

色彩空间基础



章佳杰

光学 话题的优秀回答者

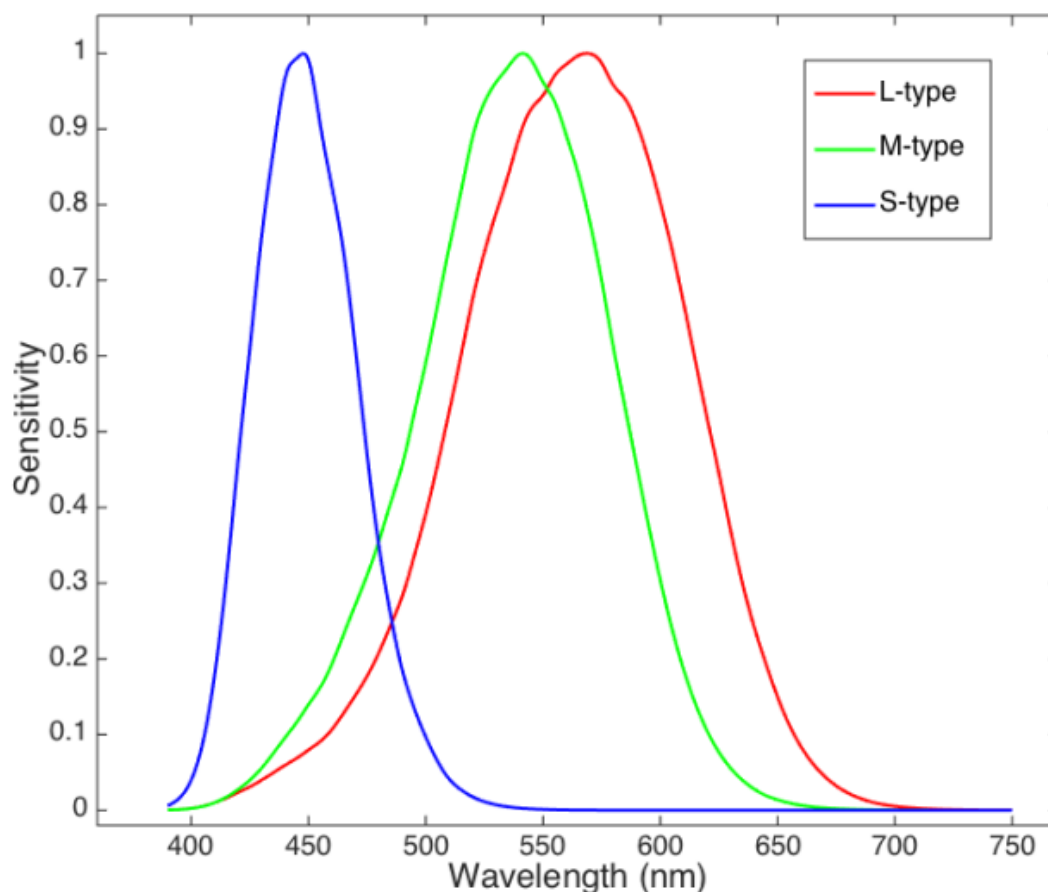
作为摄影师，你是不是遇到过这样的问题：相机设置里的色彩空间是什么意思？sRGB 和 Adobe RGB 有什么不同？为什么我的图片在 Photoshop 里看着好好的，一保存发布到网上色彩就变了？在这篇文章（以及接下来几篇）中，我将从理工科的视角，详细捋一捋有关色彩空间相关的概念，也算是自己的一点思考笔记。

本篇文章将分为几个部分。首先从几个实验规律出发，借助线性代数相关的概念，定义色彩空间，然后介绍并推导几个重要的色彩空间之间的关系。

1 实验基础

相信大家都熟悉「三原色」理论，也知道现代显示器是按照 RGB 模式来显示色彩，不知道有没有想过，为什么是「三原色」？为什么是「三」这个数字？为什么一定要选择 RGB 红绿蓝三种颜色作为原色呢？选其他行不行？用四种颜色行不行？知乎上有个问题 [红绿蓝三色是（唯一的）正交基吗？](#) 就提出了这个疑问，我也给出过 [我的回答](#)，可以参考。

追根溯源的话，得从我们人类的视网膜说起。大部分人类的视网膜上有三种感知颜色的感光细胞，叫做视锥细胞，分别对不同波长的光线敏感，称为 L/M/S 型细胞。三种视锥细胞最敏感的波长分别是橙红色（长波，Long），绿色（中波，Medium），蓝色（短波，Short）。这三种视锥细胞的归一化感光曲线如下图所示（图片数据来自 [CVLR](#)，我重新绘制），



可以看到 L 型视锥细胞与 M 型视锥细胞的感光曲线差别很小，实际上这两种视锥细胞起源于一次基因变异，在这之前人类可都是红绿色盲呢，多亏这个基因变异，让人类可以看到更加多彩的世界——这又是一个庞大的话题了，就此打住。

总之，大自然的这千千万万种颜色，在人类的眼里看到，最后传送到大脑里的信号，就只有这三种视锥细胞的电信号而已。根据这三种电信号的强弱，大脑解读成了不同的颜色。这就是三原色理论的生物学依据。

不仅如此，人类眼睛对不同颜色光线混合的反应还是 线性 的。根据 格拉斯曼定律 (Grassmann's Law)，两束不同颜色的光 C_1 和 C_2 ，假设某个视锥细胞对他们的反应分别是 r_1 和 r_2 ，现在将他们按照一个比例混合，得到第三种颜色 $C_3 = \alpha C_1 + \beta C_2$ ，那么视锥细胞对这个混合颜色的反应也将是前两个反应的线性叠加 $r_3 = \alpha r_1 + \beta r_2$ 。

格拉斯曼定律是一个实验规律，并没有物理或者生物学上的依据。然而这个规律大大简化了我们对人类彩色视觉系统的建模，并且给我们使用线性代数理论分析人类彩色视觉系统提供了一个前提和基础。

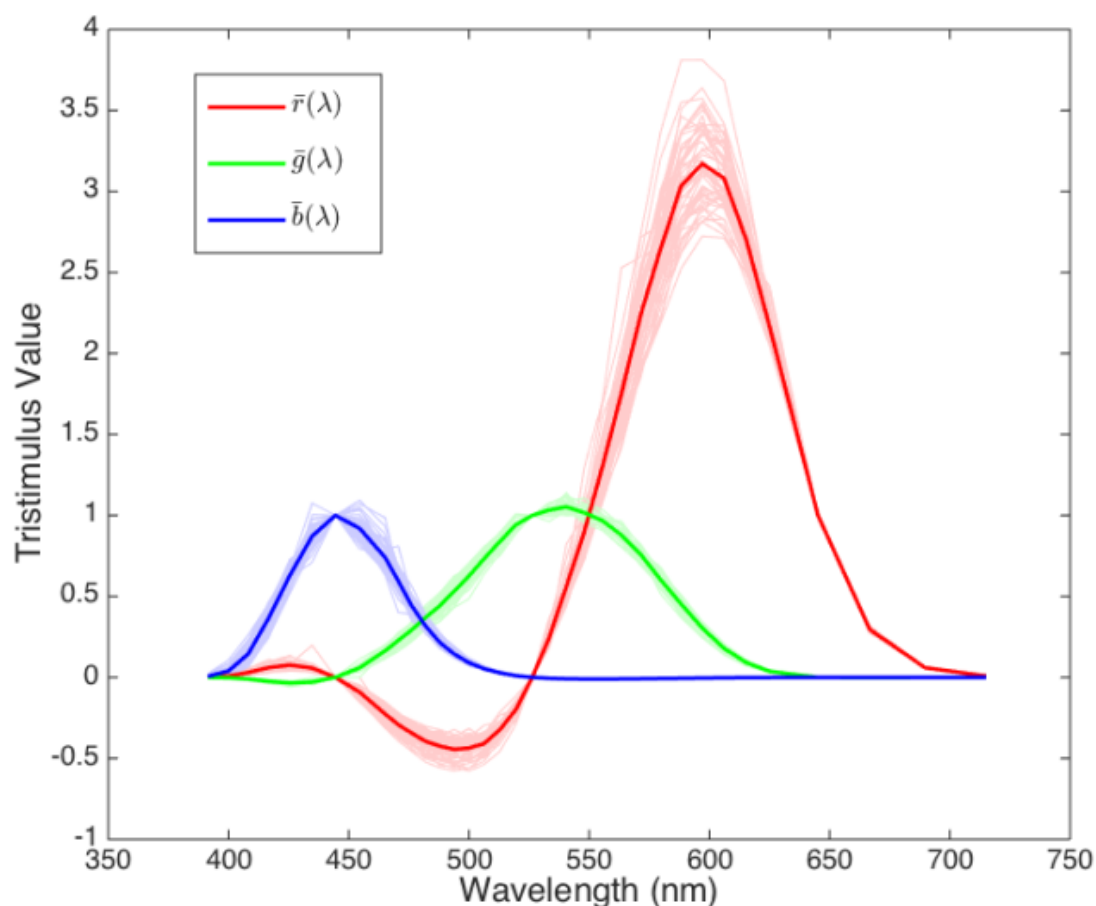
2 色匹配函数

前面已经提到，人类视网膜上有三种感知色彩的视锥细胞，所以理论上我们用三种颜色的光就可以混合出自然界中任何一种颜色来。在 20 世纪 20 年代，David Wright 和 John Guild 各自独立地领导了一些实验，通过三种颜色的光源进行匹配，得到了人眼对于不同颜色光的匹配函数。此后，多名科学家多次进行了类似的实验，加深了我们对人类彩色视觉的认识。

实验过程大致是这样的，把一个屏幕用不透光的挡板分割成两个区域，左边照射某个被测试的颜色的光线，这里记为 C （以下用大写字母表明颜色，用小写字母表明分量大小），右边同时用三种颜色的光同时照射，这里记为 R, G, B 。然后，调节右边三种颜色光源的强度，直到左右两边的颜色看上去一样为止。假设这个时候三种颜色的光源强度分别为 r, g, b ，那么根据光色叠加的线性性质，我们可以写出

$$C = rR + gG + bB$$

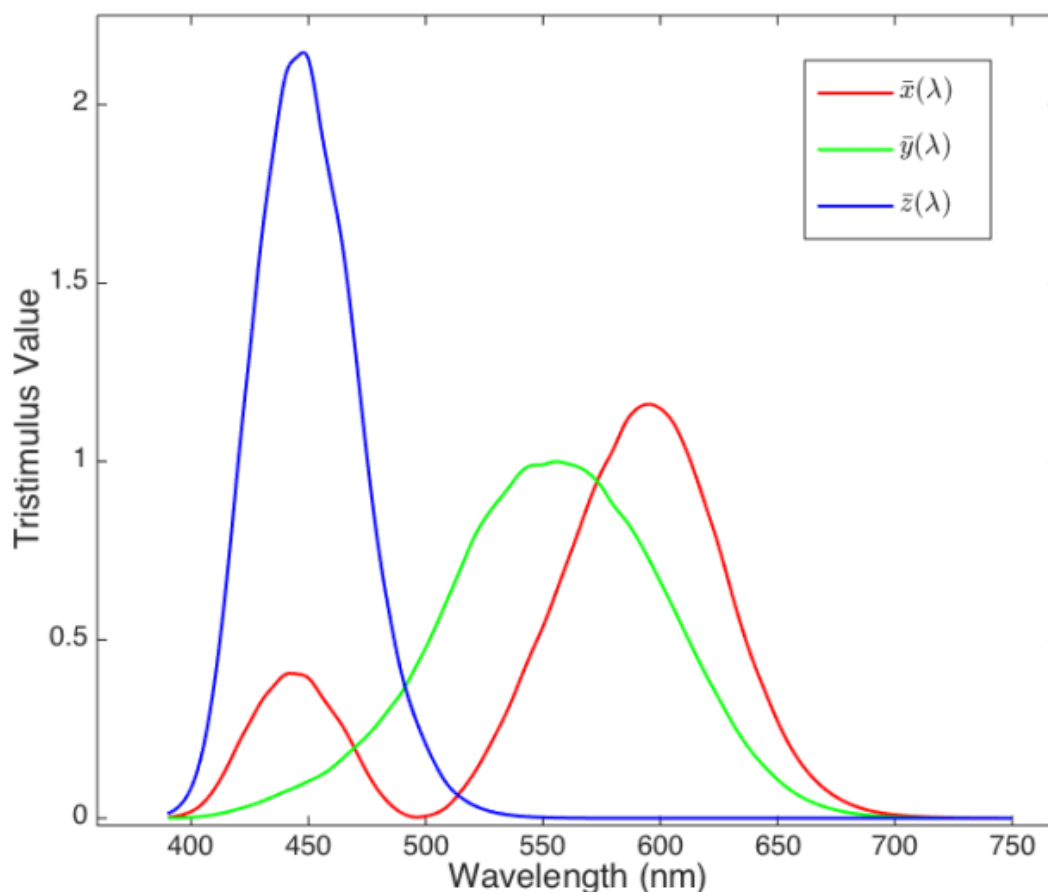
也就是说，只要按照 (r,g,b) 的分量来混合 (R,G,B) 三种颜色的光，就可以得到 C 这个颜色的光。于是在这一系列实验里，科学家们把左边的颜色按着光谱顺序，挨个测试了一遍，得到了纯光谱色的混合叠加的数据，这就是 色匹配函数 (Color Matching Function)，并且在这个基准下定义的色彩空间，就是 CIE RGB 色彩空间。下图是 CIE RGB 的色匹配函数曲线，数据来自 CVLR，我重新绘制。浅色的细线代表实验中不同参与者个人的色匹配函数曲线，中间深色的粗线代表数据的平均值。



可以看到，曲线上出现了负数，这是怎么回事？回想一下前面描述的实验过程，左边是被测试的光色，右边是可调节的三色光的混合。如果碰到一种情况，右边三色光无论如何调节比例，都不能混合出左边的颜色，比如某种颜色的光强度已经减小为 0 了，然而看趋势还需要继续减小才能与左边的光色相匹配，怎么办？这时候需要往左边的光色中混入三色光中的一种或者几种，继续调节，直到两边的颜色匹配。在左边（被测试）的光色中添加，那就是相当于在右边的混合光中减去，这就导致了色匹配函数曲线上出现了负数。实际上，这相当于就是光线的「减法」了。

比如，对于 555nm 的黄色光，色匹配函数的值是 (1.30, 0.97, -0.01)，意味着将 1.30 份的红光与 0.97 份的绿光混合放在右边，左边放上 1 份的 555nm 的黄光，以及 0.01 份的蓝光，这样左右两边的光色看上去就一样了。

因为有部分出现了负数，在使用和计算上都有不方便，因此就对这个匹配函数进行了一下线性变换，变换到一个所有分量都是正的空间中。变换后的色彩空间就是 CIE XYZ 色彩空间。（图片数据来自 [CVLR](#)，我重新绘制）



CIE RGB 色彩空间和 CIE XYZ 色彩空间是完全等价的，两者只是差了一个线性变换。由于允许「减法」的存在，因此 CIE RGB 空间是能够表示所有颜色的；同样的，CIE XYZ 空间也能。

3 从线性空间的角度理解色彩空间

以上的实验基础提示我们，色彩空间和线性代数中的线性空间之间具有某种相似性。我们可以看到，由于人类有三种感知色彩的视锥细胞，自然界千千万万的色彩被眼睛接收后，可以用三个数值来表征。而格拉斯曼定律也揭示了色彩叠加的线性性质。这似乎意味着，色彩空间就是一个 3 维的线性空间。事实上也的确如此（详细的论证参见末尾小节）。

自然界本身是没有「颜色」这个属性的，只有对不同波长光线的反射率/透过率，到达人眼中的，显然是一个连续的光谱分布函数。数学上，这是一个无穷维的函数空间（巴拿赫空间）。而人眼内的三种视锥细胞，它们的感光特性曲线相当于是在这个无穷维的函数空间中建立了三个基底。任何一个光谱分布进来，三种视锥细胞被激发。由于色视觉响应的线性性，这一过程相当于光谱分布函数与三个基底做内积，或者说，「投影」到这三个基底上。

从这个观点看，人类的色视觉，是相当于在自然界所有颜色的无穷维函数空间中取了一个三维的投影。这个三维空间的基底，既可以是视锥细胞的感光特性曲线（我们的大脑就用的是这套），当然也可以是选取三种颜色的光进行组合（CIE RGB 空间），甚至还可以是用实际中不存在的「光线」进行组合（CIE XYZ 空间）。既然这几个空间实际上是同一个线性空间，只不过由于选择了不同的基底而有不同的表达形式，那么根据线性代数的结论，这几个空间的表述形式之间，只需要通过矩阵乘法就可以完成转换，这是完全的线性变换。

当然，色彩空间并不是真正数学意义上的三维线性空间。由于不存在真正数学意义的「减法」，在实际应用中是有所限制的。数学中的「线性组合」在这里就要被替换为「锥组合」，也就是每个分量都必须是大于等于 0 的。

至此我们终于可以回答开头的部分问题了，为什么是三原色？因为人类对色彩的感知结果位于一个三维的线性空间中。最少需要三种颜色的光才能有足够的表达能力来表现各种颜色。为什么选 RGB 作为三原色？因为色彩空间不是真正数学意义上的线性空间，从工程角度考虑，以 RGB 作为三原色，能让显示器能够显示更多的颜色（此外，最初测试人眼对 RGB 三色光的色匹配曲线，也是希望能尽量单独地刺激三种视锥细胞）。

4 设备相关的 RGB 色彩空间

如前文所述，色彩空间的基底的选择有一定的任意性。事实上，如果允许真正的减法存在，那么选择哪三种颜色作为基底是无关紧要的。不过由于实际中我们不能对色光采用减法，只能使用「锥组合」而非「线性组合」，这时候到底如何选取基底就显得重要了。好的基底不仅能表达的颜色更丰富，而且工程上也易于稳定地实现。

由于 CIE XYZ 空间是一个很方便的线性空间，与具体设备无关，因此常用来做各种颜色空间转换的中间媒介。设想某个颜色的光，经过色匹配函数的计算，得到了三个 XYZ 的值，如果直接将这三个值作为 RGB 颜色显示到屏幕上，显然是不对的。我们必须把 XYZ 的值转换到屏幕的 RGB 空间中的值。

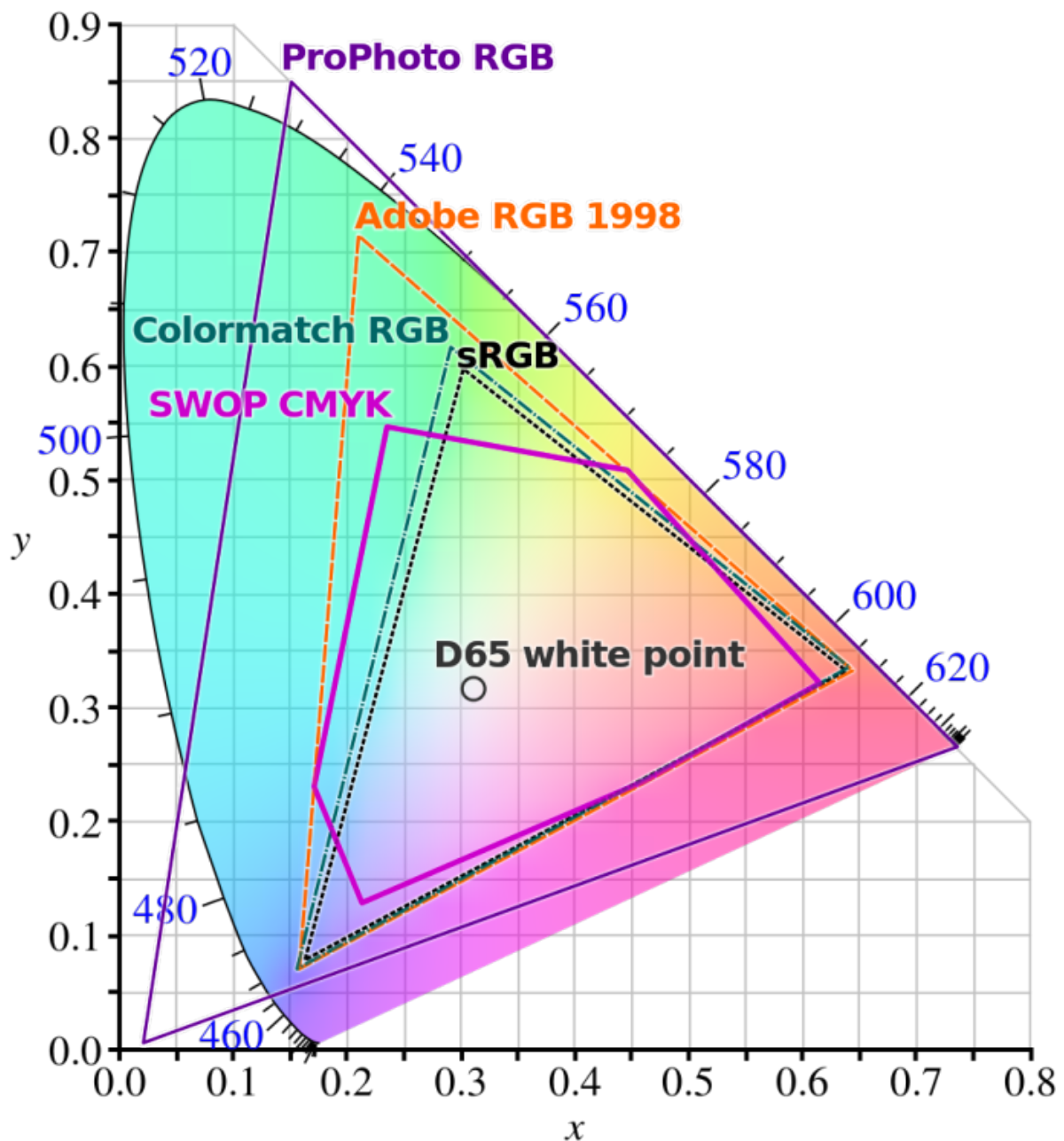
$$\begin{bmatrix} R_{\text{lin}} \\ G_{\text{lin}} \\ B_{\text{lin}} \end{bmatrix} = \mathbf{M} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

这里下标 lin 代表线性空间， \mathbf{M} 是转换矩阵。得到线性 RGB 空间的表达之后，还需要经过 gamma 校正，才是最终在屏幕上显示的 RGB 的值。一般的 gamma 校正过程为 $C = C_{\text{lin}}^{1/\gamma}$ ，其中 C 代表 RGB 的某个分量， γ 值通常为 2.2。对于如何计算的细节，我希望在 [下一篇文章](#) 中进行详细说明。

对于 RGB 色彩空间来说，关键点在于两个：1. 如何选择三个作为基底的颜色；2. 如何定义白色。一旦选好这两个关键参数，那么从 CIE XYZ 空间到设备的 RGB 空间的转换就完全确定了。我们平时常说的 sRGB 空间和 Adobe RGB 空间，他们的区别就在于这两个关键参数的定义不同。常用的 RGB 空间的转换矩阵如下表：

色彩空间	转换矩阵
Adobe RGB	$\begin{bmatrix} 2.0413690 & -0.5649464 & -0.3446944 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0134474 & -0.1183897 & 1.0154096 \end{bmatrix}$
sRGB	$\begin{bmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{bmatrix}$
Apple RGB	$\begin{bmatrix} 2.9515373 & -1.2894116 & -0.4738445 \\ -1.0851093 & 1.9908566 & 0.0372026 \\ 0.0854934 & -0.2694964 & 1.0912975 \end{bmatrix}$
CIE RGB	$\begin{bmatrix} 2.3706743 & -0.9000405 & -0.4706338 \\ -0.5138850 & 1.4253036 & 0.0885814 \\ 0.0052982 & -0.0146949 & 1.0093968 \end{bmatrix}$

不同的 RGB 空间能表示的范围如下图所示（这个图在 [以后的文章](#) 中再详细解释），



可以看到，不同的 RGB 空间所能表示的颜色范围是不一样的，并且我们可以推断出，即使是同样的 RGB 分量，在不同的 RGB 空间中所代表的颜色也是不一样的。所以我们在描述一个 RGB 颜色的时候，不仅需要描述它的 RGB 三个分量，还要说明是在哪个空间，这就是 ICC 文件的作用。

很多数码相机都可以设置色彩空间，常见的有 sRGB 和 Adobe RGB，从上面的图中我们可以看到，Adobe RGB 所能表达的色彩比 sRGB 要丰富很多。然而常见的网络环境下图片的色彩空间是 sRGB，有很多浏览器不能正确地解析图片自带的色彩空间说明，默认按照 sRGB 来进行解析。如果使用相机直出的 JPG 文件直接上传，或者说在后期处理过程中没有进行色彩空间转换，保留了相机设置的 Adobe RGB 空间，那么在浏览器中看到的图片很可能与 Photoshop 中看到的不一樣。浏览器很可能会将一张 Adobe RGB 空间中的图片解释为 sRGB 空间下的图片，引起颜色偏差。由于 sRGB 是目前屏幕

显示的「事实标准」，大多数屏幕空间都在 sRGB 内（这是颜色复现设备本身决定的），所以我的建议是，对用于网络交流目的的图片，统一转换到 sRGB 中进行保存。

5 有关色彩空间是线性空间的数学论证

回顾一下线性代数中构成线性空间的几个要素：

交换律： $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$

结合律： $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$

这两条容易理解，显然，交换不同色光叠加的顺序并不会改变叠加的结果。

加法零元： $\exists \mathbf{0}, \forall \boldsymbol{x}, \text{s.t. } \boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$

这一点也很好理解，加法零元就是全黑，什么光都没有。任何一种色光叠加一个全黑那还是它自身。

加法逆元： $\forall \boldsymbol{x}, \exists (-\boldsymbol{x}), \boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$

回想前文对实验中需要「负数」数量的光的处理手法，我们可以在左边叠加光线，这就相当于在右边减去光线。如果承认这种形式的「减法」，那么加法逆元也是顺理成章了：在另一边加入自身。

乘法幺元： $\exists 1, \forall \boldsymbol{x}, 1 \cdot \boldsymbol{x} = \boldsymbol{x}$

这也很明显，数字 1 就可以作为乘法幺元。

乘法分配率：

$\forall \boldsymbol{x}, \boldsymbol{y}, \alpha, \text{s.t. } \alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$

$\forall \alpha, \beta, \boldsymbol{x}, \text{s.t. } (\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}$

根据格拉斯曼定律，这一点也是实际上成立的。

综上所述，人类对色彩的感知在实验层面满足线性性质，人眼感知的色彩空间是一个 3 维线性空间。当然，容易想见，通常我们选取的基底（比如 CIE XYZ 空间的色匹配函数），是不正交的。