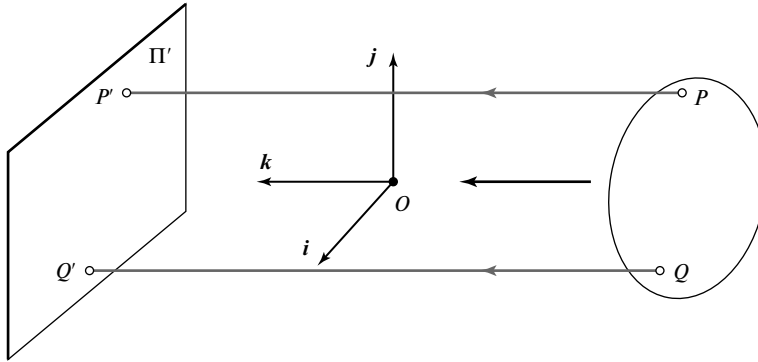


Рис. 1.6. Ортогональная проекция. В отличие от других геометрических моделей процесса формирования изображения, при ортогональном проектировании элементы изображения не переворачиваются. Соответственно, увеличение берется отрицательным, что немного неестественно, но зато упрощает проекционные уравнения

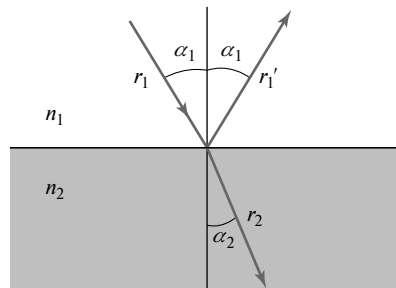


Рис. 1.7. Отражение и преломление света на границе двух однородных сред с показателями преломления n_1 и n_2

только при идеальной точечной проекции. У реальных отверстий, конечно, есть определенный размер, так что каждая точка на плоскости изображения освещается конусом световых лучей, образующих определенный телесный угол. Чем больше отверстие, тем шире конус и ярче изображение, но большие отверстия дают расплывчатое изображение. Сокращение размера отверстия повышает четкость изображения, но уменьшает количество света, который попадает на плоскость изображения, следствием чего могут быть *эффекты дифракции*. Вторая основная причина использования линз — это удержание изображения строго в фокусе при сборе света с большой площади.

Если не принимать во внимание дифракцию, интерференцию и другие физические явления, то поведение линз подчиняется законам геометрической оптики (рис. 1.7): 1) в однородной среде свет распространяется по прямой линии (*лучи*); 2) при отражении луча от поверхности падающий луч, отраженный луч и нормаль к отражающей поверхности лежат в одной плоскости, а сумма углов

между нормалью и двумя лучами равна 90° ; 3) при переходе луча из одной среды в другую происходит его преломление (т.е. луч изменяет свое направление). Согласно закону Снелла (Snell's law), если r_1 — луч, падающий на границу раздела двух прозрачных сред с показателями преломления n_1 и n_2 , а r_2 — отраженный луч, то r_1 , r_2 и нормаль к поверхности отражения лежат в одной плоскости, а углы α_1 и α_2 между нормалью и двумя лучами связаны соотношением

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2. \quad (1.4)$$

В этой главе эффекты отражения не рассматриваются, мы рассмотрим только преломление. Другими словами, мы сконцентрируем наше внимание на линзах, которые отличаются от используемых в *оптических системах с зеркальными линзами*, или *катадиоптрических системах* (например, телескопах-рефлекторах), в которые входят как отражающие (зеркала), так и преломляющие элементы. Проследить за ходом лучей света через линзу проще, если углы между этими лучами и преломляющими поверхностями линз считаются небольшими. В следующем разделе описан именно такой случай.

1.2.1. Параксиальная геометрическая оптика

В этом разделе рассмотрена параксиальная геометрическая оптика (или оптика первого порядка), где углы между всеми лучами, проходящими через линзу, и нормалью к преломляющей поверхности малы. Кроме того, мы предполагаем, что линзы обладают аксиальной симметрией относительно прямой, называемой оптической осью, и что все преломляющие поверхности сферические. Такая симметрия позволяет определить геометрию проекции, если рассматривать линзы с круглыми краями, лежащие в плоскости, которая содержит оптическую ось.

Пусть луч света проходит через точку P_1 на оптической оси и преломляется в точке P сферической поверхности радиуса R , разделяющей две прозрачные среды с показателями преломления n_1 и n_2 (рис. 1.8). Обозначим через P_2 точку, в которой преломленный луч пересекает оптическую ось во второй раз (точки P_1 и P_2 полностью симметричны), а через C — центр сферической поверхности.

Пусть через α_1 и α_2 обозначены, соответственно, углы между двумя лучами и прямой, соединяющей точки P и C . Если β_1 (соответственно, β_2) — угол между оптической осью и прямой, соединяющей точку P_1 (соответственно, точку P_2) с точкой P , то угол между оптической осью и прямой, соединяющей точки P и C (см. рис. 1.8), равен $\gamma = \alpha_1 - \beta_1 = \alpha_2 + \beta_2$. Обозначим теперь через h расстояние от точки P до оптической оси, а через R — радиус сферической поверхности. Если считать, что все углы малы и в первом приближении равны своим синусам и тангенсам, получим

$$\alpha_1 = \gamma + \beta_1 \approx h \left(\frac{1}{R} + \frac{1}{d_1} \right) \quad \text{и} \quad \alpha_2 = \gamma - \beta_2 \approx h \left(\frac{1}{R} - \frac{1}{d_2} \right).$$

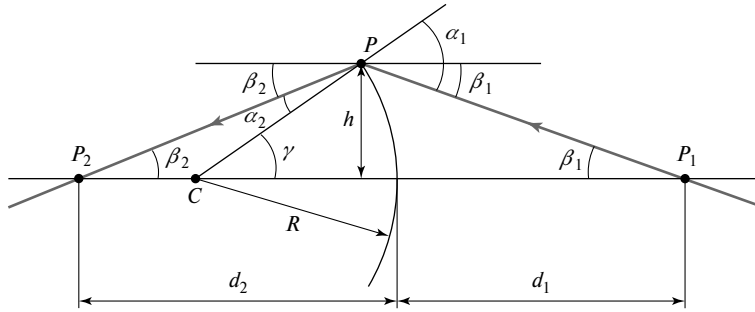


Рис. 1.8. Параксиальное преломление: луч света, проходящий через точку P_1 , преломляется в точке его пересечения со сферической поверхностью, P . Преломленный луч пересекается с оптической осью в точке P_2 . Центр сферической поверхности находится в точке C оптической оси, а ее радиус — R . Предполагается, что углы α_1 , β_1 , α_2 и β_2 малы

Если записать закон Снелла для малых углов, получим уравнение параксиального преломления:

$$n_1 \alpha_1 \approx n_2 \alpha_2 \iff \frac{n_1}{d_1} + \frac{n_2}{d_2} = \frac{n_2 - n_1}{R}. \quad (1.5)$$

Заметим, что отношение d_1 к d_2 зависит от R , n_1 и n_2 , но не от b_1 или b_2 . Это главное упрощение, которое дает параксиальное приближение. Уравнение (1.5) справедливо и тогда, когда некоторые (или все) величины d_1 , d_2 и R становятся отрицательными, что соответствует расположению точек P_1 , P_2 и C по другую сторону оптического центра.

Разумеется, настоящие линзы имеют, по меньшей мере, две преломляющих поверхности. Соответствующие пути прохождения лучей можно построить, несколько раз применяя уравнение параксиального преломления. В следующем разделе это выполнено для тонкой линзы.

1.2.2. Тонкие линзы

Рассмотрим линзу с двумя сферическими поверхностями радиуса R с показателем преломления n . Допустим, что линза находится в вакууме (или, для более точного приближения, в воздухе), показатель преломления которого равен 1, и что эта линза *тонкая* (т.е. луч, попадая на линзу и преломляясь на ее правой грани, немедленно снова преломляется на ее левой грани).

Рассмотрим точку p , расположенную на расстоянии z (в отрицательном направлении оси отсчета) от оптической оси, и обозначим через (PO) луч, проходящий через эту точку и центр O линзы (рис. 1.9). Как показано в разделе упражнений, из закона Снелла и уравнения (1.5) следует, что луч (PO) не преломляется и что все остальные лучи, проходящие через точку P , тонкая линза

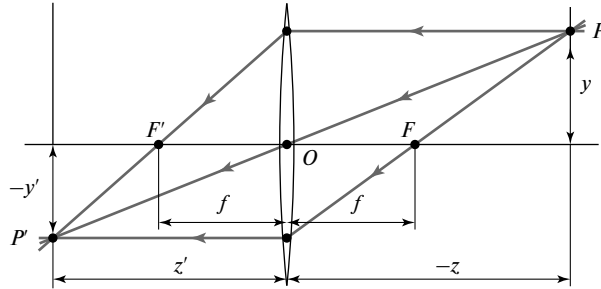


Рис. 1.9. Тонкая линза. Лучи, проходящие через точку O , не преломляются. Лучи, параллельные оптической оси, собираются в фокусе F'

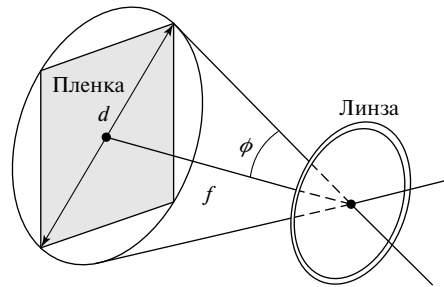


Рис. 1.10. Поле зрения камеры равно 2ϕ , где $\phi \stackrel{\text{def}}{=} \arctg \frac{d}{2f}$, d — диаметр датчика (пленки или ПЗС-датчика), а f — фокусное расстояние камеры. Если f намного меньше d , имеем широкоугольный объектив, в котором лучи отклоняются от оптической оси больше, чем на 45° . У телеобъектива маленькое поле зрения, поэтому изображения, которые он создает, ближе к аффинным

фокусирует в точке P' на расстоянии z' по направлению луча (PO) , так что

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}, \quad (1.6)$$

где $f = \frac{R}{2(n-1)}$ — фокусное расстояние линзы.

Заметим, что уравнения, связывающие точки P и P' , точно такие же, как и при точечной проекции, если принять, что $z' = f'$, поскольку P и P' лежат на луче, проходящем через центр линзы, а точки, которые находятся на расстоянии $-z$ от точки O , попадают точно в фокус только тогда, когда плоскость изображения находится на расстоянии z' от точки O с другой стороны линзы, что удовлетворяет уравнению (1.6) (т.е. уравнению тонкой линзы). Устремив $z \rightarrow -\infty$, можно показать, что f — это расстояние от центра линзы до плоскости фокусировки изображений таких объектов, как звезды, расстояние до которых можно считать равным $z = -\infty$. Две точки, F и F' , находящиеся на расстоянии f от центра линзы на оптической оси, называются *фокусами* линзы.

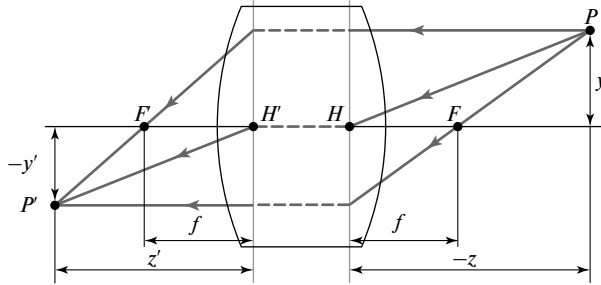


Рис. 1.11. Простая толстая линза с двумя сферическими поверхностями

На практике объекты, входящие в определенный диапазон расстояний (который называется *глубиной поля*, или *глубиной фокуса*), попадают в область, допустимую для фокуса. Как показано в разделе упражнений, глубина поля возрастает с увеличением *диафрагментарного числа* линзы (т.е. отношения фокального расстояния линзы к ее диаметру). *Поле зрения камеры* — это часть площади кадра, которая действительно проектируется на чувствительную область камеры. Оно определяется не только фокальным расстоянием, но еще зависит от эффективной площади чувствительной области (например, площадь пленки, которая освещается в фотоаппарате или площадь ПЗС-датчика в цифровой камере — рис. 1.10).

1.2.3. Реальные линзы

Более реалистичная модель простой оптической системы — это толстая линза. Уравнения, описывающие ее поведение, легко выводятся из уравнения параксиального преломления; они выглядят так же, как и уравнения для проекции и тонких линз, за исключением смещения (рис 1.11): если через H и H' обозначить главные точки линзы, тогда уравнение (1.6) сохраняется при условии, что $-z$ (соответственно, z') — расстояние от точки P (соответственно, P') до плоскости, перпендикулярной к оптической оси и проходящей через H (соответственно, H'). В этом случае единственный непреломляющийся луч проходит вдоль оптической оси.

Простые линзы имеют некоторые *абберации*. Чтобы понять, почему это происходит, вначале напомним, что уравнение параксиального преломления (1.5) — это только приближение, справедливое при условии, что угол α между любым проходящим лучом и оптической осью системы мал, и $\sin \alpha \approx \alpha$. Для больших углов разложение синуса в ряд Тейлора до членов третьего порядка дает такое уточнение параксиального уравнения:

$$\frac{n_1}{d_1} + \frac{n_2}{d_2} = \frac{n_2 - n_1}{R} + h^2 \left[\frac{n_1}{2d_1} \left(\frac{1}{R} + \frac{1}{d_1} \right)^2 + \frac{n_2}{2d_2} \left(\frac{1}{R} - \frac{1}{d_2} \right)^2 \right].$$

Здесь через h обозначено (как и на рис. 1.8) расстояние от оптической оси до точки, в которой падающий луч пересекается с границей раздела. В частности, лучи, которые попадают на границу раздела дальше от оптической оси, фокусируются ближе к границе раздела.

Тот же эффект имеет место и для линз, он является причиной двух типов *сферических aberrаций* (рис. 1.12, *a*). Рассмотрим точку P на оптической оси и ее парааксиальное изображение P' . Расстояние от точки P' до точки пересечения оптической оси с лучом, выходящим из точки P и преломляемым линзой, называется *продольной сферической aberrацией* этого луча. Заметим, что если плоскость изображения Π' поставить в точку P , то луч пересекал бы эту плоскость на некотором расстоянии от оси, что называется *поперечной сферической aberrацией* луча. Все лучи, проходящие через точку P и преломляемые линзой, при пересечении с плоскостью Π' образуют пятно рассеяния с центром в точке P . Размер этого пятна будет меняться, если двигать Π' вдоль оптической оси. Пятно минимального размера называется *пятном наименьшего рассеяния*, и (в общем случае) оно не находится в точке Π' .

Кроме сферической aberrации, существует еще четыре других типа *первичных aberrаций*, вызванных отличием оптических явлений первого и третьего порядка, а именно: *кома*, *астигматизм*, *кривизна поля изображения* и *дисторсия*. Точное определение этих aberrаций в данной книге мы приводить не будем. Достаточно сказать, что подобно сферическим aberrациям, они ухудшают изображение, делая расплывчатой каждую точку объекта. Искажение дает несколько иной эффект, изменяя форму изображения в целом (рис. 1.12, *b*). Это вызвано тем, что в различных частях линзы фокусное расстояние немного меняется. Все упоминавшиеся ранее aberrации монохроматичны (т.е. они не зависят от преломления линзой волн различной длины). В то же время, показатель преломления прозрачной среды зависит от длины волны (рис. 1.12, *в*); из уравнения тонкой линзы (1.6) следует, что фокусное расстояние также зависит от длины волны. Таким образом, возникает явление *хроматической aberrации*: преломленные лучи разной длины волны пересекают оптическую ось в разных точках (*продольная хроматическая aberrация*) и образуют разные пятна рассеяния на одной и той же плоскости изображения (*поперечная хроматическая aberrация*).

Аберрации можно свести к минимуму, сочетая несколько простых линз с хорошо подобранными формой и показателями преломления, разделенных соответствующими промежутками. Такие составные линзы можно точно так же смоделировать с помощью уравнений толстой линзы. Впрочем, у них есть еще один недостаток, важный в контексте компьютерного зрения: пучок света, исходящий от точек объекта, которые находятся не на оси, частично перекрывается различными диафрагмами (включая отдельные компоненты линзы), находящимися внутри линзы для ограничения aberrаций (рис. 1.13). Это явление, которое называют *виньетированием*, приводит к снижению яркости на периферии изображения. Виньетирование может создавать проблемы для программ автоматического анализа изображения, но оно не настолько важно в фотографии

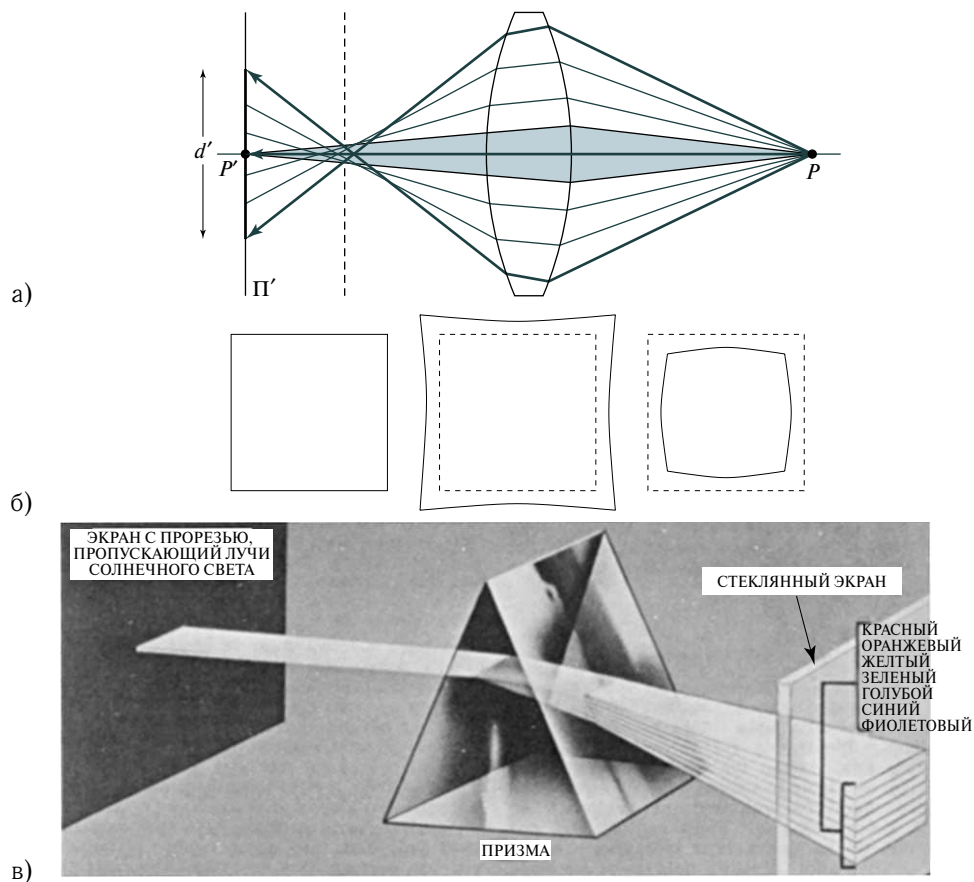


Рис. 1.12. Аберрации; а) сферическая аберрация: область серого цвета — параксиальная зона, где лучи, выходящие из точки P , пересекаются в ее параксиальном изображении P' . Если поставить плоскость изображения Π' в точку P , то изображение P' точки в этой плоскости будет выглядеть как пятно рассеяния диаметром d' . Фокальная плоскость, которая дает пятно наименьшего рассеяния, обозначена пунктирной линией; б) искажение. Слева направо: номинальное изображение прямо-параллельного квадрата, подушкообразное искажение и бочкообразное искажение; в) хроматическая аберрация: показатель преломления прозрачной среды зависит от длины волны (или цвета) падающих лучей. Здесь белый свет, проходя через призму, распадается на спектр цветов. Рисунок из руководства *US NAVY MANUAL OF BASIC OPTICS AND OPTICAL INSTRUMENTS*, подготовленного Комитетом по управлению военно-морским персоналом США, перепечатанный корпорацией *Dover Publications*, (1969)

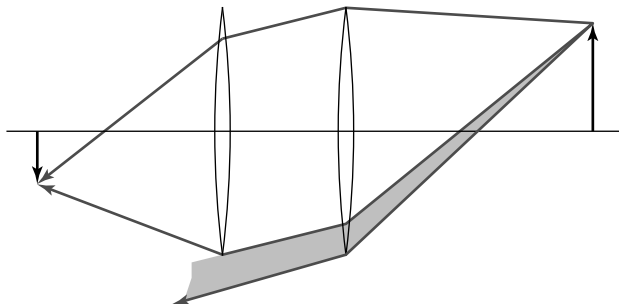


Рис. 1.13. Эффект виньетирования в системе, состоящей из двух линз. Затененная часть пучка никогда не попадет на вторую линзу. Дополнительные диафрагмы и промежутки еще больше усиливают виньетирование

благодаря замечательной способности человеческого глаза сглаживать градиент яркости. Пожалуй, если уж мы заговорили о человеческом глазе, стоит рассмотреть его подробнее.

1.3. ЧЕЛОВЕЧЕСКИЙ ГЛАЗ

Ниже предлагается краткий обзор анатомического строения человеческого глаза (по большей части он основан на представлении Ванделла [Wandell, 1995]). На рис. 1.14 (слева) изображено глазное яблоко в разрезе по его вертикальной плоскости симметрии и показаны главные части глаза: *радужная оболочка* и *зрачок*, контролирующие количество света, который проникает в глазное яблоко; *роговица* и *хрусталик*, на которых преломляется свет для создания изображения на сетчатке; и, наконец, собственно *сетчатка*, где образуется изображение.

Несмотря на свою шарообразную форму, человеческое глазное яблоко по функциональным возможностям напоминает камеру с полем зрения 160° по ширине и 135° по высоте. Как и любая другая оптическая система, глаз подвержен различным видам геометрических и хроматических аберраций. Было предложено несколько моделей глаза, которые описываются законами геометрической оптики первого порядка. На рис. 1.14 (справа) показана, например, схематическая *модель глаза по Гельмольцу*. Здесь имеется только три преломляющих поверхности с бесконечно тонкой роговицей и однородным хрусталиком. На рис. 1.14 изображен глаз, сфокусированный на бесконечности (*несфокусированный взгляд*). Разумеется, данная модель позволяет лишь приближенно описать реальные оптические характеристики глаза

Обратимся еще раз к составляющим частям глаза: роговица — это прозрачное, сильно изогнутое, преломляющее свет окно, через которое свет попадает в глаз до того, как он частично перекрывается цветной и непрозрачной поверхностью радужной оболочки. Зрачок — это отверстие в центре радужной обо-

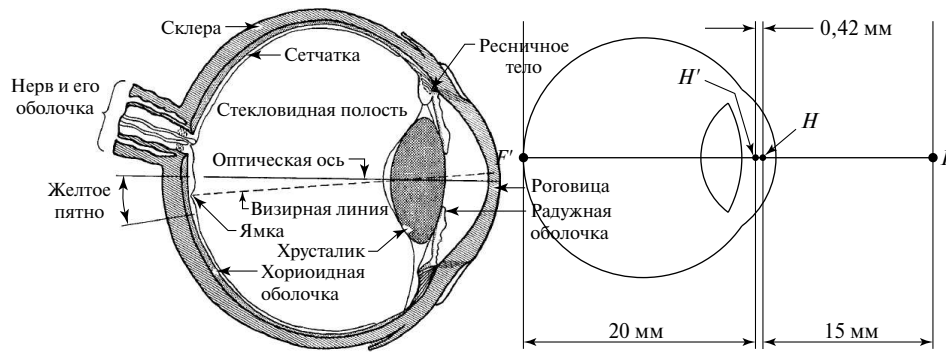


Рис. 1.14. Слева: основные части человеческого глаза. Перепечатано с разрешения American Society for Photogrammetry and Remote Sensing из: A.L. Nowicki, "Stereoscopy". MANUAL OF PHOTOGRAMMETRY, edited by M.M. Thompson, R.C. Eller, W.A. Radlinski, and J.L. Speert, third edition, pp. 515–536. Bethesda: American Society of Photogrammetry, (1966). Справа: схема глаза по Гельмольцу, модифицированная Лаурансом (согласно [Driscoll and Vaughan, 1978]). Расстояние от полюса роговицы до передней главной плоскости — 1,96 мм, а радиусы роговицы, передней и задней поверхностей линз — 8 мм, 10 мм и 6 мм, соответственно

лочки, диаметр которого меняется от 1 до 8 мм в зависимости от освещения. При слабом освещении зрачок расширяется, чтобы увеличить энергию, которая попадает на сетчатку, при нормальном освещении зрачок сужается, чтобы противостоять сферическим aberrациям в глазном яблоке и получить более четкое изображение. Преломляющая сила (аналог фокусного расстояния) глаза по большей части зависит от преломления на границе раздела воздух/роговица и хорошо регулируется изменением формы хрусталика, который аккомодируется для того, чтобы удерживать объект строго в фокусе. У здорового взрослого человека она меняется от 60 (неаккомодированный глаз) до 68 диоптрий (1 диоптрия = 1 м^{-1}), что соответствует изменению фокусного расстояния в пределах 15–17 мм. Сама сетчатка — это тонкая, однослойная мембрана, состоящая из двух видов фоторецепторов — палочек и колбочек, — которые реагируют на свет с длиной волны 330–730 нм (от фиолетового до красного). Существует три вида колбочек с различной спектральной чувствительностью, и это играет ключевую роль при восприятии цвета. В человеческом глазе приблизительно 100 млн палочек и 5 млн колбочек. Их распределение по поверхности сетчатки меняется: желтое пятно — это область в центре сетчатки, где концентрация колбочек особенно высока, и изображение попадает точно в фокус, если глаз фокусирует внимание на объекте (рис. 1.14). Самая высокая концентрация колбочек в ямке — углублении в центре желтого пятна, где она достигает значения $1,6 \times 10^5/\text{мм}^2$, когда расстояние между центрами двух соседних колбочек составляет всего полминуты угла зрения (рис. 1.15). И наоборот, в центре ямки нет палочек, зато их плотность растет по направлению к периферии поля зрения. Еще на сетчатке есть слепое пятно, в котором узел аксона соединяется с сетчаткой и образует оптический нерв.

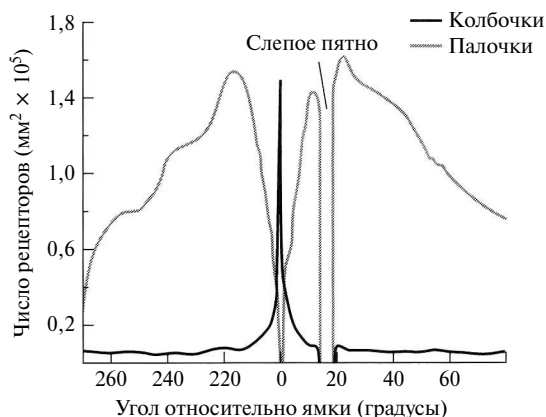


Рис. 1.15. Распределение палочек и колбочек по сетчатке. *Перепечатано из В. Wandell, FOUNDATIONS OF VISION, Sinauer Associates, Inc., 1995*

Палочки представляют собой чрезвычайно чувствительные фоторецепторы; они способны отреагировать на один фотон, но, несмотря на их большое число, палочки дают относительно мало информации о пространственном расположении объекта, поскольку многие из них связаны с одним и тем же нейроном на сетчатке. Колбочки, наоборот, становятся активными при более высоком уровне освещенности, но сигнал, который подает каждая колбочка в ямке, расшифровывается несколькими нейронами, что сказывается на высоком разрешении в этой области. В общем случае, область сетчатки, которая влияет на реакцию нейронов, обычно называют *чувствительным полем сетчатки*, хотя теперь этим словом обозначают также реальную электрическую реакцию нейронов на оптические изображения.

Конечно, о человеческом глазе можно (и нужно) сказать намного больше — например, о том, как оба наши глаза согласуют свою работу и фиксируются на цели, взаимодействуют друг с другом при стереозрении и т.п. Кроме того, в этой телекамере человеческого мозга зрение только начинается, что ставит захватывающую (и по большей части еще не решенную) задачу выяснения роли различных участков нашего мозга в зрении человека. Некоторые аспекты этой проблемы обсуждаются в следующих главах.

1.4. ВОСПРИЯТИЕ

Отличие камер (в сегодняшнем смысле этого слова) от переносной камеры-обскуры XVII века заключается в способности записывать изображения, которые формируются на ее задней стенке. Хотя еще в Средние века было известно, что некоторые соли серебра быстро темнеют под действием солнечного света, только в 1816 году Нипс (Niépce) получил первые настоящие фотографии, подержав бумагу, обработанную хлоридом серебра, действием лучей света, по-

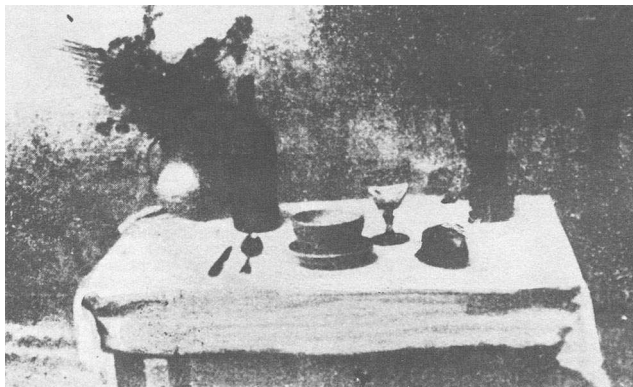


Рис. 1.16. Наиболее ранняя из сохранившихся фотографий, *la table servie*, полученная Ничипором Нипсом (Nicéphore Niépce) в 1822 году. Коллекция Харлингед-Буолемм (*Harlinge-Viollet*)

падающих на плоскость изображения камеры-обскуры, а затем зафиксировав изображения с помощью азотной кислоты. Эти первые изображения были негативами, и Нипс вскоре перешел на другие чувствительные к свету химикаты, чтобы получить позитивное изображение. Самые первые фотографии были утеряны, а самая первая из сохранившихся — это *la table servie* (накрытый стол), воспроизведенная на рис. 1.16.

Нипс изобрел фотографию, а Дагерре (Daguerre) можно считать ее первым популяризатором. С 1826 года эти ученые стали работать вместе. Дагерре продолжил разрабатывать свой собственный фотографический процесс, пользуясь парами ртути для закрепления и проявления скрытого изображения, образовавшегося на меди, покрытой йодидом серебра. Такие *дагерротипы* пользовались большим успехом, когда Араго (Arago) представил процесс Дагерре во Французской Академии Наук в 1839 году, через три года после смерти Нипса. Другими вехами в длинной истории фотографии были: введение мокроколлоидного процесса создания негатива/позитива Легре (Légray) и Арчером (Archer) в 1850 году (фотографию требовалось создавать сразу же, но при этом оставались отличные негативы); изобретение желатинового процесса Маддоксом (Maddox) в 1870 году (после этого процесс создания фотографии уже не требовалось выполнять немедленно после съемки); введение в 1889 году фотопленки (которая заменила стеклянные пластинки в большинстве современных аппаратов) Истманом (Eastman); изобретение кино братьями Люмьер (Lumière) в 1895 году и цветной фотографии в 1908 году.

Изобретение телевидения в 1920-х Беардом (Baird), Фансвортом (Farnsworth) и Зворыкиным (Zworykin) было, конечно, главным толчком к развитию электронных сенсоров. *Видикон* — это общий тип телевизионной электронно-лучевой трубки. Это стеклянная колба с электронной пушкой на одном конце и плоским экраном на другом. Задняя сторона экрана покрыта тонким слоем

фоторезистора, на который наложена прозрачная пленка из положительно заряженного металла. Это двойное покрытие представляет собой *мишень*. Трубку окружают фокусирующие и отклоняющие катушки, которые используются для многократного прохождения электронного пучка, выходящего из пушки, по мишени. Этот пучок оставляет слой электронов на мишени, компенсируя, таким образом, ее положительный заряд. Когда свет попадает на маленькую область экрана, он выбивает электроны из этого места, тем самым заряд мишени уменьшается. Проходя по этой области, электронный пучок восполняет утраченные электроны, и возникает ток, пропорциональный интенсивности падающего света. Затем схема видеодикона преобразует эти изменения силы тока в видеосигнал.

1.4.1. ПЗС-камеры

Рассмотрим теперь, чем отличаются *камеры на приборах с зарядовой связью (ПЗС-камеры)*, которые были предложены в 1970 году и заменили видеодиконы в большинстве современных устройств, от портативных видеокамер, рассчитанных на широкого потребителя, до специализированных камер, предназначенных для применения в микроскопии или астрономии. В ПЗС-датчике используется прямоугольная решетка из узлов, где собираются электроны, покрытая тонкой кремниевой пластинкой, для регистрации количества световой энергии, попадающей на каждый из них (рис. 1.17). Каждый узел образуется путем наращивания слоя диоксида кремния на пластинке и последующим осаждением на диоксид проводящей затворной структуры. Когда фотон падает на кремниевую пластинку, рождается пара электрон-дырка (имеет место *фотопереход*), и электроны попадают в *потенциальную яму*, полученную вследствие приложения положительного электрического потенциала к соответствующему затвору. Электроны, появляющиеся на каждом узле, собираются за фиксированный период времени T .

На данном этапе заряды, хранимые на отдельных узлах, перемещаются посредством *зарядовой связи*: для перемещения зарядовых пакетов от одного узла изменяются потенциалы затворов, при этом пакеты распространяются по отдельности. Изображение считывается с прибора с зарядовой связью построчно, каждая строка передается параллельно к регистру последовательного вывода, содержащему по одному элементу в каждом столбце. Между считыванием двух рядов регистр последовательно передает свои заряды на выходной усилитель, который генерирует сигнал, пропорциональный полученному заряду. Этот процесс продолжается, пока не будет считано все изображение. Он может повторяться до 30 раз за секунду (телевидение) или с намного меньшей скоростью, предоставляя большой промежуток времени (секунды, минуты и даже часы) для сбора электронов в устройствах, работающих при низком уровне освещенности (например, астрономия). Следует отметить, что цифровая выходящая информация в большинстве ПЗС-камер преобразуется в аналоговый видеосигнал, а лишь потом попадает в *механизм захвата кадра*, который создает окончательное цифровое изображение.

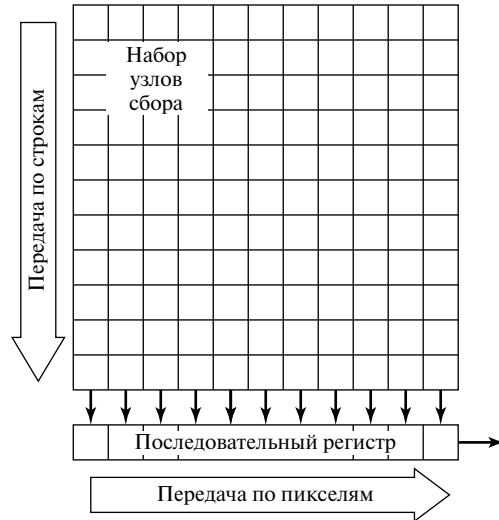


Рис. 1.17. Устройство ПЗС

В цветных ПЗС-камерах потребительского уровня, по сути, используются те же микросхемы, что и в черно-белых, за исключением того, что задействованные строки или столбцы датчиков чувствительны к красному, зеленому или синему свету (это часто достигается с помощью фильтра, который пропускает только нужный цвет). Возможны и другие модели фильтров, например, мозаика из блоков 2×2 , состоящих из двух зеленых, одного красного и одного синего рецепторов (*модель Байера*). Пространственное разрешение камер на одном ПЗС, конечно, ограничено, и в высококачественных камерах применяется расщепитель пучка, чтобы распределять изображение по трем различным ПЗС через цветные фильтры. Затем отдельные цветные каналы оцифровываются также отдельно (RGB-сигнал) или объединяются в сложный цветной видеосигнал (сигнал *NTSC* в США, *SECAM* или *PAL* в Европе и Японии) или в составной видеоформат, где информация о цвете и яркости разделяется.

1.4.2. Модели датчиков

В этом разделе описаны черно-белые ПЗС-камеры, поскольку цветные камеры можно считать точно такими же, если рассмотреть каждый цветовой канал в отдельности и непосредственно учесть связанную с этим реакцию фильтров.

Количество электронов I , зарегистрированных в клетке, расположенной в строке r и в столбце c схемы ПЗС, можно представить как

$$I(r, c) = T \int_{\lambda} \int_{p \in S(r, c)} E(p, \lambda) R(p) q(\lambda) dp d\lambda,$$

где T — время сбора электронов, а интеграл вычисляется по всей области клетки $S(r, c)$ и диапазону длин волн, на которые ПЗС дает ненулевую реакцию. В этом интеграле E — сила света на единицу площади и единицу длины волны (т.е. *освещенность*, формальное определение см. в главе 4) в точке p , R — пространственная реакция узла, а q — *квантовый выход* устройства (т.е. количество рождающихся электронов на единицу энергии падающего света). В общем случае, и E , и q зависят от длины волны λ , а E и R зависят от расположения точки p в области $S(r, c)$.

Усилитель на выходе ПЗС преобразует заряд, собранный на каждом узле, в напряжение, которое можно измерить. В большинстве камер это напряжение затем с помощью электроники камеры преобразуется в низкочастотный¹ видеосигнал, пропорциональный I . Аналоговое изображение можно еще раз преобразовать в цифровое посредством механизма захвата кадра, который пространственно сверяет видеосигнал и квантует значение яркости в каждой точке изображения, или *пикселе* (от *picture element* — элемент изображения).

Существует несколько физических явлений, для учета которых следует корректировать идеальную модель камеры, приведенную выше: если источник света, освещающий узел сбора, настолько ярок, что заряд, накапливаемый на узле, распространяется и на соседние узлы, изображение *размывается*. Этого можно избежать, контролируя освещение, но остальные факторы, такие как дефекты изготовления, тепловые и квантовые эффекты, квантовый шум — неотъемлемая часть процесса создания изображения. Как будет показано дальше, эти факторы учитываются в соответствующих простых статистических моделях. Эффекты квантовой физики вносят внутреннюю неопределенность в процесс фотоперехода на каждом узле (*дробовый шум*). Более точно, количество электронов, возникающих в результате этого процесса, можно обозначить случайным целым числом $N_I(r, c)$, которое подчиняется закону распределения Пуассона, со средним значением $\beta(r, c)I(r, c)$, где $\beta(r, c)$ — число от 0 до 1, которое показывает изменение пространственной реакции и квантовый выход по всему изображению, а также подсчитывает плохие пиксели. Электроны, вылетающие из кремния в результате теплового движения, добавляются к заряду каждого узла сбора. Их вклад называют *теневым током*, и его можно обозначить случайным целым числом $N_{DC}(r, c)$, среднее значение $\mu_{DC}(r, c)$ которого растет с температурой. Эффект теневого тока можно контролировать охлаждением камеры. Дополнительные электроны исходят из электроники ПЗС (*ток смещения*), и их число также можно обозначить случайной, подчиняющейся распределению Пуассона, величиной $N_B(r, c)$ со средним значением $\mu_B(r, c)$. Усилитель на выходе добавляет шум считывания, его можно смоделировать действительным числом R (которое подчиняется распределению Гаусса) со средним значением μ_R и среднеквадратическим отклонением σ_R .

Есть и другие источники неопределенности (например, эффективность переноса заряда), но ими часто можно пренебречь. Наконец, дискретизация анало-

¹Грубо говоря, усредненный по времени или пространству; подробнее об этом — далее.

гового напряжения механизмом захвата кадра порождает как геометрические эффекты (дрожание строк), которые можно устранить с помощью калибровки, так и шум квантования, который можно обозначить случайным числом $Q(r, c)$, среднее значение которого равно нулю, с однородным распределением в интервале $[-\frac{1}{2}\delta, \frac{1}{2}\delta]$ и дисперсией $\frac{1}{12}\delta^2$, где δ — шаг квантования. Это дает такую модель цифрового сигнала $D(r, c)$:

$$D(r, c) = \gamma(N_I(r, c) + N_{DC}(r, c) + N_B(r, c) + R(r, c)) + Q(r, c).$$

В приведенном выше уравнении γ — это общий коэффициент усиления усилителя и камеры. Статистические свойства данной модели можно оценить с помощью радиометрической калибровки камеры: например, теневой ток можно оценить, измеряя число пробных изображений в темноте ($I = 0$).

1.5. ПРИМЕЧАНИЯ

Отличное введение в геометрическую оптику — классический учебник [Hecht, 1987], где подробно рассматривается парааксиальная оптика, а также различные аберрации, о которых вкратце упоминалось в этой главе (см. также [Driscoll and Vaughan, 1978]). Виньетирование проанализировано в [Horn, 1986] и [Russ, 1995]. Работа [Wandell, 1995] дает прекрасное описание процесса образования изображения в зрительной системе человека. Схематическая модель глаза по Гельмольцу подробно описана в [Driscoll and Vaughan, 1978].

ПЗС-устройства представлены в [Boyle and Smith, 1970] и [Amelio, *et al.*, 1970]. Научное применение ПЗС-камер в микроскопии и астрономии обсуждается в [Aiken, *et al.*, 1989], [Janesick, *et al.*, 1987], [Snyder, *et al.*, 1993] и [Tyson, 1990]. Статистическая модель датчика, представленная в этой главе, описана согласно работе [Snyder, *et al.*, 1993] с добавлением понятия шума квантования, взятого из [Healey and Kondepudy, 1994]. В этих двух статьях рассказывается о применении моделирования датчиков для восстановления изображений в астрономии и радиометрической калибровки камер в компьютерном зрении.

Основные уравнения, выведенные при изложении данной главы, представлены в табл. 1.1.

Задачи

- 1.1. Постройте уравнение перспективной проекции для мнимого изображения, расположенного на расстоянии f' перед отверстием.
- 1.2. Докажите геометрически, что проекции двух параллельных прямых, лежащих в плоскости Π , сходятся на горизонтальной прямой H , образованной пересечением плоскости изображения с плоскостью, параллельной к Π и проходящей через отверстие.
- 1.3. Докажите то же самое алгебраически, воспользовавшись уравнением перспективной проекции (1.1). Для простоты можно принять, что плоскость Π перпендикулярна к плоскости изображения.

ТАБЛИЦА 1.1. Справочная таблица: модели камер

Перспективная проекция	$\begin{cases} x' = f' \frac{x}{z} \\ y' = f' \frac{y}{z} \end{cases}$	x, y : координаты реального объекта ($z < 0$) x', y' : координаты изображения f' : расстояние от отверстия до чувствительной поверхности
Слабо-перспективная проекция	$\begin{cases} x' = -mx \\ y' = -my \\ m = -\frac{f'}{z_0} \end{cases}$	x, y : координаты реального объекта x', y' : координаты изображения f' : расстояние от отверстия до чувствительной поверхности z_0 : глубина точки отсчета (< 0) m : коэффициент увеличения (> 0)
Ортографическая проекция	$\begin{cases} x' = x \\ y' = y \end{cases}$	x, y : координаты реального объекта x', y' : координаты изображения
Закон Снелла	$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$	n_1, n_2 : показатели преломления α_1, α_2 : углы между лучом и нормалью
Параксиальное преломление	$\frac{n_1}{d_1} + \frac{n_2}{d_2} = \frac{n_2 - n_1}{R}$	n_1, n_2 : показатели преломления d_1, d_2 : расстояние от точки до границы раздела R : радиус кривизны границы раздела
Уравнение тонкой линзы	$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$	z : глубина точки объекта (< 0) z' : глубина точки изображения (> 0) f : фокусное расстояние

- 1.4. Воспользуйтесь законом Снелла и покажите, что лучи, которые проходят через оптический центр тонкой линзы, не преломляются. Выведите уравнение тонкой линзы. *Подсказка:* рассмотрите луч r_0 , проходящий через точку P , и постройте лучи r_1 и r_2 , которые получаются при преломлении луча r_0 на правой грани линзы и преломлении луча r_1 на ее левой грани.

- 1.5.** Рассмотрите камеру с тонкой линзой, плоскость изображения которой находится на расстоянии z' , а плоскость точек кадра в фокусе — на расстоянии z . Представьте, что плоскость изображения перемещается в точку \hat{z}' . Покажите, что диаметр соответствующего пятна рассеяния равен

$$d \frac{|z' - \hat{z}'|}{z'}$$

где d — диаметр линзы. Воспользуйтесь этим результатом, чтобы показать, что глубина поля (т.е. расстояния между дальней и ближней плоскостями, на которых диаметр пятна рассеяния не превышает некоторого порогового значения ε) задается как

$$D = 2\varepsilon f z (z + f) \frac{d}{f^2 d^2 - \varepsilon^2 z^2},$$

и убедитесь, что для *фиксированного* фокусного расстояния глубина поля возрастает с уменьшением диаметра линзы, т.е. с увеличением дифракционного числа. *Подсказка:* решайте задачу для глубины \hat{z} точки, изображение которой находится в фокусе на плоскости изображения в точке \hat{z}' , рассматривая случаи, когда $\hat{z}' > z'$, и $\hat{z}' < z'$.

- 1.6.** Выполните геометрическое построение изображения P' точки P , если заданы два фокуса F и F' тонкой линзы.
- 1.7.** Выведите уравнение толстой линзы для случая, когда у обеих сферических поверхностей линзы одинаковые радиусы кривизны.