

第3章 彩色成像

本章中描述的成像过程涉及三个过程（照明、材料反射和检测/观察），这些过程相互作用以生成最终的彩色图像。该过程从照亮视觉场景的光开始。光被描述为具有一定强度的电磁辐射，它由包含一定波长能量的粒子（光子）组成，各个光子都沿一定方向传播。当许多光子沿相同方向传播时，光线被引导并形成光束。当所有光子沿随机方向传播时，光就会扩散。光通常由光源发出。光源可以通过光束被引导的方式以及在整个波长范围内发出的光子光谱来表征。当发射的短波长的光子比长波长的光子更多时，光源的彩色为蓝色；当发射更多的长波长的光子时，彩色为红色。对于**烛光和卤素照明**，发射光谱遵循所谓的**黑色辐射体**^[25]的光谱。对于该辐射体，光滑的发射光谱可以通过一个唯一的数字来表征，即辐射体的温度。由于许多自然光源发出的光谱与这种黑色辐射体的彩色相似，因此光源的彩色由“**相关色温**”定义，即黑色辐射体的温度，在该温度下可以感觉到类似的彩色。但是，请记住，有许多非自然光源（例如**荧光灯**）的彩色可能与黑色辐射体很相似，但是其光谱与光滑的黑色辐射体的光谱非常不同。

图像形成的第二个过程涉及材料。场景中的材料与入射光相互作用，引起反射(图 3.1)。材料吸收光子，仅反射照射在材料上的部分光。对于“白色”材料，大多数光子都会被反射。对于“黑色”材料，大多数光子都会被吸收。因此，材料以某种方式调节光的强度。此外，材料吸收的光粒子的数量可能取决于光子的波长。根据材料特性，某些波长的光子可能会被吸收，而其他波长的光子会被反射。在该情况下，这种材料对光进行了光谱调制。例如，当所有中短波长都被吸收而其他波长被反射时，光看起来会发红。3.3 节中介绍的**库贝卡·蒙克 (Kubelka-Munk)** 理论对这种效果进行了详细的建模。

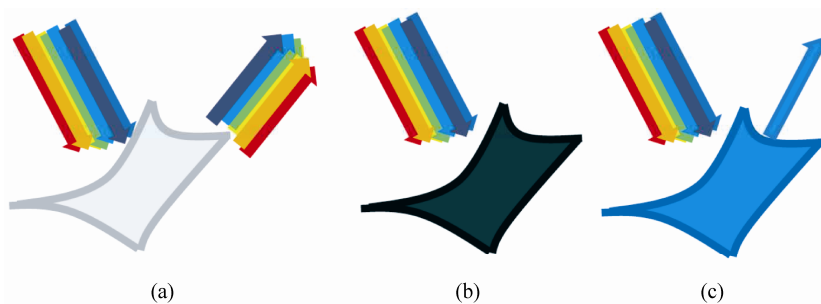


图 3.1 白光与材料相互作用的示例。白光包含可见光谱所有波长的能量。(a)与完美的白色表面互动时，所有光都会反射；(b)对于完美的黑色物体，光会被吸收；(c)对于蓝色材料，仅代表蓝色光的光谱被反射，其他波长被该材料吸收

除了材料内部的吸收以外，还要考虑光击中材料的效果。当光线照射到材料上时，光线改变介质从空气中穿过材料边界。这样一来，一部分光线会在空气和材料之间的“界面”处反射，这会引起**菲涅耳反射**或**界面反射**，本书中使用后者。界面反射导致材料上的镜面反射（也称为**高光**）。在 3.2 节的双色反射模型中考虑了这一因素。

在图像形成的第三个过程中，通过相机或眼睛记录或观察光。此时通过在一定带宽、一定空间区域和一定时间段内的积分能量来记录光子。对于眼睛（2.2.1 节），积分是在三个光谱宽带上执行的，覆盖了光谱的短、中和长的波长。积分时间约为 50ms，视敏度取决于光线落在视网膜上的位置。视网膜中央（中央凹）区域的视力最高，并向周边逐步下降。彩色相机或多或少地模仿眼睛的时间和光谱特性，在大约 50ms 内记录三个谱带，并且通常具有百万像素级的均匀空间分辨率。

对光线、材料和观察过程之间相互作用的建模是反射模型的主要目标。这样的模型旨在简化所涉及的物理学，以便于理解过程的某些方面。不同的模型进行了不同的简化，因此更适用于不同条件下的使用，或者比替代模型相比更适用于不同的数学框架。在本章中，我们讨论计算机视觉中使用的最相关的模型。

除了对反射建模外，还需要量化所得到的彩色信息。因为计算机处理数字，所以应将相机每个像素处的光谱信息配准凝练为数字。只要每个数字唯一地定义彩色，任何任意的编号方案都可以使用。例如，编码方案可以按照与彩虹中出现的彩色相同的顺序对彩色进行编号，比如从零开始是深红色，而一百万可以是深紫色。由于历史原因，商用相机按 RGB 方案产生结果。本章的 3.5 节介绍了常见的彩色空间，这些彩色空间将 RGB 信息重新排序为替代方案。所有这些方案本质上都描述了相同的彩色信息。但是，类似于彩色成像模型，某个方案可能比另一个方案更有利于突出显示彩色信息中的某些属性。例如，众所周知的色调-饱和度-值模型将 RGB 值分解为正交配色方案，其中（无色）强度信息独立于色度信息。如本章所示，有许多彩色编码方案，每种彩色编码方案用于计算机视觉时都有其优点和缺点。

3.1 朗伯反射模型

许多计算机视觉应用都基于朗伯反射率的假设，这意味着表面反射的光强度与视角无关。表面亮度可认为是各向同性的，具有此属性的材料称为无光泽/磨砂材料，朗伯反射率的示例是粉笔、纸和未加工的木材。

考虑场景的照度由 $e(\lambda, \mathbf{x})$ 给出，其中 λ 是波长， $\mathbf{x}$ 是图像中的空间位置。通常，我们假设光源的光谱空间分布在整个场景中是均匀的。在这种情况下，可写成 $e(\lambda)$ 。从表面反射的能量（即辐射） E 由下式给出

$$E(\lambda, \mathbf{x}) = m^b(\mathbf{x})s(\lambda, \mathbf{x})e(\lambda, \mathbf{x}) \quad (3.1)$$

其中， s 为描述材料光谱反射特性的表面反射率； m^b 为反射率的几何依赖性， $m^b = \cos \alpha$ ，取决于光源的方向和表面的方向，其中 α 是表面法线和照明方向之间的角度； $\mathbf{x}$ 表示图像的空间坐标，这里使用黑斜体表示矢量。

通过在可见光谱 ω 上积分来对具有光谱灵敏度 $\rho^c(\lambda)$ ， $c \in \{R, G, B\}$ 的相机的观测值 $f_{\text{RGB}} = (R, G, B)$ 进行建模：

$$f^c(\mathbf{x}) = \int_{\omega} E(\lambda, \mathbf{x}) \rho^c(\lambda) d\lambda \quad (3.2)$$

$$= m^b(\mathbf{x}) \int_{\omega} s(\lambda, \mathbf{x}) e(\lambda) \rho^c(\lambda) d\lambda \quad (3.3)$$

这也可以写成矢量

$$\mathbf{f}(\mathbf{x}) = m^b(\mathbf{x}) \mathbf{c}^b(\mathbf{x}) \quad (3.4)$$

其中体反射率

$$\mathbf{c}^b(\mathbf{x}) = \int_{\omega} s(\lambda, \mathbf{x}) e(\lambda) \rho^c(\lambda) d\lambda \quad (3.5)$$

朗伯反射模型预测出，单个彩色目标上的像素位于通过 RGB 立方体原点的线上。请注意，对于许多材料，朗伯假设在严格意义上并不成立。例如，材料可能是有光泽的，从而导致材料上某些位置出现镜面反射。但对于这些材料，朗伯假设可以是一个很好的近似值，因为镜面反射通常只占目标的一小部分。不过，在这些情况下，材料属性也可能有更好的近似值，如下文所述。

3.2 双色反射模型

朗伯模型不包括镜面反射（高光）之类的反射。**双色反射模型（DRM）**包括界面反射或菲涅耳反射，它允许各向异性的镜面反射。

DRM 是由谢弗^[26]提出的，除了基于**朗伯定律**的模型之外，它也是计算机视觉中最受欢迎的反射模型之一。该模型侧重于光反射的彩色方面，并且仅在场景几何恢复中有限地使用。该模型假定场景中只有一个光源。它将反射率分解为表面的体反射率和界面反射率。该模型对于非均质材料类别有效，它涵盖了广泛的材料，例如木材、油漆、纸张和塑料（但不包括金属等均质材料）。DRM 是体反射率（上标 b ）和界面反射率（上标 i ）的总和：

$$\mathbf{f}^c(\mathbf{x}) = m^b(\mathbf{x}) \int_{\omega} s(\lambda, \mathbf{x}) e(\lambda) \rho^c(\lambda) d\lambda + m^i(\mathbf{x}) \int_{\omega} i(\lambda) e(\lambda) \rho^c(\lambda) d\lambda \quad (3.6)$$

请注意，对于 $m^i(\mathbf{x}) = 0$ ，此式等于式(3.3)。我们将考虑中性界面反射（NIR），这意味着**菲涅耳反射率** i 与 λ 无关。因此，我们将在其他方程式中省略 i 。反射率的几何相关性由 m^b 项和 m^i 项来描述，它们取决于视角、光源方向和表面朝向。

在许多情况下，我们可以假定场景中的照明为白色，因此 $e(\lambda) = i$ 是恒定的。例如，可以通过本书第 3 部分中讨论的白平衡或光源估计来获得此效果。消除对 $e(\lambda)$ 的依赖而得到：

$$\mathbf{f}^c(\mathbf{x}) = m^b(\mathbf{x}) \int_{\omega} s(\lambda, \mathbf{x}) \rho^c(\lambda) d\lambda + m^i(\mathbf{x}) \int_{\omega} \rho^c(\lambda) d\lambda \quad (3.7)$$

其中常数因子 i 包含在几何项 m^b 和 m^i 中。当我们进一步假设灵敏度函数 ρ 下的面积近似相同时，称为**积分白色条件**，即， $\int_{\lambda} \rho_R(\lambda) d\lambda = \int_{\lambda} \rho_G(\lambda) d\lambda = \int_{\lambda} \rho_B(\lambda) d\lambda = 1$ ，式(3.7)简化成：

$$\mathbf{f}^c(\mathbf{x}) = m^b(\mathbf{x}) \int_{\omega} s(\lambda, \mathbf{x}) \rho^c(\lambda) d\lambda + m^i(\mathbf{x}) \quad (3.8)$$

解释 DRM 的一种有洞察力的方法是将其表示为矢量符号。然后我们可以将式(3.6)写为

$$\mathbf{f}(\mathbf{x}) = m^b(\mathbf{x}) \mathbf{c}^b(\mathbf{x}) + m^i(\mathbf{x}) \mathbf{c}^i(\mathbf{x}) \quad (3.9)$$

并将式(3.8)写为

$$f(\mathbf{x}) = m^b(\mathbf{x})\mathbf{c}^b(\mathbf{x}) + m^i(\mathbf{x})\mathbf{c}^i \quad (3.10)$$

光的反射包括两个部分：①体反射部分 $m^b(\mathbf{x})\mathbf{c}^b$ ，它描述了与表面反射率相互作用后反射的光；②界面反射部分 $m^i(\mathbf{x})\mathbf{c}^i$ ，它描述了直接在表面反射并引起镜面反射的光。这两个部分都包括一个几何项，该几何项取决于场景中的位置；以及一个取决于光谱波长的光谱项。双色反射模型将单个彩色目标 $f(\mathbf{x})$ 的像素值投影到平行四边形上（图 3.2）。平行四边形上的位置由体反射率和界面反射率确定。

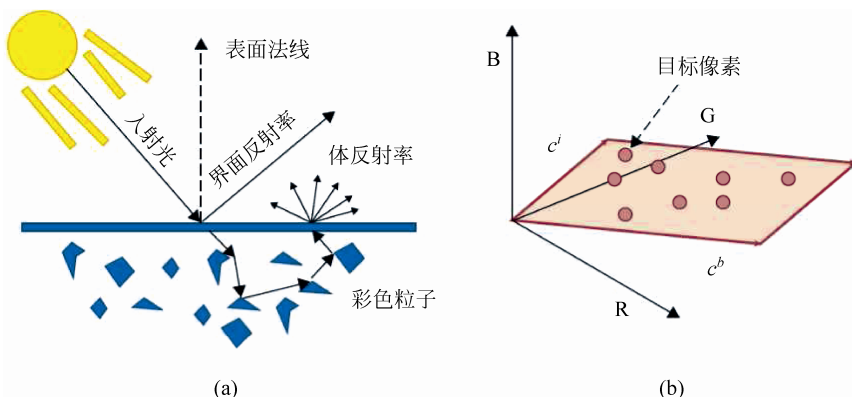


图 3.2 根据双色反射模型的光与材料相互作用。(a)反射光由体反射和界面反射两部分组成；(b)双色反射模型预测单个彩色目标的像素值位于由体反射矢量和界面反射率矢量形成的平行四边形上

如果不满足原始 DRM 的假设，则需要更复杂的反射模型。一种这样的情况是存在着环境光，即，光来自所有方向。环境光存在于室外场景中，在该场景中，主要光源（即太阳）旁边存在来自天空的漫射光。类似地，在室内情况下也会出现这种情况，因为墙壁和天花板的反射会导致散射光。谢弗^[26]通过增加第三项来模拟漫射光 $\mathbf{a}$

$$f(\mathbf{x}) = m^b(\mathbf{x})\mathbf{c}^b(\mathbf{x}) + m^i(\mathbf{x})\mathbf{c}^i(\mathbf{x}) + \mathbf{a} \quad (3.11)$$

后来的工作改进了建模^[27]，并表明环境项会导致目标-彩色-相关的偏移，这对于处理彩色阴影至关重要。

对 DRM 的最初应用是将阴影与镜面反射分离^[26]。取决于场景偶然事件（例如视点和表面法线）的镜面反射可以被去除掉，以简化对彩色图像的理解。去除镜面反射可以改进分割算法^[28-29]。在本书中，我们将看到 DRM 的几种应用，例如彩色恒常性、光度不变特征计算和彩色图像分割。

3.3 库贝卡·蒙克模型

物理学的一个较旧模型是著名的库贝卡·蒙克（Kubelka-Munk）光传输理论。该模型实质上与谢弗（Shafer）的 DRM 模型类似。但是，它通过已建立的实验工作提供了更好的背景，以及对材料反射光所涉及的物理原理的更好理解。因此，我们在这里简要介绍一下该理论。对于更复杂的解释，我们请读者参考 Judd 和 Wyszecki 的出色著作^[30]。

刻画材料光传输特征的是三个基本过程：吸收、散射和发射。吸收是将辐射能转化为热量的过程；散射是辐射能向不同方向扩散的过程；发射是创建新辐射能的过程（本书未考虑）。库贝卡·蒙克理论在假设 1-D 光通量的情况下对这些过程的效果进行了建模，从而揭示了材料内的各向同性散射^{[25][30-32]}。在这种假设下，材料层（即物体表面）的特征取决于与波长有关的散射系数和吸收系数。该理论适用的材料类别包括染色纸和纺织品、不透明塑料、漆膜、液体、搪瓷和牙科硅酸盐水泥。该模型可以应用于反射材料和透明材料。

考虑具有均匀厚度 d 和无穷小面积（图 3.3）的均匀着色材料片，其特征取决于其吸收系数 $K_a(\lambda)$ 和散射系数 $K_s(\lambda)$ 。当材料被具有光谱分布 $e(\lambda)$ 的光照射时，材料内部的光散射会导致漫射体反射，而菲涅耳界面反射发生在表面边界处。当片层的厚度达到一定厚度而进一步增加厚度不影响反射的彩色时，可以忽略在背面的菲涅耳反射。在这种情况下，可以认为该材料具有无限的光学厚度。

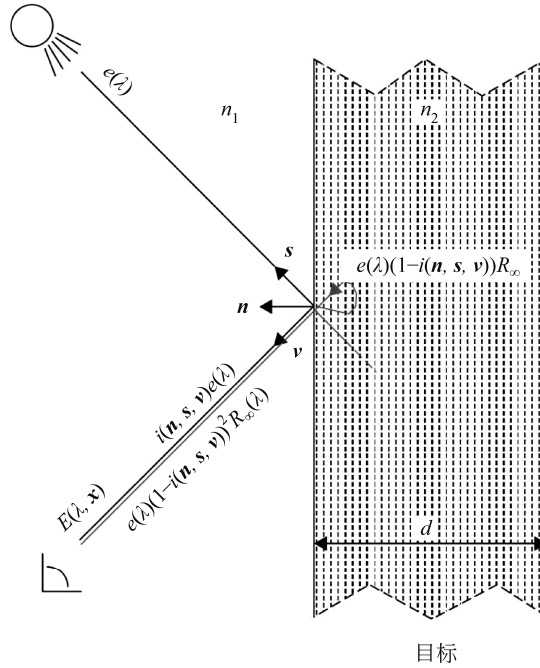


图 3.3 库贝卡·蒙克模型涉及的各个方面

入射光有一部分在正面反射，有一部分进入材料，被各向同性散射，其中一部分再次通过正面边界。忽略边界内反射后的二次散射，在观察方向 $\mathbf{v}$ 上的反射光谱由下式给出：

$$E(\lambda) = e(\lambda)(1 - i(\lambda, \mathbf{n}, \mathbf{s}, \mathbf{v}))^2 R_{\infty}(\lambda) + e(\lambda)i(\lambda, \mathbf{n}, \mathbf{s}, \mathbf{v}) \quad (3.12)$$

其中， $\mathbf{n}$ 是表面片的法线， $\mathbf{s}$ 是光源的方向。此外， i 是在视线方向 $\mathbf{v}$ 上的菲涅耳界面反射系数。体反射 $R_{\infty}(\lambda) = a(\lambda) - b(\lambda)$ 取决于吸收系数 $K_a(\lambda)$ 和散射系数 $K_s(\lambda)$ 。

$$a(\lambda) = 1 + \frac{K_a(\lambda)}{K_s(\lambda)} \quad (3.13)$$

$$b(\lambda) = \sqrt{a^2(\lambda) - 1} \quad (3.14)$$

注意，体反射 $R_{\infty}(\lambda)$ 等于前面讨论的表面反射 $s(\lambda)$ 。

考虑中性界面反射率 (NIR) (3.2 节), 则式(3.12)的反射模型可简化如下:

$$E(\lambda) = e(\lambda)(1 - i(\mathbf{n}, \mathbf{s}, \mathbf{v}))^2 R_\infty(\lambda) + e(\lambda)i(\mathbf{n}, \mathbf{s}, \mathbf{v}) \quad (3.15)$$

使用此模型, 通过代入 $i(\bullet) = 0$ 可以建模出完美的漫射表面, 而通过代入 $i(\bullet) = 1$ 可以建模出完美的光源镜像。实际上, $i(\bullet)$ 将假设介于 0 和 1 之间的某个值, 这将导致光谱色 $E(\lambda)$, 它是光源色和完全漫反射体的反射色的加性混合物, 类似于 DRM 模型。

由于将 3-D 坐标投影到 2-D 图像平面上, 因此矢量 $\mathbf{n}$ 、 $\mathbf{s}$ 和 $\mathbf{v}$ 取决于它们在图像中的位置。图像平面上空间位置 $\mathbf{x}$ 处的入射光谱的能量为

$$E(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x})[1 - i(\mathbf{x})]^2 R_\infty(\lambda, \mathbf{x}) + e(\lambda, \mathbf{x})i(\mathbf{x}) \quad (3.16)$$

其中每个点 $\mathbf{x}$ 的光谱分布是从特定的材料片生成的。如果进行以下替换:

$$c^b(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x})R_\infty(\lambda, \mathbf{x}) \quad (3.17)$$

$$c^i(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x}) \quad (3.18)$$

$$m^b(\mathbf{x}) = [1 - i(\mathbf{x})]^2 \quad (3.19)$$

$$m^i(\mathbf{x}) = i(\mathbf{x}) \quad (3.20)$$

式(3.16)简化为谢弗^[26]提出的双色反射模型:

$$E(\lambda, \mathbf{x}) = m^b(\mathbf{x})c^b(\lambda, \mathbf{x}) + m^i(\mathbf{x})c^i(\lambda, \mathbf{x}) \quad (3.21)$$

因此

$$f^c(\mathbf{x}) = \int_{\omega} E(\lambda, \mathbf{x}) \rho^c(\lambda) d\lambda \quad (3.22)$$

但是, 请注意, 可以从上面的公式得出的系数 m^b 和 m^i 是相互依赖的。通过假设仅有磨砂或无光泽表面, 忽略镜面反射 (即 $i(\mathbf{x}) \approx 0$), 可以进一步把式(3.16)简化为用于漫射体反射的朗伯模型:

$$E(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x})R_\infty(\lambda, \mathbf{x}) \quad (3.23)$$

库贝卡·蒙克理论还推广了光透射的情况。假设材料层的光学厚度有限, 则光将在它进入一侧的相反侧离开该材料。在这种情况下, 材料中的吸收和散射会导致光强度呈指数衰减, 从而产生众所周知的比尔·朗伯方程:

$$E(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x}) \exp\{-d(\mathbf{x})c(\mathbf{x})\alpha(\lambda, \mathbf{x})\} \quad (3.24)$$

其中, d 是层的局部厚度, c 是着色剂颗粒的浓度, α 表示着色剂颗粒的吸收和散射系数。同样, e 是发光体的发射光谱。该方程在诸如透射光显微镜中起着重要作用。

3.4 对角模型

场景中的彩色会随着光源彩色的变化而明显变化。本书的第 3 部分致力于估计场景的光源。在这里, 我们简短地讨论**对角模型**, 该模型可预测在光源变化下相机响应的变化。

对角变换或冯·克里斯模型 (von Kries model)^[33]由下式给出

$$\begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} R_u \\ G_u \\ B_u \end{bmatrix} \quad (3.25)$$

或者简写成

$$\mathbf{f}_c = \mathbf{D}_{u,c} \mathbf{f}_u \quad (3.26)$$

其中, $\mathbf{f}_u$ 是在未知光源下拍摄的图像; $\mathbf{f}_c$ 是变换后的同一图像, 因此看起来好像是在标准光源下拍摄的; 而 $\mathbf{D}_{u,c}$ 是对角矩阵, 用于将在未知光源 u 下拍摄的彩色映射为在标准光源 c 下对应的彩色。通常, 将白色光源 (例如 D65) 用作标准参考光源。

对角模型可以通过假设相机灵敏度 $\rho^c(\lambda) = \delta(\lambda_c)$ 的狄拉克 (Dirac) 三角函数来得出。如果将其代入式(3.3), 则在波长 λ_c 处, 有

$$\mathbf{f}^c(\mathbf{x}) = m^b(\mathbf{x}) s(\lambda_c, \mathbf{x}) e(\lambda_c) \quad (3.27)$$

如果我们考虑两个不同的光源 e^1 和 e^2 , 则两者之间的关系为

$$\frac{\mathbf{f}_1^c(\mathbf{x})}{\mathbf{f}_2^c(\mathbf{x})} = \frac{m^b(\mathbf{x}) s(\lambda_c, \mathbf{x}) e^1(\lambda_c)}{m^b(\mathbf{x}) s(\lambda_c, \mathbf{x}) e^2(\lambda_c)} = \frac{e^1(\lambda_c)}{e^2(\lambda_c)} \quad (3.28)$$

因此, 可以通过 $\mathbf{f}_1 = \mathbf{D}_{1,2} \mathbf{f}_2$ 来模拟两个光源下相机响应之间的关系, 其中 $\mathbf{D}_{1,2}$ 是对角矩阵。

为了包括漫射光项, 芬莱森等^[34]将对角模型扩展了一个偏移量(o_1, o_2, o_3), 得到了对角偏移模型:

$$\begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} R_u \\ G_u \\ B_u \end{bmatrix} + \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix} \quad (3.29)$$

尽管此模型只是光源变化的近似值, 可能无法准确地模拟光度变化, 但它已被广泛接受为彩色校正模型, 并且是许多彩色恒常性算法的基础。还有一些对角变换的改进, 例如更改彩色基^[35]和应用光谱锐化^[36]。

3.5 彩色空间

光线从物体表面反射后, 可以由人类观察者或彩色相机进行检测和“测量”。到达传感器 (眼睛或相机) 的光是光源的光谱功率分布 $e(\lambda)$ 与物体的光谱反射率分布 $s(\lambda)$ 之间相互作用的结果, 并可使用 3.4 节中的任意一种 E 方程进行建模。这些分布可以转换为实际的彩色信号, 如下所示:

$$\mathbf{f}^c(\mathbf{x}) = \int_{\omega} E(\lambda, \mathbf{x}) \rho^c(\lambda) d\lambda \quad (3.30)$$

其中最简单的模型是朗伯模型, 它通过对可见光谱 ω 的每个波长处的三个分量的乘积进行积分来获得三刺激值 $\mathbf{f}^c$ 。这三个分量是光源的光谱功率分布 $e(\lambda)$ 、物体的反射率分布 $s(\lambda)$ 和传感器的灵敏度函数 $\rho^c(\lambda)$ 。三基色理论表明, 需要三个通道来生成人类全部可见的光范围, 因此需要定义三个相机灵敏度, 以指定传感器对输入光谱功率分布的灵敏度。

3.5.1 XYZ 系统

从式(3.30)可以明显看出, 相机的感光度会显著影响地最终色彩值。正如人们希望相机像人眼一样“感知”光谱, 因此需要对人体敏感度进行规范。但是, 正在运行的人类视觉

系统只能被视为一个黑匣子，可以向其询问诸如“这两种彩色是否相似”之类的问题。这正是如何获得有关眼睛敏感度的知识^[37]。需要展示两个面板，一个由任意彩色的测试光源来照亮，另一个由3个基色光源的混合来照亮。现在，要求观察者改变基色光源的强度，直到两个测试面板看起来相等为止，即不能在相对的面板之间区分彩色边缘。现在，3个基色光源的强度值（三刺激值）指示“彩色匹配对”。当选择了3个单色（窄波段）的基色光源，并采用各种波长的单色测试光源时，每个测试波长得到的强度值产生了所谓的彩色匹配函数。这些彩色匹配函数实质上描述了人类的彩色响应。不幸的是，对于上述实验中的某些波长，无法获得令人满意的彩色匹配。在这种情况下，实验者不得不将一个面板上的基色之一移动到另一个面板上，以添加到测试光源中。这将导致彩色匹配函数的值为负。为了避免这种情况，国际照明委员会（CIE）引入了3个虚构的原色 X 、 Y 和 Z ，以使相关的彩色匹配函数的三刺激值为正。

为了尽可能地匹配人类视觉系统，CIE 引入了两个标准：CIE 1931 的 2° 标准观察者（简称为 2° 标准观察者）和 CIE 1964 的 10° 标准观察者（简称为 10° 标准观察者）。第一个标准用于为视野狭窄的观察者建模（阅读距离在 10 英寸时约为 0.4 英寸），而后一个标准对应于较大视野的视觉匹配（阅读距离在 10 英寸时约为 1.9 英寸）。

使用标准观察者的 3 个彩色匹配函数，可以计算出 3 个数字（称为三刺激值），它们等效于标准观察者的响应：

$$X = \int_{\lambda} E(\lambda, \mathbf{x}) \bar{x}(\lambda) d\lambda \quad (3.31)$$

$$Y = \int_{\lambda} E(\lambda, \mathbf{x}) \bar{y}(\lambda) d\lambda \quad (3.32)$$

$$Z = \int_{\lambda} E(\lambda, \mathbf{x}) \bar{z}(\lambda) d\lambda \quad (3.33)$$

其中， $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 和 $\bar{z}(\lambda)$ 是 2° 标准观察者或 10° 标准观察者的 CIE 彩色匹配函数(图 3.4)。这些 XYZ 值可以转换为色度坐标以描述彩色的色度。

$$x = \frac{X}{X+Y+Z} \quad (3.34)$$

$$y = \frac{Y}{X+Y+Z} \quad (3.35)$$

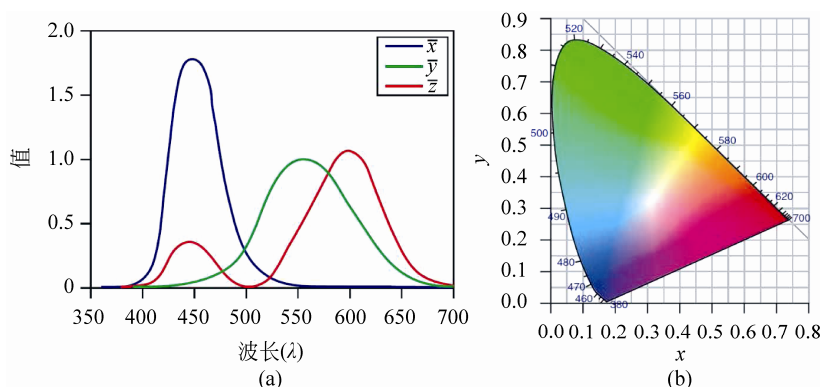


图 3.4 (a) CIE 1931 的 2° 标准观察者的彩色匹配函数。(b) CIE 1931 的 xy 色度图。马蹄形表示 xy 平面中可见彩色的色域。在外侧曲线（光谱轨迹）上指示了波长

$$z = \frac{Z}{X+Y+Z} \quad (3.36)$$

由于排除了彩色的强度并且这些色度值的总和等于 1，因此仅两个色度值足以描述彩色。但是，为了保留有关彩色和强度的完整信息，通常根据两个色度通道 x 和 y 以及强度通道 Y 进行指定，从而产生 xyY 彩色空间。可见光谱在 xy 平面上形成马蹄形，如图 3.4 所示。马蹄形以外的值对于人类是不可见的。从 xyY 到原始 XYZ 坐标的转换指定如下：

$$X = \frac{xY}{y} \quad (3.37)$$

$$Y = Y \quad (3.38)$$

$$Z = \frac{(1-x-y)Y}{y} \quad (3.39)$$

使用式(3.31)～式(3.36)，可以客观地将精确的数值分配给标准观察者的色觉。实际上，CIE 引入的 XYZ 系统是客观彩色测量的科学基础。在下一部分中，彩色系统（RGB）被推导出来以表达用于显示器和数码相机的彩色。

3.5.2 RGB 系统

电视行业已根据“标准”选择了 RGB 灵敏度，该灵敏度能够很好地匹配人眼。但是，由于一些制造商的光电传感器略有不同，因此 RGB 灵敏度曲线取决于设备。考虑用相机拍摄图像（未知的光源，或者最好情况下是对基色的估计），并在监视器上复制图像（已知但不同的光源和基色），或者在纸上打印图像（已知但不同的基色，未知的“光源”——纸张的反射率函数与观察纸张时的光源相结合）的情况，这点就立即显而易见。

XYZ 和 RGB 之间的转换由一组彩色基色（ xyY 坐标），参考白色和 γ 校正函数表示。 γ 校正函数用于在非线性显示设备上可视化线性彩色值，通常指定为幂定律表达式：

$$f_{\text{out}} = f_{\text{in}}^{\gamma} \quad (3.40)$$

其中 γ 需选择以匹配显示设备。所有 RGB 彩色空间都是为特定设备（或一组设备）定义的，并使用 γ 的特定值。唯一值得注意区别的是 sRGB 空间，稍后将进行讨论。

参考白色是名义上白色物体刺激的彩色。此值通常为未知值，此时必须进行假设。由于彩色的基色部分地取决于参考白色的值，因此 XYZ 和 RGB 之间的转换表达为彩色基色和参考白色的组合，其中参考白色通常是通过参考一种标准 CIE 光源（例如，D65、D50、C 等）来定义的。已经提出了许多不同的彩色匹配函数集，以能够在显示器上可视化彩色，从而产生了不同的 RGB 标准，例如 NTSC-RGB、PAL-RGB 和 sRGB。给定这些数据， XYZ 和 RGB 之间的转换可如下进行：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.41)$$

其中， (X_r, Y_r, Z_r) 、 (X_g, Y_g, Z_g) 和 (X_b, Y_b, Z_b) 是使用式(3.39)根据彩色基色计算的。此外，使用以下公式计算 (S_r, S_g, S_b) ：

$$\begin{bmatrix} S_r \\ S_g \\ S_b \end{bmatrix} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad (3.42)$$

其中 (X_w, Y_w, Z_w) 是参考白色, 通常用标准 CIE 光源表示。请注意, RGB 值必须是线性的并且在标称范围 $[0, 1]$ 中。从 XYZ 到 RGB 的转换只是式(3.41)的逆:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{bmatrix}^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.43)$$

存在不同的 RGB 工作空间, 每个工作空间导致不同的变换矩阵。在此指定两个经常使用的工作空间。NTSC 通常用于数码相机和视频, 它们使用 C 光源作为参考白色。彩色基色是 $\mathbf{xyY}_r = (0.6700, 0.3300, 0.2988)^T$ 、 $\mathbf{xyY}_g = (0.2100, 0.7100, 0.5868)^T$ 和 $\mathbf{xyY}_b = (0.1400, 0.0800, 0.1144)^T$, 可以进行如下 RGB_{NTSC} 和 XYZ 之间的转换:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.6069 & 0.1735 & 0.2003 \\ 0.2989 & 0.5866 & 0.1145 \\ 0.0000 & 0.0661 & 1.1162 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.44)$$

从 XYZ 到 RGB_{NTSC} 的转换是使用

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.9100 & -0.5325 & -0.2882 \\ -0.9847 & 1.9992 & -0.0283 \\ 0.0583 & -0.1184 & 0.8976 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.45)$$

另外, sRGB 是指定用于显示器、打印机和互联网的标准工作空间。该彩色空间使用略有不同的基色 $\mathbf{xyY}_r = (0.6400, 0.3300, 0.2127)^T$ 、 $\mathbf{xyY}_g = (0.3000, 0.6000, 0.7152)^T$ 和 $\mathbf{xyY}_b = (0.1500, 0.0600, 0.0722)^T$, 并使用 D65 作为参考白色。 RGB_{sRGB} 和 XYZ 之间的转换指定为

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9503 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.46)$$

从 XYZ 到 RGB_{sRGB} 的转换是使用

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2405 & -1.5371 & -0.4985 \\ -0.9693 & 1.8760 & -0.0416 \\ 0.0556 & -0.2040 & 1.0572 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.47)$$

还有其他多种 RGB 工作空间, 但超出了本书的范围。

3.5.3 对立彩色空间

一旦已知相机的光谱灵敏度与人类观察者相匹配, 便可以将任何编号方案分配给彩色。只要分配给每个可能的彩色值是一个唯一的数字, 不同的方案就能描述相同的彩色信息。但是, 不同的编码方案可能会突出显示某些彩色属性, 这将在后续部分看到。将相似的数字分配给不同的值甚至可能是有利的。例如, 通过将一种彩色的所有阴影都分配成单一值, 就可以获得由阴影和影调引起的强度变化的不变性, 如第 4 章所述。

RGB 的特性之一是 3 个通道的数值高度相关（例如，3 个通道之一的大数值通常也对应于其他两个通道的大数值）。对 RGB 彩色空间进行解相关会导致出现对立彩色空间。对立彩色理论始于大约 1500 年，达·芬奇（Leonardo da Vinci）得出的结论是，彩色是由黄色和蓝色、绿色和红色以及白色和黑色的混合物产生的。亚瑟·叔本华（Arthur Schopenhauer）也注意到相同的红-绿、黄-蓝和白-黑色的对立。对立彩色理论最终由爱德华·赫林（Edwald Hering）完成，他得出结论，眼睛的工作基于 3 种相对立的彩色。所谓的残像（残留影像）给出了对立彩色理论的证明：观察绿色样品一定时间会产生红色残像（另请参见图 2.3）。考虑彩色通道（即红-绿色和蓝-黄色），它们以两种不同的方式对立。首先，似乎没有彩色是任何对立色对的两个成员的混合体（例如，没有彩色看上去是黄蓝色，而经常遇到绿蓝色）。其次，一对对立色中的每个成员都表现出另一个，即，通过平衡添加两种对立彩色部分，将得到灰色。对立彩色理论已于 1950 年得到证实，当时在眼睛和大脑之间的光学连接中检测到对立彩色信号。

已经提出了几种模型来模拟对立彩色理论。在本书中，最简单的模型之一被称为**对立彩色空间**，可以通过简单地旋转 RGB 彩色系统来计算：

$$o_1 = \frac{R-G}{\sqrt{2}} \quad (3.48)$$

$$o_2 = \frac{R+G-2B}{\sqrt{6}} \quad (3.49)$$

$$o_3 = \frac{R+G+B}{\sqrt{3}} \quad (3.50)$$

请注意， o_1 大致对应于红-绿色通道， o_2 对应于黄-蓝色通道，而 o_3 对应于强度通道。除了直观之外，此彩色系统的另一个优点是它在很大程度上消除了 RGB 彩色通道之间的相关性。此外，对立彩色空间取决于设备并且在感知上不统一。

3.5.4 感知均匀彩色空间

为了克服这些缺点，CIE 提出了两种对立彩色系统，这些彩色系统被设计为在感知上是一致的（即，两种彩色之间的数字距离可能与感知上的差异有关），但这以不直观为代价。这两个系统是根据 XYZ 计算得出的，因此与设备无关。第一种彩色系统旨在描述光源彩色，称为 CIE $L^*u^*v^*$ ：

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_w} \right)^{1/3} - 16, & \frac{Y}{Y_w} > \varepsilon \\ 903.3 \left(\frac{Y}{Y_w} \right), & \frac{Y}{Y_w} \leq \varepsilon \end{cases} \quad (3.51)$$

$$u^* = 13L^* (u' - u'_w) \quad (3.52)$$

$$v^* = 13L^* (v' - v'_w) \quad (3.53)$$

$$u' = \frac{4X}{X+15Y+3Z} \quad (3.54)$$

$$v' = \frac{9Y}{X + 15Y + 3Z} \quad (3.55)$$

$$u'_w = \frac{4X_w}{X_w + 15Y_w + 3Z_w} \quad (3.56)$$

$$v'_w = \frac{9Y_w}{X_w + 15Y_w + 3Z_w} \quad (3.57)$$

其中 $\varepsilon = 216/24389 = 0.008856$ 。当强度低时（即 $(X + 15Y + 3Z)$ 接近零时），彩色通道 u^* 和 v^* 变得不稳定且毫无意义。第二种彩色系统旨在用于表面彩色，称为 CIE $L^*a^*b^*$ ：

$$L^* = 116f\left(\frac{Y}{Y_w}\right) - 16 \quad (3.58)$$

$$a^* = 500\left[f\left(\frac{X}{X_w}\right) - f\left(\frac{Y}{Y_w}\right)\right] \quad (3.59)$$

$$b^* = 200\left[f\left(\frac{Y}{Y_w}\right) - f\left(\frac{Z}{Z_w}\right)\right] \quad (3.60)$$

$$f(t) = \begin{cases} t^{\frac{1}{3}}, & t > \varepsilon \\ \frac{903.3t + 16}{116}, & t \leq \varepsilon \end{cases} \quad (3.61)$$

一方面，这些彩色空间在计算机视觉应用（例如图像检索或图像质量评估）中具有优势，这些应用的目的是与人类视觉保持一致。另一方面，在与人类视觉没有直接联系的应用中使用这些彩色空间（如体视或运动跟踪）并没有多大意义。在这些情况下，不执行耗时的从 RGB 到 $L^*a^*b^*$ 的非线性变换会更有效。

3.5.5 直观彩色模型

除了对立彩色空间外，到目前为止讨论的彩色系统还没有直观的表述。不同的彩色通道没有直观的含义。为此，引入基于艺术家推理的不同彩色系统。所有这些彩色系统都根据色调、饱和度和强度来表达彩色。不过，这些术语存在许多不同的定义，而且这些定义均未标准化。此外，通常将不同的缩写用于不同的定义，例如 HSV、HSI、HSL 等。在本书中，下标将用于指示相应的定义。

色调的常见定义之一是用光谱功率分布的主波长来描述，也就是说，色调借助我们通常用来描述任何给定彩色的词来描述：红色、蓝色、橙色、黄色等。在数学上，可以使用以下笛卡儿坐标到极坐标的转换按角度来计算：

$$H_{\text{RGB}} = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)(R-B)}\right) \quad (3.62)$$

此外，色调也可定义为参考线（例如水平轴）与彩色点之间的角度：

$$H_{\text{rgb}} = \arctan\left(\frac{r-1/3}{g-1/3}\right) \quad (3.63)$$

可以根据 CIE $L^*u^*v^*$ 或 CIE $L^*a^*b^*$ 计算与设备无关的色调版本：

$$H_{uv} = \arctan\left(\frac{v^*}{u^*}\right) \quad (3.64)$$

$$H_{ab} = \arctan\left(\frac{b^*}{a^*}\right) \quad (3.65)$$

请注意，对于非彩色，色调是不确定的（例如， $R = G = B$ ， $u^* = v^* = 0$ 或 $a^* = b^* = 0$ ）。

饱和度通常定义为彩色的纯度，当将更多的非彩色混合到彩色中时，饱和度会降低。完全不饱和的彩色与灰色轴重合，而完全饱和的彩色与纯色重合。数学上，饱和度定义为从一种彩色到非彩色轴的距离，但是可以使用不同的方式计算该距离。例如，

$$S_{\text{rgb}} = \sqrt{(r-1/3)^2 + (g-1/3)^2 + (b-1/3)^2} \quad (3.66)$$

$$S_{\text{RGB}} = 1 - \frac{\min(R, G, B)}{R + G + B} \quad (3.67)$$

$$S_{\text{HSL}} = \max(R, G, B) - \min(R, G, B) \quad (3.68)$$

$$S_{\text{HSV}} = 1 - \frac{\min(R, G, B)}{\max(R, G, B)} \quad (3.69)$$

请注意，对于暗像素（即 $R + G + B = 0$ ），饱和度是不确定的。与饱和度有关的另一个彩色通道是**色品**，可以从 CIE $L^*u^*v^*$ 或 CIE $L^*a^*b^*$ 彩色系统出发计算如下：

$$C_{uv}^* = \sqrt{(u^*)^2 + (v^*)^2} \quad (3.70)$$

$$C_{ab}^* = \sqrt{(a^*)^2 + (b^*)^2} \quad (3.71)$$

与饱和度类似，它们描述了彩色的纯度。

在坐标变换的背景下，通过对由 O_1 和 O_2 构成的对立彩色平面执行极坐标变换，可以计算出色调-饱和度-强度，具体如下：

$$\begin{bmatrix} h \\ s \\ i \end{bmatrix} = \begin{bmatrix} \arctan(O_1 / O_2) \\ \sqrt{O_1^2 + O_2^2} \\ O_3 \end{bmatrix} = \begin{bmatrix} \arctan\left[\frac{\sqrt{3}(R-G)}{(R+G-2B)}\right] \\ \sqrt{\frac{4}{6}(R^2 + G^2 + B^2 - RG - RB - GB)} \\ (R+G+B)/\sqrt{3} \end{bmatrix} \quad (3.72)$$

这种转换如图 3.5 所示。请注意，在此变换中谈论饱和度的强度更为正确，因为饱和度尚未像上面的公式中那样通过强度进行归一化。

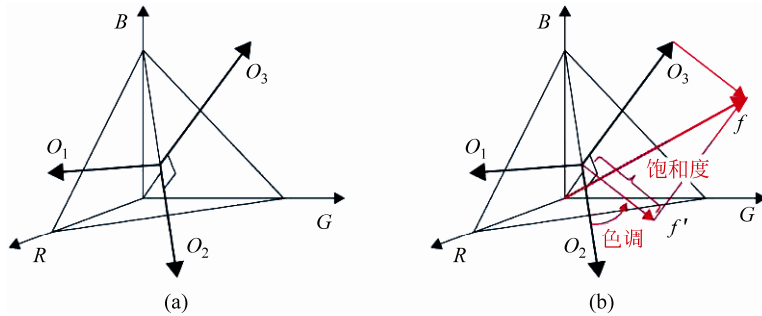


图 3.5 (a)显示了对立轴；(b)展示了对立坐标系与色调-饱和度-强度坐标系之间的关系。 f 在 O_3 上的投影是强度。令 f' 为 f 在 O_1O_2 平面上的投影；那么它的长度就是饱和度，平面中的角度就是色调

3.6 本章小结

本章中介绍了彩色图像形成的过程，讨论了几种基于物理的反射率模型。它们提供了分析导致彩色测量的各种物理原因的工具。更准确地说，它们将使我们能够区分图像中的阴影、镜面反射和材质变化。我们将看到此信息可用于彩色图像理解的许多阶段，例如改进的特征检测、特征描述和有意义的图像分割。

此外，列举了一组相关的彩色空间。根据手头的任务，一个彩色空间可能比另一个彩色空间更好用。本章介绍的反射率模型在第4章中用于分析彩色空间的光度特性。这将提供进一步了解哪种彩色空间最适合各种应用的信息。