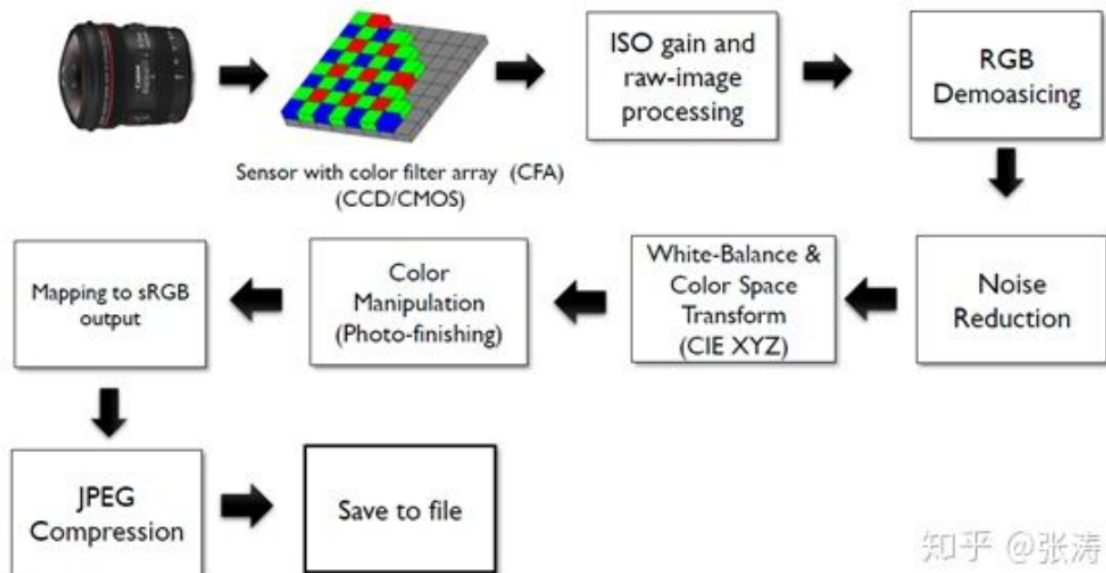


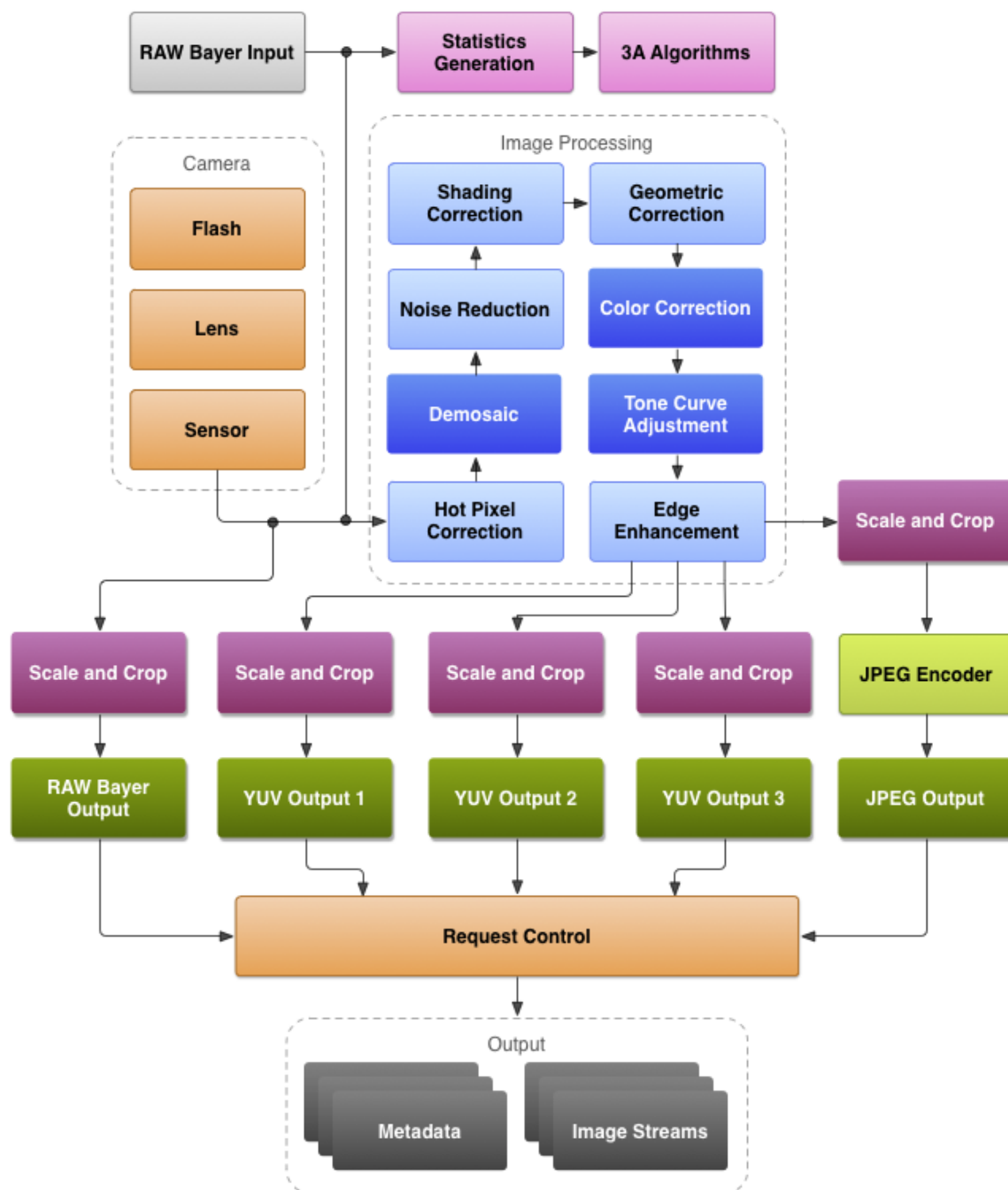
ISP(Image Signal Processor)

by Wang Jiawei

contact: wangjiawei@mail.ustc.edu.cn

ISP(Image Signal Processor), 即图像处理, 主要作用是对前端图像传感器输出的信号做后期处理, 主要功能有**线性纠正**、**噪声去除**、**坏点去除**、**内插**、**白平衡**、**自动曝光控制**等, 依赖于ISP才能在不同的光学条件下都能较好的还原现场细节, ISP技术在很大程度上决定了摄像机的成像质量。





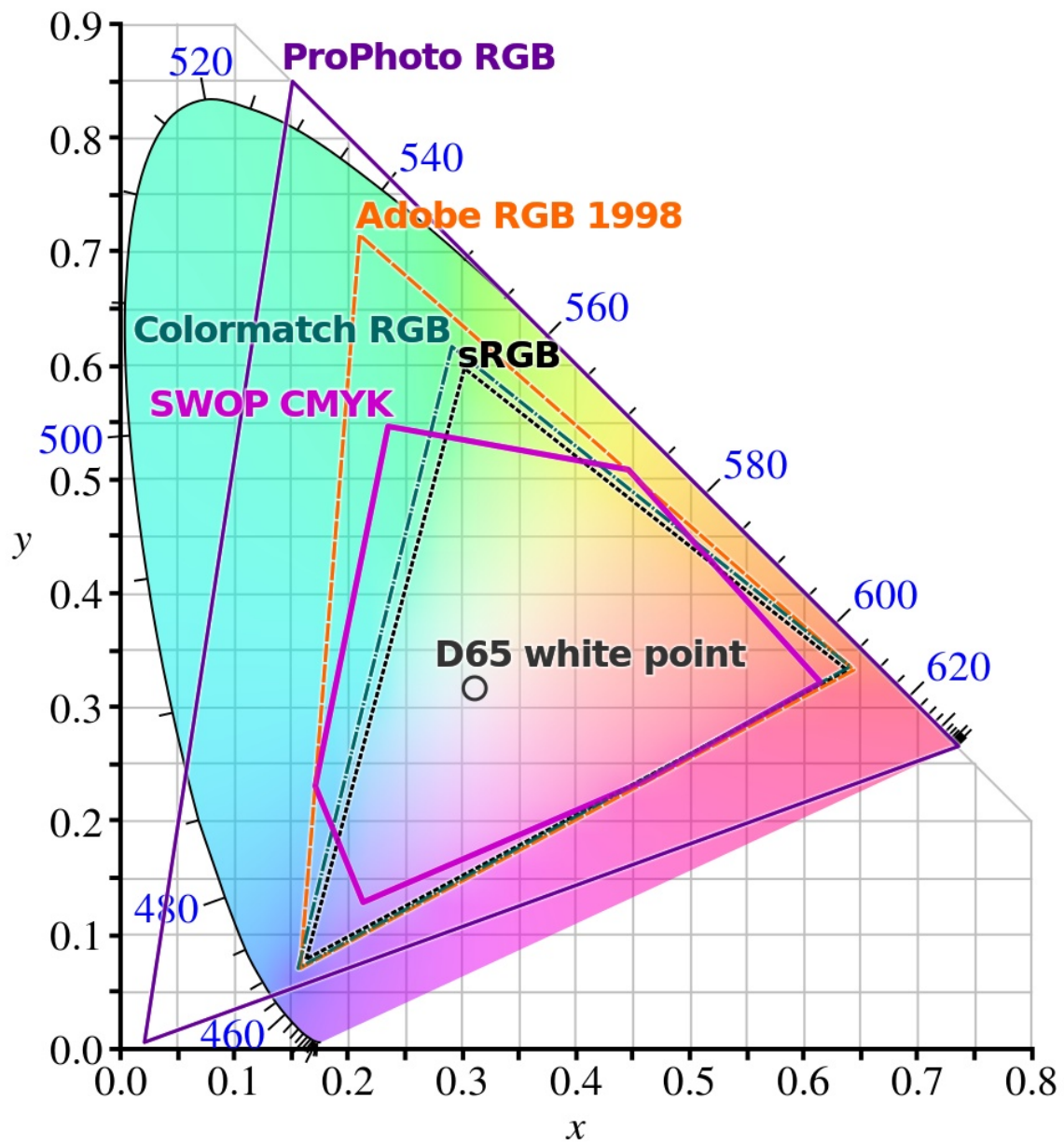
<http://kernel.meizu.com/camera-isp-intro.html>

<https://www.zhihu.com/question/61075213/answer/954905390>

成像原理

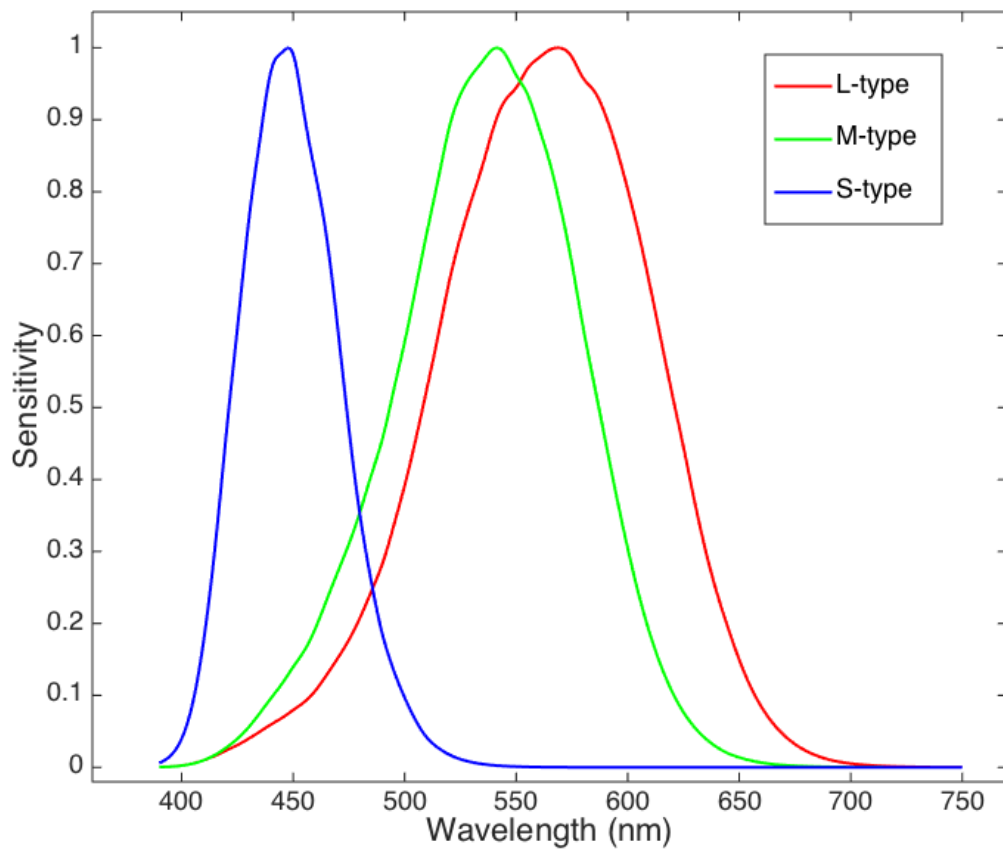
色彩空间

(参考 <https://zhuanlan.zhihu.com/p/24214731>)



1.

人类视网膜上的三种感知颜色的感光细胞（视锥细胞）分别对特定波长的光最敏感：



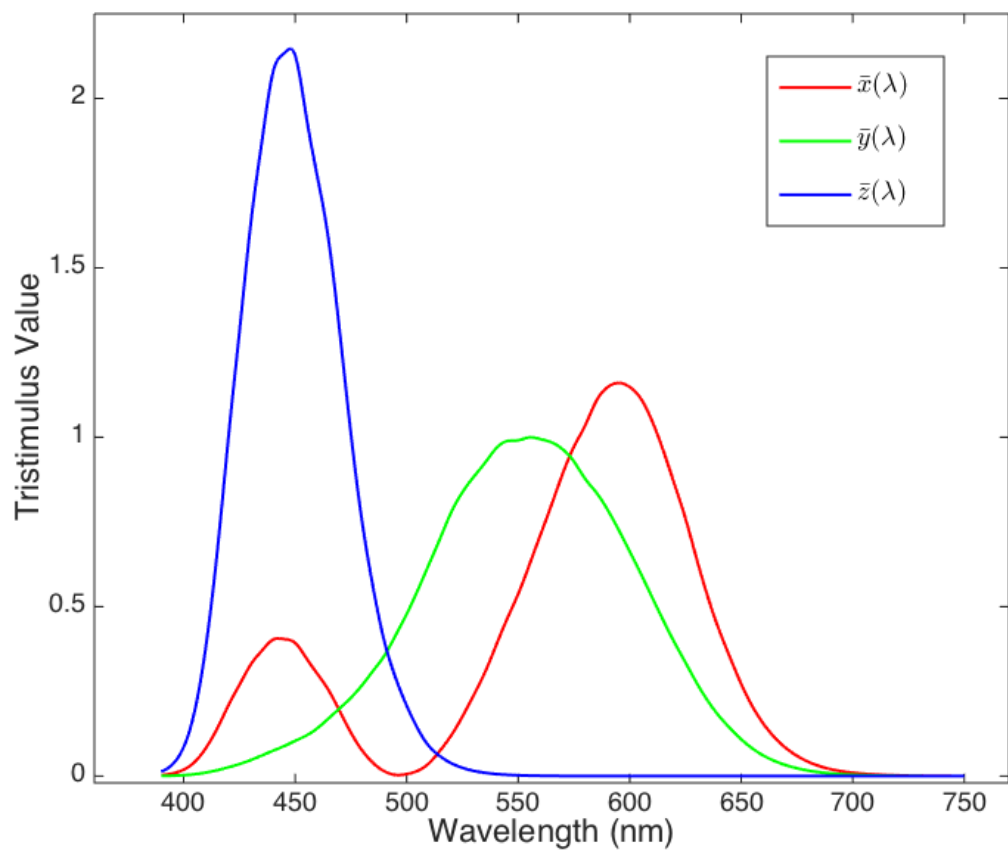
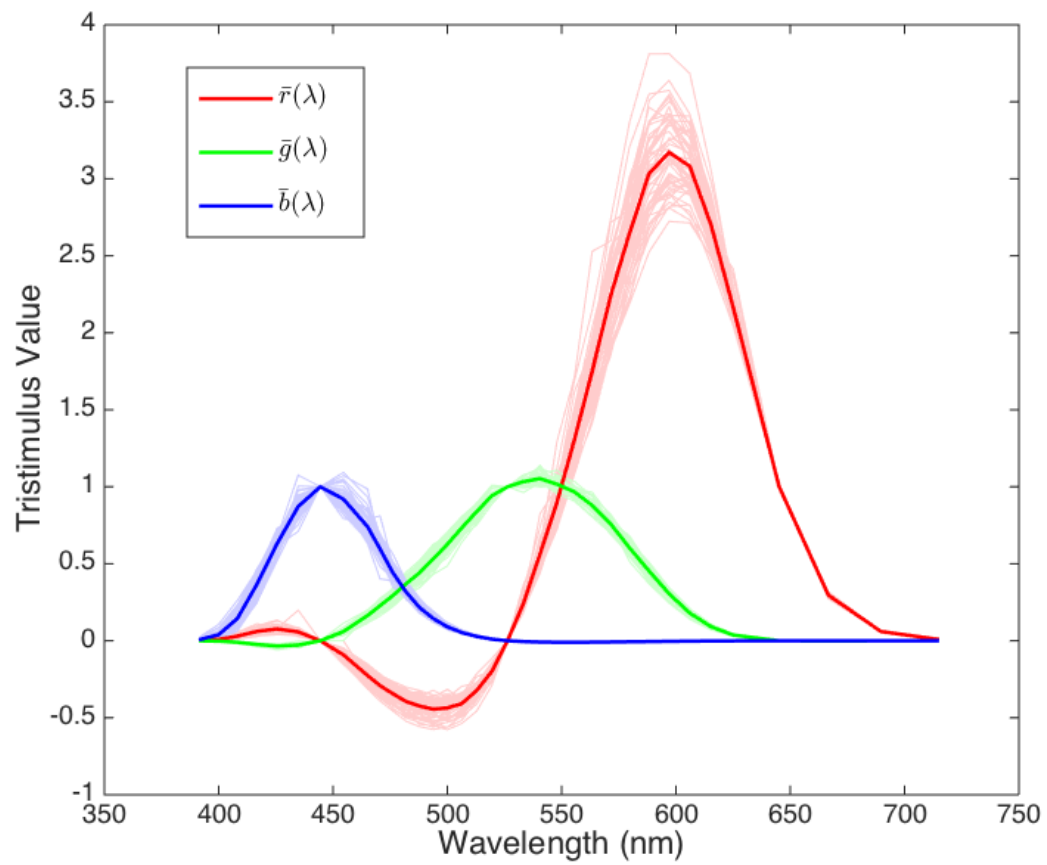
根据三种视锥细胞的电信号，大脑会解读成不同的颜色。

2.

人类眼睛对不同颜色光线混合的反应是线性的，根据一个实验定律:格拉斯曼定律，对混合颜色的反应会是各自的线性叠加。

3.

通过实验，得出色匹配函数：对于不同波长的光，都可以使用三原色的混合（匹配函数）得出。由于是实验，取了平均得到下面的深色的线，这定义的色彩空间为CIE RGB。由于有些光没办法通过完全正的混合得到，而可以使用光线的减法得到，所以存在强度为负数的情况。又可通过线性变换，变换到分量全是正的空间，得到CIE XYZ



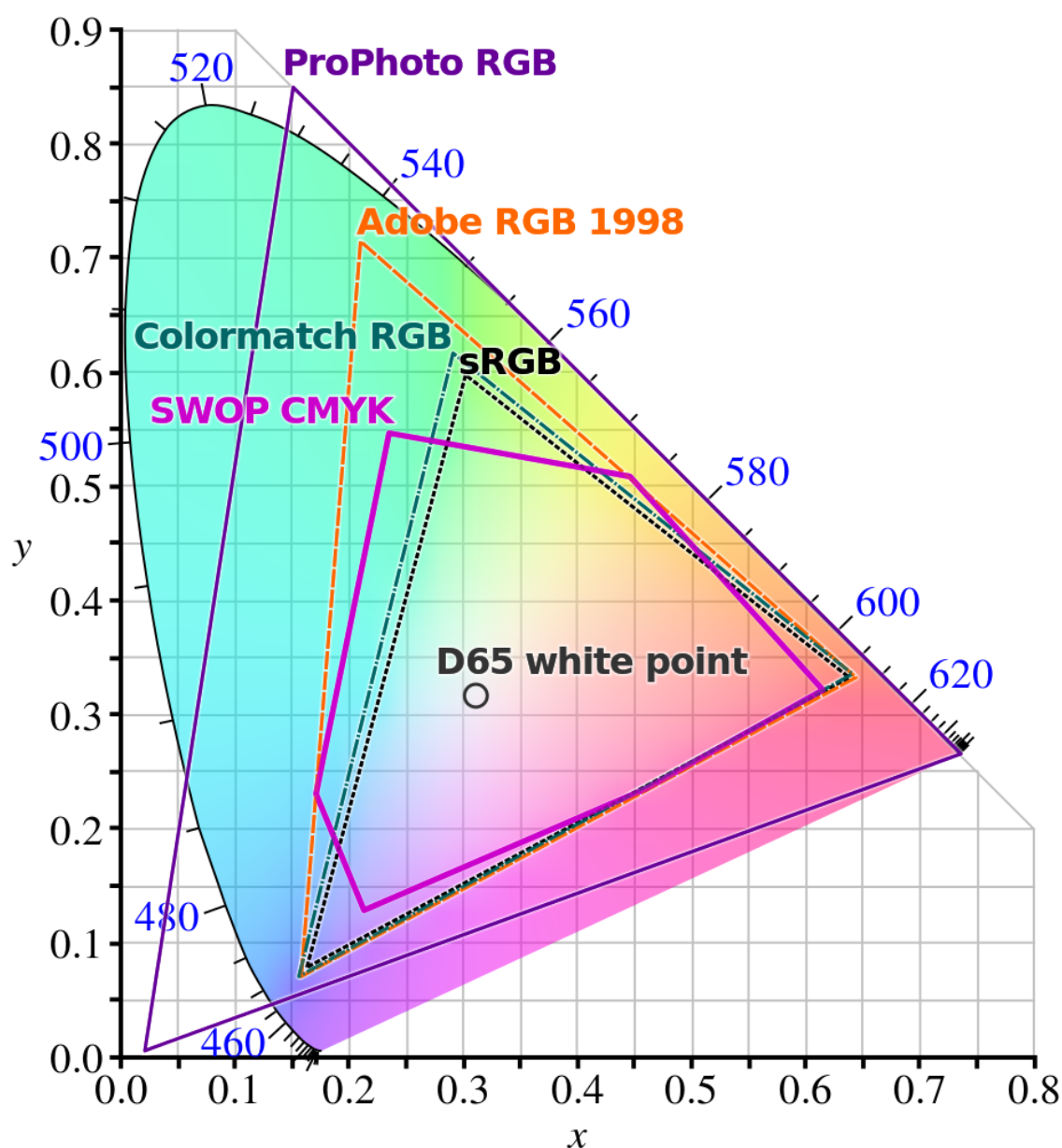
自然界本身是没有「颜色」这个属性的，只有对不同波长光线的反射率/透过率，到达人眼中的，显然是一个连续的光谱分布函数。数学上，这是一个无穷维的函数空间（巴拿赫空间）。而人眼内的三种视锥细胞，它们的感光特性曲线相当于是在这个无穷维的函数空间中建立了三个基底。任何一个光谱分布进来，三种视锥细胞被激发。由于色视觉响应的线性性，这一过程相当于光谱分布函数与三个基底做内积，或者说，「投影」到这三个基底上。

而这个基底可以取很多种，只不过表达形式不同，通过矩阵乘法便可以完成转换。又色彩空间并不是真正数学意义上的三维线性空间，比如不存在减法，所以实际应用有所限制。

又因为RGB能让显示器显示更多颜色，所以选择RGB作为三原色。

5.

CIE XYZ空间经常作为中间媒介，通过线性变换将XYZ值转换到不同的RGB颜色空间中，每种RGB颜色空间由白色的定义以及基底颜色所影响。所以有很多种RGB颜色空间。



REVERSE IMAGING PIPELINE FOR RAW RGB IMAGE AUGMENTATION

看到的一篇图像增强论文，感觉思路很有意思，反向获得光谱图，然后再生成质量更高的图片，没太细看。 http://vision.cs.tut.fi/data/publications/icip2019_reverse_imaging_pipeline.pdf

色彩恒常性

什么是色彩恒常性？ what is Color Constancy?

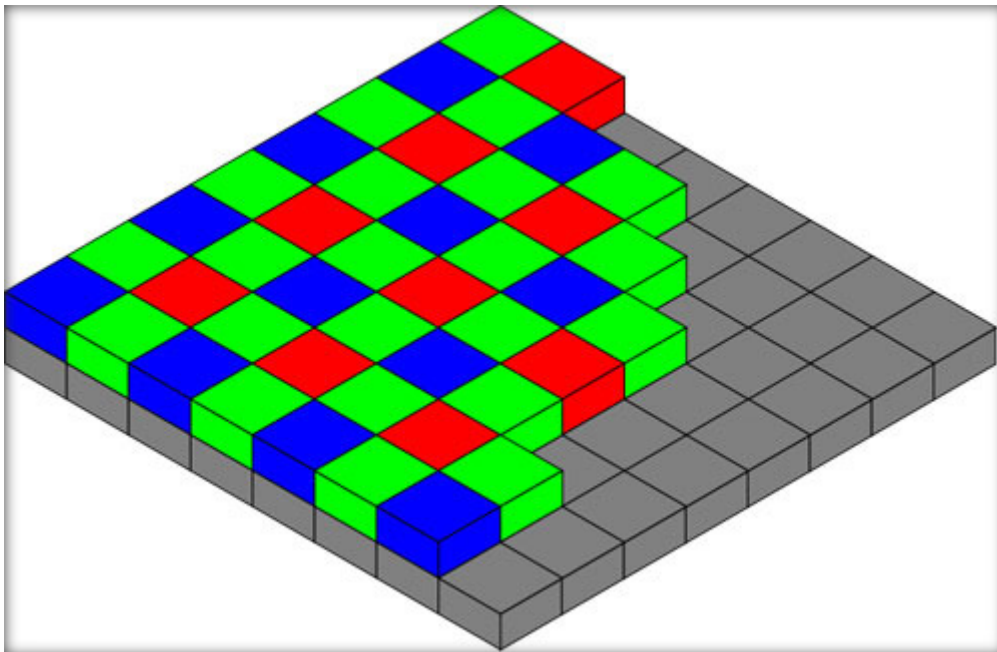
首先，颜色并不是一个实实在在的东西，它只是我们大脑和视网膜处理的结果。我们的视觉系统有一个特殊的功能就是在光源变化的情况下，依然能够确定物体的真实颜色，这种鲁棒性是计算机等处理机器所不具有的。色彩的恒常性，也就可以简单的理解为，在外界光源变化的情况下，依然能够保留物体原始的色彩。我们研究色彩恒常性，提出各种算法，目的就是让计算机也具有人类视觉系统的这个特殊功能，在不同光源的照射下还原图像的本来颜色。White Patch Retinex算法就是其中的一个最简单的算法。在后面的实验结果中，我们也可以看出，由于黄色光照下而整体颜色偏黄的餐桌，经过White Patch Retinex算法处理后，桌布恢复了白色。

原文链接：<https://blog.csdn.net/zhangjing11292/article/details/51089160>

拜尔阵列

<https://blog.csdn.net/youmingyu/article/details/52572977>

关于拜尔阵列一个很有意思的科普



首先光线进入镜头被这些滤镜分解成一个个单色的光，传感器记下每个值，得到图片的Raw 数据，再通过一定方法猜出每个像素点的RGB值(去马赛克)，再进行更多的后期处理，最后得到呈现出的照片。

去马赛克

Demosaicing的意思是如何将采集到的Bayer单通道像素值转成RGB三通道的数据。例如下图，这是我们采集到的单通道图像，我们需要估计R5位置处的G5和B5。目前常用的方法是：

1. 简单插值



$$\begin{aligned} G5 &= \frac{G2 + G4 + G6 + G8}{4} \\ B5 &= \frac{B1 + B3 + B7 + B9}{4} \end{aligned}$$

2. edge-aware插值

If $(|G2-G8| \&\& |(G4-G6)| \text{ both } < \text{Thres})$:

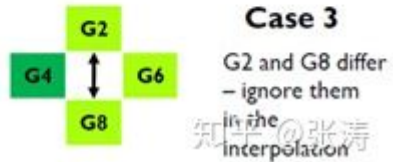
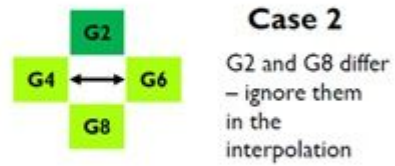
$$G5 = \frac{G2 + G4 + G6 + G8}{4}$$

elseif $(|G2-G8| > \text{Thres})$:

$$G5 = \frac{G4 + G6}{2}$$

else:

$$G5 = \frac{G2 + G8}{2}$$



Deep Joint Demosaicing and Denoising

白平衡

传统相机中的自动白平衡是通过灰度世界原理，假设RGB三通道的均值相同。

Hu, Yuanming, Baoyuan Wang, and Stephen Lin. "Fc4: Fully convolutional color constancy with confidence-weighted pooling." CVPR 2017.

Afifi, Mahmoud, et al. "When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images." CVPR 2019.

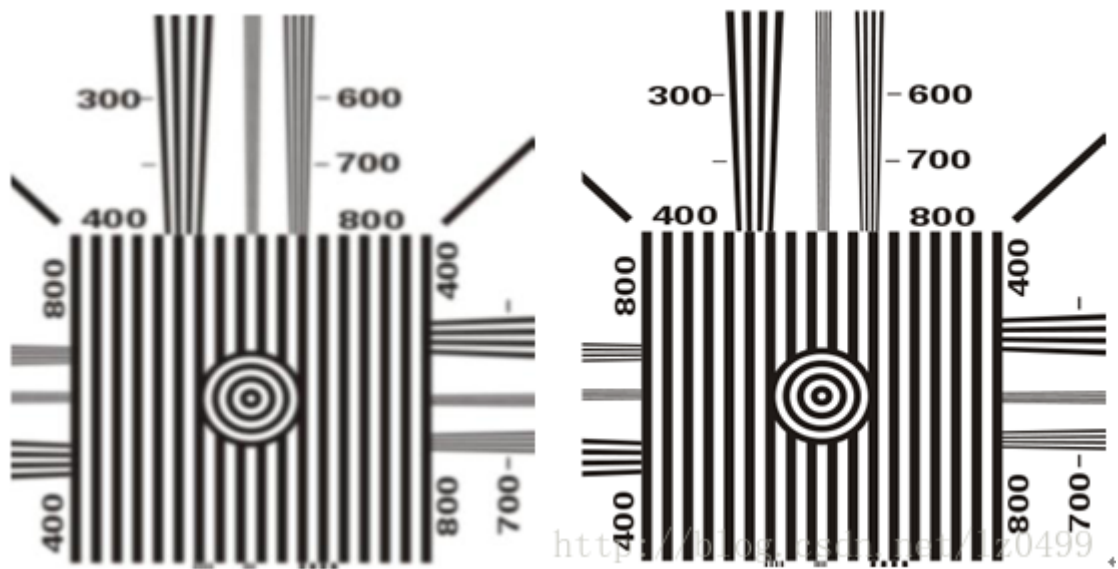
Denoise

使用 cmos sensor 获取图像，光照程度和传感器问题是生成图像中大量噪声的主要因素。同时，当信号经过 ADC 时，又会引入其他一些噪声。这些噪声会使图像整体变得模糊，而且丢失很多细节，所以需要去噪。传统的去噪方法有均值滤波、高斯滤波等。

但是，一般的高斯滤波在进行采样时主要考虑了像素间的空间距离关系，并没有考虑像素值之间的相似程度，因此这样得到的模糊结果通常是整张图片一团模糊。所以，一般采用非线性去噪算法，例如双边滤波器，在采样时不仅考虑像素在空间距离上的关系，同时加入了像素间的相似程度考虑，因而可以保持原始图像的大体分块，进而保持边缘。

Sharp锐化

CMOS输入的图像将引入各种噪声，有随机噪声、量化噪声、固定模式噪声等。ISP降噪处理过程中，势必将在降噪的同时，把一些图像细节给消除了，导致图像不够清晰。为了消除降噪过程中对图像细节的损失，需要对图像进行锐化处理，还原图像的相关细节。如下图所示，左图是未锐化的原始图像，右图是经过锐化之后的图像。



<https://blog.csdn.net/lz0499/article/details/71156291>