

# HDR技术调研综述

by Wang Jiawei

contact: [wangjiawei@mail.ustc.edu.cn](mailto:wangjiawei@mail.ustc.edu.cn)

## HDR技术调研综述

Introduction

Basic Theory And Correlation Techniques

动态范围测量

感知动态范围

动态范围的不同表示方法

最小可分辨差异 (JND)

显示模型和gamma 校正

gamma修正

显示模型

HDR存储格式及压缩

编码格式

HDR 文件格式

色调映射

分类

代数解释

色调映射的主要方法

光照和反射的分离

Forward visual model

Forward and inverse visual models

约束映射问题

Perceptual effects for the enhancement of tone-mapped images

逆色调映射

动态范围恢复

LDR像素值线性化

动态范围扩展

抑制轮廓和量化误差

恢复过曝和欠曝区域纹理

HDR图像评价指标

Visual difference predictor for HDR images

Luminance-independent

Metrics

Tone-mapping metrics

**整理思路参考:** High Dynamic Range Imaging Rafał K. Mantiuk, Karol Myszkowski and Hans-Peter Seidel  
April 18, 2016

## Introduction

---

**动态范围**（英语：dynamic range）是可变化信号（例如声音或光）最大值和最小值的比值。也可以用以10为底的对数（**分贝**）或以2为底的对数表示。在计算机图形学和数字摄影中，**动态范围指一张照片中所记录下来的最暗的部分和最亮的部分之间的亮度差异，用指数表示**，比如jpeg格式的照片，亮度从0到255总共256级， $2^8=256$ ，也就是说用jpeg格式记录的照片最大动态范围只有8档。而**宽容度是相机记录动态范围的能力**，一般相机的记录能力都超过8档，超过8档的数据，jpeg是没办法记录的，这时候就要靠raw格式了，raw目前用得最多的是14bit的（也有12bit的raw），也就是说理论上raw可以记录14档动态范围。**高动态范围成像**（英语：High Dynamic Range Imaging，简称HDRI或HDR），在计算机图形学与电影摄影术中，是用来实现比普通数位图像技术更大曝光**动态范围**（即更大的明暗差别）的一组技术。高动态范围成像的目的就是要正确地表示真实世界中从太阳光直射到最暗的阴影这样大的范围亮度。



上图摘自 <https://zhuanlan.zhihu.com/p/39557326>

如上图所示，普通成像技术记录的只是粉红色的色彩空间sRGB的部分，而HDR成像技术则记录的是整个视觉的色彩空间（灰色的部分）。HDR成像与普通成像的区别是：HDR成像系统始终以高精度记录数据，而且与显示设备无关。只有在显示阶段，图像的色彩空间才因为显示的局限被压缩了。而普通成像则在图像采集，处理和存储阶段就已经压缩了色彩空间。

**HDR成像包括了以下技术：**

1. HDR图像的采集
2. HDR图像的编码和压缩；
3. tone mapping技术（HDR图像在普通显示器上显示需要的处理）；
4. 反tone mapping技术；（已经处理了的HDR图像在HDR显示器上显示需要的处理）
5. HDR显示技术；
6. HDR图像质量的评价标准

...

## Basic Theory And Correlation Techniques

### 动态范围测量

图像里往往是有很少的非常亮的和非常暗的像素，所以动态范围的测量一般是把图像先做low pass filter，再去测量，这样计算的结果比较可靠，因为经过低通滤波后的图像把最小亮度的像素求了平均会更好的估计出noise floor，而且平滑了特别亮的单点，这样测量的结果才比较稳定。low pass filter会过滤图片中高频部分，图片中高频部分是变化明显的部分，所以使用低通滤波器可以保留图片平滑的部分。

## 感知动态范围

<https://www.zhihu.com/question/21492596>

人的视觉动态范围大，主要实现机制是在**视觉系统各层级灵活的增益控制（gain control）**。人眼感知的动态范围是个比较有争议的话题。当光线散射到人眼表面，人眼可以有效地把最大亮度对比度减小到2-3个log10单位。人眼可以被看做一个非常灵敏的sensor，它可以迅速地改变注视点并且针对局部进行适应，所以人可以感知大于4个log10单位的动态范围。这种感知能力是与场景相关的，所以不能说一个绝对的数字来评价人眼的动态范围。不过，有研究证明，人能感知并且喜欢对比度高于1000:1的图像，这种人喜欢的图像的峰值亮度很高，像OLED显示器，它虽然动态范围很大，但是它的峰值亮度不高，它所呈现的内容大多在低亮度的范围，而这个范围人眼对色彩是不那么敏感的。所以这种显示器并不是最理想的。

另外，人眼中感光细胞也主要有两种，一种是视锥细胞，一种是视杆细胞。在中心凹处大约有七百万个视锥细胞。视锥细胞对 $10^{-2}$ — $10^8$  cd/m<sup>2</sup>（明亮的环境）等级的光线更为敏感，负责感知快速运动和高频信息。进一步细分，视锥细胞可以分为：短波视锥细胞，对435nm左右的光线最为敏感；中波视锥细胞，对530nm左右的光线最为敏感；长波视锥细胞，对580nm左右的光线最为敏感。

视杆细胞大约有七千五百万到一亿五千万个。视杆细胞主要对 $10^{-6}$ — $10$  cd/m<sup>2</sup>的光线（较暗的环境）比较敏感，但是视杆细胞无法提供颜色信息，这也就是为什么在低照度环境下我们难以区分物体的颜色。视锥细胞和视杆细胞会对原始光线动态范围进行压缩以缩小动态范围，压缩规则如公式（1）所示：

$$\frac{R}{R_{\max}} = \frac{I^n}{I^n + \sigma^n} \quad (1)$$

其中R代表感光细胞的响应， $R_{\max}$ 是最大的响应，I代表光线强度，变量 $\sigma$ 和n分别是半饱和常数和灵敏度控制指数，对于视锥细胞和视杆细胞是不同的。

## 动态范围的不同表示方法

**【动态范围】**主要是用来衡量受光者（如相机里的感光器件）与发光者（如显示器）的，并不用于已数字化的图片。因为无论感光器件有多高级，动态范围有多大，即使它接收的图像有足够大的亮度差（例如蓝天中的一片白云与近处岩石后面的阴影被摄入同一幅图像），只要它被数字化为8位图像，那么这幅图像仍然只有256级灰度。这也就是说，感光器件的质量水平不同（就是动态范围不同），它们提供的256级灰度反映的“实际的亮度差”是不一样的。从这一点而言，当然感光（发光）器件的动态范围越大，就越能更真实的记录（显示）自然界的亮度变化。下表是根据实际应用的不同有以下几种动态范围的测量：

name	formula	example	context
contrast ratio	$CR = (Y_{\text{peak}}/Y_{\text{noise}}) : 1$	500:1	displays
log exposure range	$D = \log_{10}(Y_{\text{peak}}) - \log_{10}(Y_{\text{noise}})$	2.7 orders	HDR imaging, photography
	$L = \log_2(Y_{\text{peak}}) - \log_2(Y_{\text{noise}})$	9 stops	
peak signal to noise ratio	$PSNR = 20 \cdot \log_{10}(Y_{\text{peak}}/Y_{\text{noise}})$	53 [dB]	digital cameras

**【对比度】**是一个相对值。就一幅图片而言，它反映了图片上最亮处与最黑处的比值。假设这幅图片上最亮与最黑之间的变化是均匀的，如果这个比值没有超过显示这幅图片的显示器的动态范围，那么我们看这幅图时会说它有充分的层次。如果这幅图片上最亮处与最黑处的比值超过了显示这幅图片的显示器的动态范围，那毫无疑问，偏亮的部分就已经达到了显示器的亮度极限，那更亮的部分也不会更亮，这部分的层次就会丢失。同理，图片中较暗的地方显示器已成全黑，那更暗的部分也无法显示。实际上，数字图像处理中的对比度拉伸是当一幅图像中的亮暗差较小时（例如只有50~150），我们比较均匀的扩展它的对比度，将它的幅值拉伸到10~180或5~240。这样，暗处显得更黑，亮处显得更亮些，由于是均匀拉伸，原本不易分辨的灰度差（中间灰度）也拉开了距离，显得图片更富有层次。

【对数域曝光范围】（log exposure range）是HDR技术中常用于测量场景动态范围的度量。这里主要考虑的是给定场景中的最亮和最暗亮度，亮度的计算是在对数域下进行的。使用对数值的优点是，它们能够更好地描述动态范围中感知的差异。

【信噪比】（SNR）用于描述数码相机的动态范围。它利用使图像传感器恰好饱和的信号强度 $Y_{peak}$ 和恰好使图像传感器能够感知的信号强度 $Y_{noise}$ 来进行定义。单位是dB。

## 最小可分辨差异（JND）

人眼对于光线的感知并不是线性的，另外，实际上，将现有的LDR图像处理算法适配到HDR图像的最简单的方法是对对数像素值进行操作。所以相对对比度变化的敏感度与对数函数的关系以及最小可分辨差异很重要。

公式(2)为亮度对比度的定义

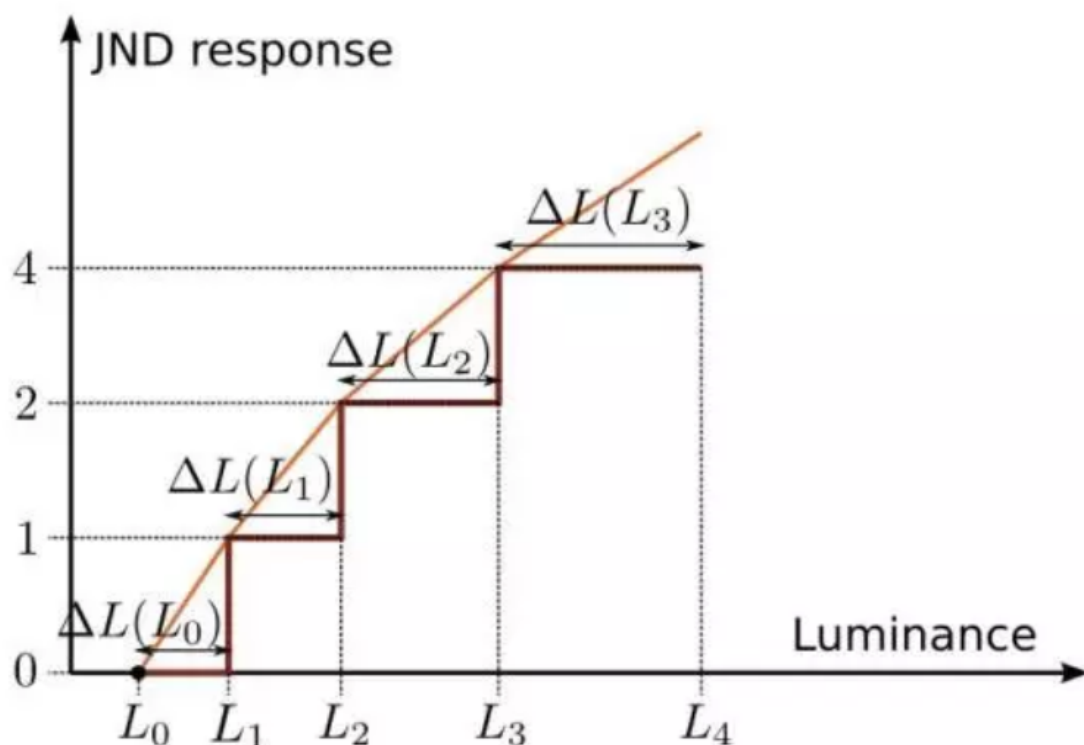
$$c = \frac{\Delta L}{L} \quad (2)$$

来源于韦伯定律；我们希望构建一个函数 $R(L)$ 用于近似描述人眼对光线的响应，假设可以恰好被人眼分辨的两个亮度等级之间差异是单位1，如公式(3)所示：

$$R(L + \Delta L) - R(L) = 1 \Leftrightarrow \frac{\Delta L}{L} = c \quad (3)$$

响应空间 $R$ 可以通过迭代的方式导出，可以计算得

$$L_t = L_0 * (c + 1)^{R_t} \quad (4)$$



亮度到JND尺度响应的映射函数即如上图。

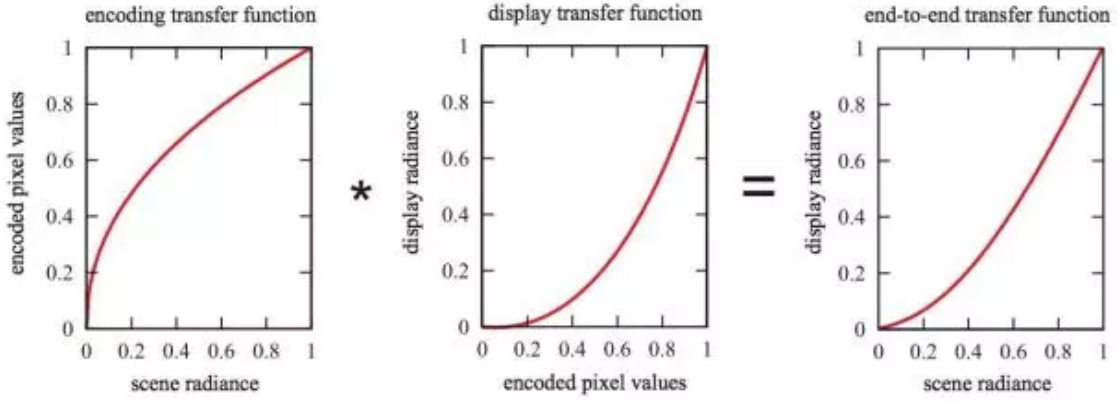
## 显示模型和gamma 校正

### gamma修正

大多数低动态范围的图像和视频格式都会使用伽马修正将亮度或RGB光谱颜色强度转换为整数，以方便编码。伽马修正通常以指数形式给出：

$$intensity = signal^\gamma \quad (5)$$

其中intensity表示原始的亮度或光谱强度，signal表示修正后的信号强度。 $\gamma$ 的取值在1.8到2.8之间。伽马修正一开始是为了减少相机噪声并且适应当时的CRT显示器的水平，但是之后发现，经过伽马修正后显示器所显示的信息能线性近似于真实场景的信息，这是一个令人惊喜的巧合。因为CRT显示器的输入电平和显示值之间也存在一个指数关系，而像素的编码值决定了输入电平值的大小。我们可以使用两个不同伽马值定义的转换函数(transfer function)来描述整个显示系统。编码伽马值(encoding gamma)定义了编码转换函数，用于描述设备采集的场景辐射值与像素编码值之间的关系。另一个是显示伽马值(display gamma)，定义了显示转换函数，用于描述输入的编码像素值与显示设备显示光谱值之间的关系。两个转换函数的乘积得到的全局伽马值(overall gamma)或端到端伽马值(end-to-end gamma)可以很好描述整个显示系统，如下图所示。



全局伽马值一般不设为单位一，根据不同的设备和使用条件不同而不一样，比如在影院中一般设为1.5，在办公室中一般设为1.125[6]，以此来保证显示辐射值与真实场景辐射值的近似关系。

## 显示模型

以上介绍的显示伽马修正，只是显示模型的简化，为了进一步定量分析，就需要介绍完整的显示模型，伽马偏移增益(gamma-offset-gain, GOG)。GOG模型更准确地描述了输入的像素值与显示的值之间的关系。式(6)给出了在灰度图情况下像素的亮度值与显示亮度之间的关系，

$$L = (L_{peak} - L_{black}) V^\gamma + L_{black} + L_{refl} \quad (6)$$

$L$ 是显示亮度， $V$ 是像素值亮度。 $L_{peak}$ 是显示设备在完全黑暗环境中的亮度峰值， $L_{black}$ 是显示设备所能发出的最暗光线的亮度， $L_{refl}$ 表示从显示设备表面反射的环境光值。

$$L_{refl} = \frac{k}{2\pi} E_{amb} \quad (7)$$

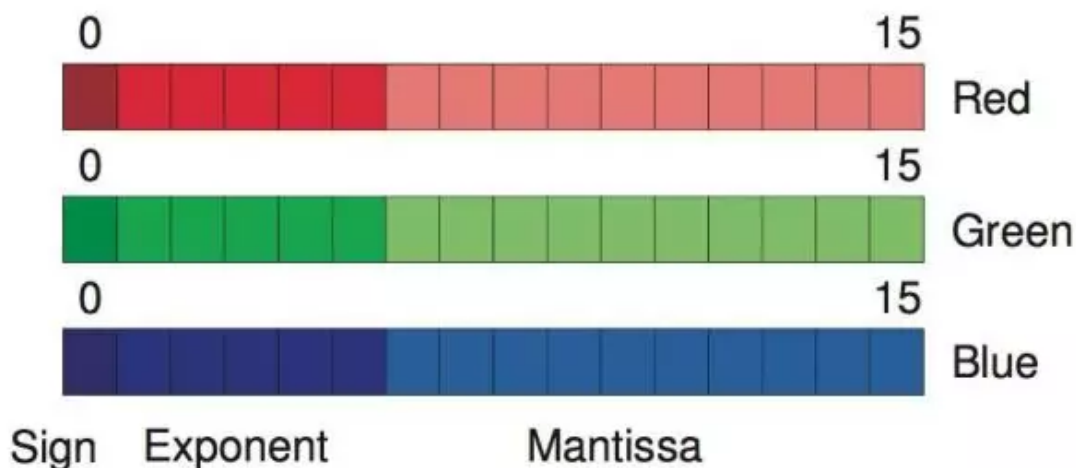
$k$ 表示显示设备的反射率， $E_{amb}$ 表示环境光的亮度。

## HDR存储格式及压缩

由于HDR图像本身携带更多信息，所占存储空间更大，比如相同分辨率的RAW格式图像是JPEG图像所占空间的60倍到250倍。所以我们对于HDR的音视频我们需要一种更好的表征方式。

## 编码格式

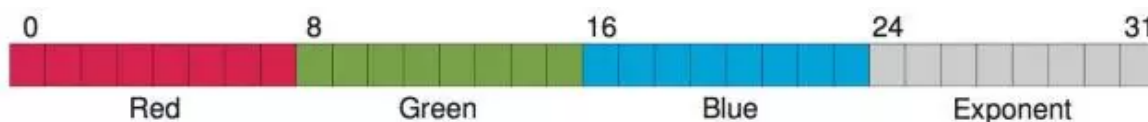




### 【半精度浮点数】

半精度浮点数可以灵活方便地对线性亮度值和辐射值进行编码，能覆盖较大的动态范围，同时比直接用32位精度浮点数更节省空间。所以HDR图像文件存储格式OPENEXR就采用的这种像素编码格式。但是由于精度的原因，这种表示方式最大只能表示65504，所以若HDR图像包含亮度值极大的部分则需要放缩后再用半精度浮点数存储编码。

### 【RGBE】



RGBE的像素编码格式主要用在Radiance的HDR图像文件格式中。利用四个字节编码，前三个字节表示红绿蓝三种颜色，最后一个字节存储公用的指数。之所以共用一个指数是因为三种颜色在RGB空间中有很强相关性，数值上至少都在同一数量级，因而不需要为每个通道单独存储指数。转换到三原色的公式如下：

$$(r, g, b) = \begin{cases} \frac{(R, G, B) + 0.5}{256} 2^{E-128} \frac{\text{exposure}}{E_w} & \text{if } E \neq 0 \\ (0, 0, 0) & \text{if } E = 0 \end{cases} \quad (8)$$

反变换：

$$E = \begin{cases} \lceil \log_2(\max\{r, g, b\}) + 128 \rceil & \text{if } (r, g, b) \neq 0 \\ 0 & \text{if } (r, g, b) = 0 \end{cases}$$

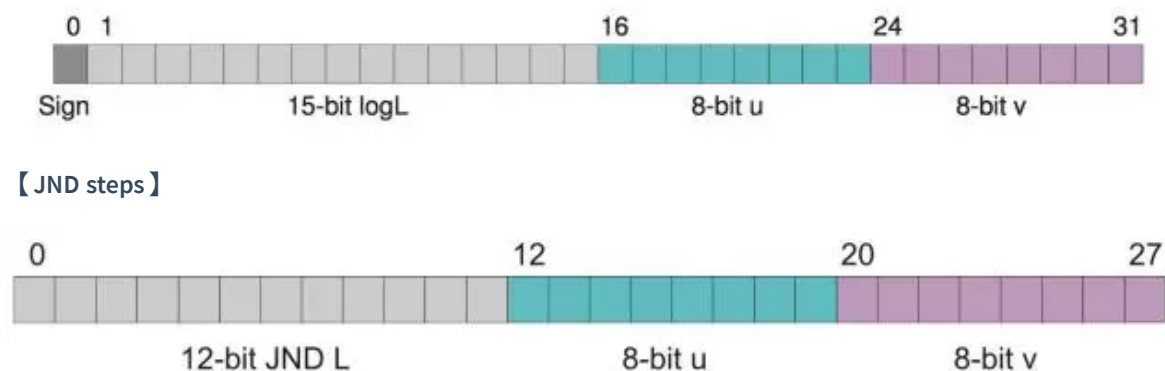
$$(R, G, B) = \left\lfloor \frac{256(r, g, b)}{2^{E-128}} \right\rfloor$$

因此它不能表示高度饱和的颜色，因为转换到RGB空间会变成负数。

### 【LogLuv】

我们知道人眼对不同亮度范围的亮度差异的感知是不一样的，所以结合JND用整数编码HDR图像更适合压缩。

LogLuv编码格式就是用整数对人眼可见的整个动态范围和色域进行编码。人眼对亮度变化的感知是非线性的，对暗处的亮度变化更为敏感，这种特性可以用对数形式近似模拟。LogLuv用32位二进制数进行编码，其中两个字节编码亮度，另外两个字节编码色度。但是这种格式也有缺点，它对于使用算术编码进行的压缩是无效的。具体形式如下图所示。



JND(最小可辨别差异)可以将真实亮度映射为一个JND单元，此单元需要更高的编码精度，因此而提出了JND STEP的编码格式。

## HDR 文件格式

### 【Radiance】

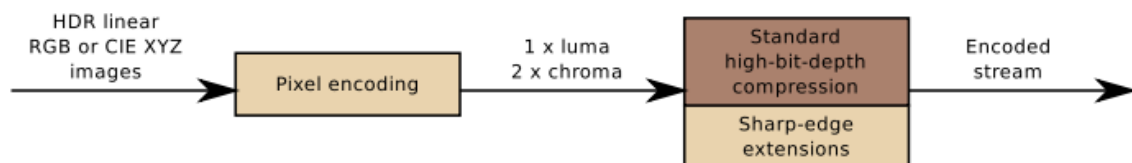
Radiance格式是最早的HDR文件格式之一，在1989年引入Radiance Rendering Package(<http://radsite.lbl.gov/radiance/>)之后更为流行。文件主要由文本头文件以及之后的编码像素构成。像素编码所采用的格式是4.1.2中介绍的RGBE或者XYZE编码格式，两种编码方式差异在于前者所用的基元是RGB(三原色)，后者用的是CIE 1931 XYZ中的XYZ。后者可以编码人眼可见的所有颜色。

### 【OpenEXR】

OpenEXR格式（或扩展范围格式），由文件名扩展名.EXR识别，该格式已被许多开源和商业应用所改编，并成为HDR图像实际上的标准，在特效行业应用广泛。主要特性：

- 支持16位、32位浮点数和32位整数像素。实际应用中16位半精度浮点数使用得最广泛。
- 支持多种无损和低损的压缩算法。
- 具有良好的扩展性。新的压缩编解码器和图像类型可以通过扩展OpenEXR软件发行版中包含的C++类轻松添加。

### 【High bit-depth encoding for HDR】



[The HDR pixels need to be encoded into one luma and two chroma channels to ensure good decorrelation of color channels and perceptual uniformity of the encoded values ]

HDR图片也可以存储在任何支持更高比特位深度的标准压缩格式中，按照上图流程进行压缩存储。一个难点在于对比度尖锐的边缘，因为几乎所有现代压缩算法都使用频率转换，边缘会导致频率系数很高。当这种系数用量子化表示，解码图片在边缘处经常出现**振铃**效应。这可以通过在每一个8x8的尖锐对比度边缘部分编码来缓解。见下方文献。

[Rafał Mantiuk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Perception-motivated high dynamic range video encoding. ACM Transactions on Graphics (Proc. SIGGRAPH), 23(3):730–738, 2004. ]

## 色调映射

色调映射是指将高对比度和宽色域的场景在对比度和色彩显示有限的显示器上进行渲染的过程。通常它包括将高动态范围的图像转换为可以在普通显示器上显示的像素值。

## 分类

### 【 Visual System Simulators(VSS) 】

VSS主要是模仿实验视觉系统的限制和性质，比如会加入眩光(glare)、模仿人眼在夜间的视觉效果或者减少图像色彩降低图像对比度，对图像在真实世界场景的条件和观看条件之间的差异（包括色彩方面）进行调整。

### 【 Scene reproduction(SRP) operators 】

SRP是尝试在色域、对比度和峰值亮度都受限的显示设备上保留原场景的样子，包括对比度，锐度和色彩。这类操作不会模拟由于感知影响而导致的外观变化，例如夜间视力和色觉的丧失。相反，他们致力于克服显示媒介的局限性，并尝试在有限的色域和动态范围内实现最佳匹配。

### 【 Best subjective quality(BSQ) operators 】

BSQ主要是基于主观喜好和艺术特点来产生最适宜的图像。这样的操作通常包含一系列的自适应参数，比较好的例子就是一般的照片编辑软件。

正因为有不同的优化目标，所以没有单一的最佳色调映射算法。

## 代数解释

简单而言，色调映射的代数解释就是将HDR的真实场景亮度转换到可以被LDR显示器现实的亮度区间。

For simplicity, we restrict our consideration to gray-scale images. The tone mapping function will be denoted by:

$$\hat{L}_p = T(L_p), \quad (16)$$

where  $L$  is the HDR pixel luminance,  $p$  is an index of a pixel, and  $\hat{L}$  is luminance that should be shown on a display. The values of  $\hat{L}$  should be transformed by a display model (refer to Sec. 2.2) to get LDR pixel values that could be sent to a display. This step is often confusingly called *gamma correction*, though it is meant to transform luminance into luma values, as discussed in Sec. 2.2 rather than correct any aspect of an image.

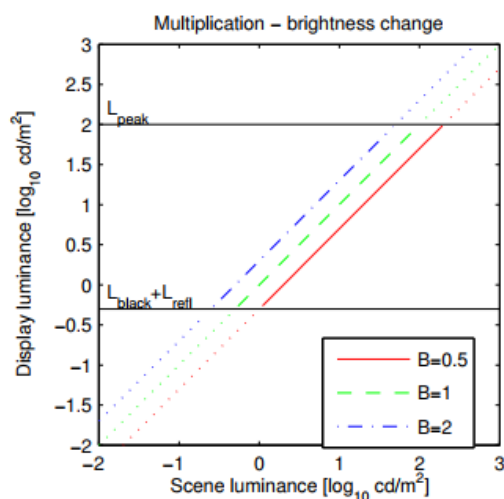
有一些色调映射的方法直接将HDR的像素值 $L_p$ 转换为LDR的像素值 $V_p$ ，而忽略了显示模型。这样做的缺点就是不能针对不同的显示器做出调整。

### 乘法操作——亮度变化

乘法操作是最简单的操作之一，但是对HDR图像做乘法操作只会改变图像的整体亮度而不会影响图像的对比度或者动态范围，所以这样的色调映射函数，类似于相机中的曝光变化，所以也称为曝光调整。

$$T(L_p) = B \cdot L_p \quad (9)$$





(a) Tone mapping function



(b) B=0.5



(c) B=1



(d) B=2

对数域上则如上图a:

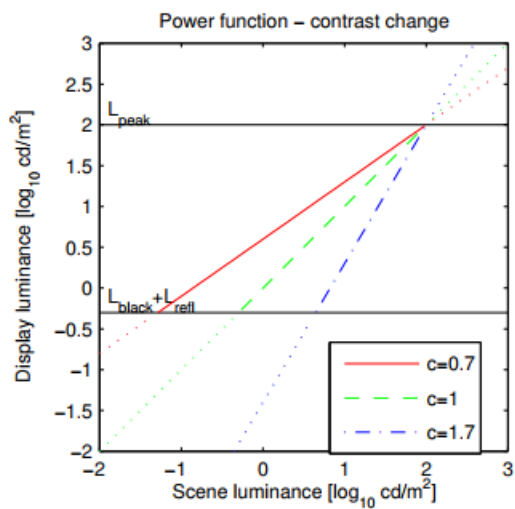
$$t(l_p) = l_p + b \quad (10)$$

### 功率函数(Power Function)——改变对比度

功率函数可以控制一张图片的动态范围，动态范围有时与图像对比度可互换使用，因为降低动态范围也会降低整体图像对比度。公式如下，

$$T(L_P) = \left( \frac{L_P}{L_{white}} \right)^c \quad (11)$$

其中c是对比度调节因子，该变化是相对于参考白点 $L_{white}$ 的亮度，使得对比度将朝向或远离该点收缩或扩大。 $L_{white}$ 一般是场景中最亮的点，会映射到显示器的峰值亮度。例子如下图所示。这种变换由于形式与显示模型类似，所以也被称为伽马修正。



(a) Tone mapping function



(b)  $c=0.7$



(c)  $c=1$



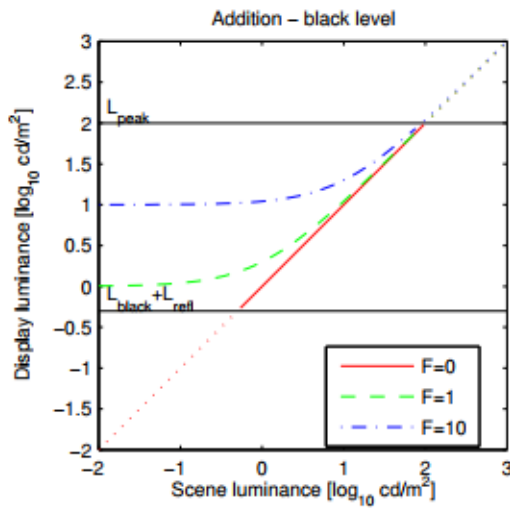
(d)  $c=1.7$

对数域上:

$$t(l_p) = c(l_p - l_{white}) \quad (12)$$

Addition — black level, flare, fog

由Weber定律，我们为图像加一个定值，则对暗处的影响远大于对亮处的影响，这会形成雾化效果或者光晕。如下图：



(a) Tone mapping function



(b) F=0



(c) F=1



(d) F=10

## 色调映射的主要方法

### 光照和反射的分离

如果我们把到达眼睛的光线看作是光照(illumination)和物体表面的反射的积，并且对于视觉系统而言反射提供了比光照更多的信息。反射可以提供物体的形状颜色和纹理特征，并且对于大部分情况下都是不变的。相反光照很容易受环境的影响，比如，室内和室外。事实上，有证据表明，人类视觉系统中的存在几种生理机制是为了减少光照的影响，颜色适应就是一个典型的例子。如果光照不那么重要，那么很可能对图像的光照分量的修改比对反射的修改更容易保证图像原有的性质。

漫反射表面反射率的变化范围很大，从黑色天鹅绒的1%到白色油漆的90%。即使在同一场景中两种物体都存在，单独反射产生的最大动态范围也要小于两个数量级，但是光照部分对动态范围的影响很容易就能超过四个数量级。这也就是说对图像中光照部分进行压缩将能极大程度的减小图像的动态范围。

$$intensity = reflectance \times illumination \quad (13)$$

若光照和反射知道如何分离，就可以得出简单的色调映射操作算符，只会影响光照而不会对反射部分造成失真。

$$intensity_d = reflectance \times T(illumination) \quad (14)$$

上式中， $intensity_d$ 表示经过色调映射后的HDR图像像素值，最早提出这种算法的是下面这篇文献：

[AV Oppenheim, R.W. Schafer, and T.G. Stockham. Nonlinear filtering of multiplied and convolved signals. Proceedings of the IEEE, 56(8):1264–1291, 1968. ]

【 Low-pass filter decomposition 】

In contrast to the reflectance, the illumination in a scene usually varies smoothly between pixels. The only sharp discontinuities can be expected at the boundaries of hard shadows and light sources. The easiest way to extract the smoothly changing part of an image is to convolve it with Gaussian filter of a large extent:

$$I_p \approx \sum_{t \in \Omega} f(p-t) L_t \quad (15)$$

$I_p$ 是像素位置 $p$ 的光照度分量， $L$ 为线性强度值（光亮度）， $\Omega$ 为像素局部区域， $f$ 为高斯函数。虽然比较粗糙，但在很多情况下效果满意。此方法的提出来自于如下文献：

[K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In Graphics Interface, pages 245–253. Citeseer, 1993. ]

他们提出一个压缩HDR图像对比度的空间上非归一化映射函数：

$$T(L_p) = \frac{L_p}{kL_p} \quad (16)$$

$k$ 是常数，取值于2-8。 $k$ 影响整体图片亮度，但不影响光照度组分的对比度。

高斯滤波分离也被用于流行的*unsharp masking algorithm*,提升图片细节。

### 【Bilateral filter decomposition】

高斯滤波器在光照分离操作中主要的限制是它不能检测锐利的光照变化，这会在锐利阴影的边界和光源的边界处出现。

As a result, the illumination component is smoothed out across these boundaries and halo artifacts start to appear in an image when strong contrast modification is applied （smooth out有点难理解，halo是光晕）



但是双边滤波器会检测锐利边界并大大减少不希望出现的光晕现象。

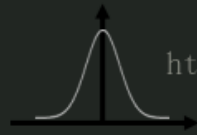
# Spatial Parameter



input

$$GB[I]_p = \sum_{q \in S} G_{\sigma}(\|p - q\|) I_q$$

size of the window

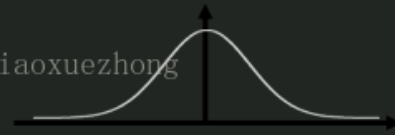


small  $\sigma$

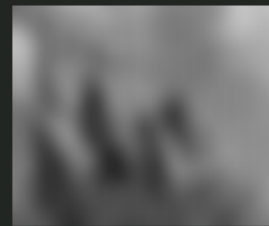


limited smoothing

<http://blog.csdn.net/piaoxuezhong>



large  $\sigma$



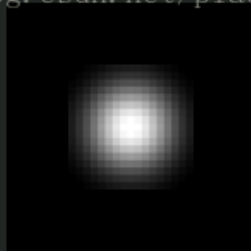
strong smoothing

$$BF[I]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|I_p - I_q|) I_q$$

new      not new      new

normalization factor      space weight      range weight

<http://blog.csdn.net/piaoxuezhong>



来源: <https://blog.csdn.net/piaoxuezhong/article/details/78302920>

$$I_p \approx \frac{1}{k_s} \sum_{t \in \Omega} f(p - t) g(L_p - L_t) L_t \quad (17)$$

$$k_s = \sum_{t \in \Omega} f(p - t) g(L_p - L_t) \quad (18)$$

$k_s$ 为正则化项。具体算法实现:Frédo Durand and Julie Dorsey. Fast bilateral filtering for the display of highdynamic-range images. ACM Transactions on Graphics (Proc. SIGGRAPH), 21(3):257-266, 2002. 在此之后也有很多优化。A. Adams, J. Baek, and M.A. Davis. Fast High-Dimensional Filtering Using the Permutohedral Lattice. Computer Graphics Forum, 29(2):753-762, 2010 这篇文章中有不同算法的比较分析。



除了上面介绍的滤波方法来分离光照分量和反射分量外，还有一种方法——Retinex算法。Retinex算法最初由Land和McCann提出，用来解释颜色恒常现象。Retinex将模拟HVS（人类视觉系统）在光照发生变化的情况下从我们感知的世界中提取可靠信息的能力。McCann将Retinex算法从理论变为形式化，并表明该问题等同于求解泊松方程。该算法本质上试图通过抑制小的梯度来分离反射和照明。

E.H.Land and J.J.McCann.Lightness and retinex theory.Journal of the Optical society of America, 61(1):1-11, 1971.

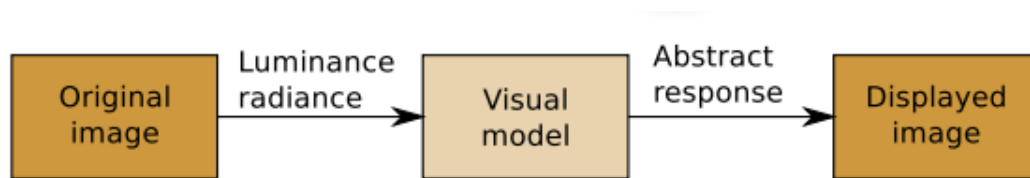
Berthold K.P. Horn. Determining lightness from an image. Computer Graphics and Image Processing, 3(4):277-299, 1974.

Anya Hurlbert. Formal connections between lightness algorithms. Journal of the Optical Society of America A, 3(10):1684, 1986.

还有基于梯度和对比度的方法，不用将图像分成反射层和照明层，可以在用伽玛函数（或对数域中的线性缩放）压缩图像对比度之前，增强图像中的细节（反射率）。这种方法的好处在于：对梯度操作会理性的增加局部对比度，而不是引入令人反感的对比度反转，比如光晕效应。不过局部操作也会引起全局图像亮度的不连续性在距离比较远的区域之间。因此更新的工作引入多规模图像结构。

Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edgepreserving decompositions for multi-scale tone and detail manipulation. ACM Transactions on Graphics (Proc. SIGGRAPH), 27(3):no. 67, 2008.

## Forward visual model



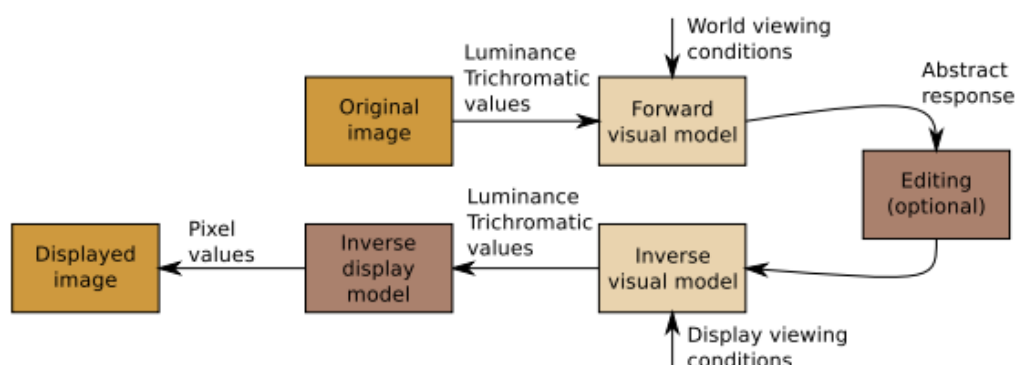
由于视网膜和视觉皮层之间的神经连接会只发送动态范围有限的信号，视觉系统在发送视觉信息到大脑之前需要实施一种有效的动态范围压缩。因此，一种色调映射方法是可以视觉系统中模拟这样的过程去缩小图片的动态范围。具体步骤就是物理信号被转换成一种视觉系统里抽象的内部表示，比如 lightness, brightness, photoreceptors，之后这样一种响应被映射到显示器的像素值中。如上图所示。

但是问题在于：

1. 人眼想看到的是亮度而不是抽象内部表示，因此，这种仅向前的色调映射方法可以认为是受感知启发的，而不是在感觉上合理的。
2. 而且很难了解这些操作的真实意图因为这个操作没有显示的形成一个原始和显示器感知上的匹配。

【Photoreceptor models】不是很了解原理

## Forward and inverse visual models



如流程所示，假设观察者适应原始场景中的观看条件，首先由正向视觉模型处理原始HDR图像。这意味着适应户外场景的较高亮度和夜景的较低亮度。然后，视觉模型的结果可以被进一步编辑调整，例如减小动态范围或提高细节的可见性。再下一步，抽象的表达通过基于特定的显示器的逆显示模型被转换回亮度值或三色值。最后，使用反向显示模型将物理亮度值或三色值转换为像素值。通过这样的处理就得到了伽马修正后的RGB值，可以直接显示在屏幕上。

这类方法包含前馈和反馈的模型，在物理上和感知上都是合理准确的。这种方法的一个主要优点是它可以根据真实世界场景和显示器之间的观看条件差异来调整图像外观。例如，夜景可以真实地在显示器上显示，显示的结果要比原始场景明亮。这是通过模拟视觉系统中的夜视（所谓的暗视觉和中视觉视觉）来实现的。前馈和反馈视觉模型还可以补偿由于色彩适应引起的色彩偏移，或者由于适应光照或黑暗环境而导致的视觉暂时丢失。

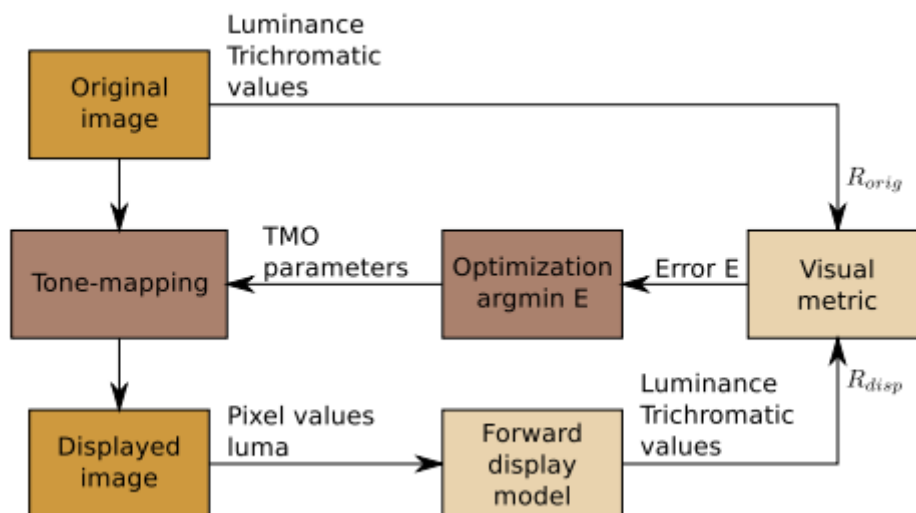
问题：

1. 它是建立在标准显示器可以再现更亮或更暗的场景的视觉感觉的前提下。根据所使用的视觉模型，逆视觉模型的结果可能产生位于色域之外的颜色，超出了目标显示器的动态范围，从而导致这样的操作在减小动态范围方面往往不是最有效的。
2. 许多复杂的视觉模型很难求逆，不能直接用于这种方法。

## 约束映射问题

Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images.

IEEE Computer Graphics and Applications, 13(6):42–48, 1993. 中提出的色调映射问题的原始目标之一是在显示器上再现一个场景，使得显示的图像的亮度感觉与现实世界的亮度感觉相等或接近匹配。由于显示设备能够提供的亮度，动态范围（对比度）和色域有所不足，所以原始图像和其在显示器上显示几乎不可能完美匹配。因此，输出设备上的渲染是以牺牲其他图像特征为代价的折衷。基于这些考虑，色调映射可以被表述为如下图所示的优化问题。



将原始HDR图像作为输入，目标是生成能最佳呈现原始场景的显示器适应的图像。可以这样假设，如果HVS对显示器上显示的图像的响应 $R_{disp}$ ，与对原始场景响应 $R_{orig}$ 近似相等，那么目标是可以达到的。由于显示器的性能限制两种响应是不可能一样的。同时考虑观看显示器显示时会受观看条件，如环境光，影响，是两者更难一致。所以这个优化问题的求解的方法就是优化色调映射的参数以最小化两者的误差，显示模型在上图中，引入了显示器的颜色和亮度方面的物理约束。

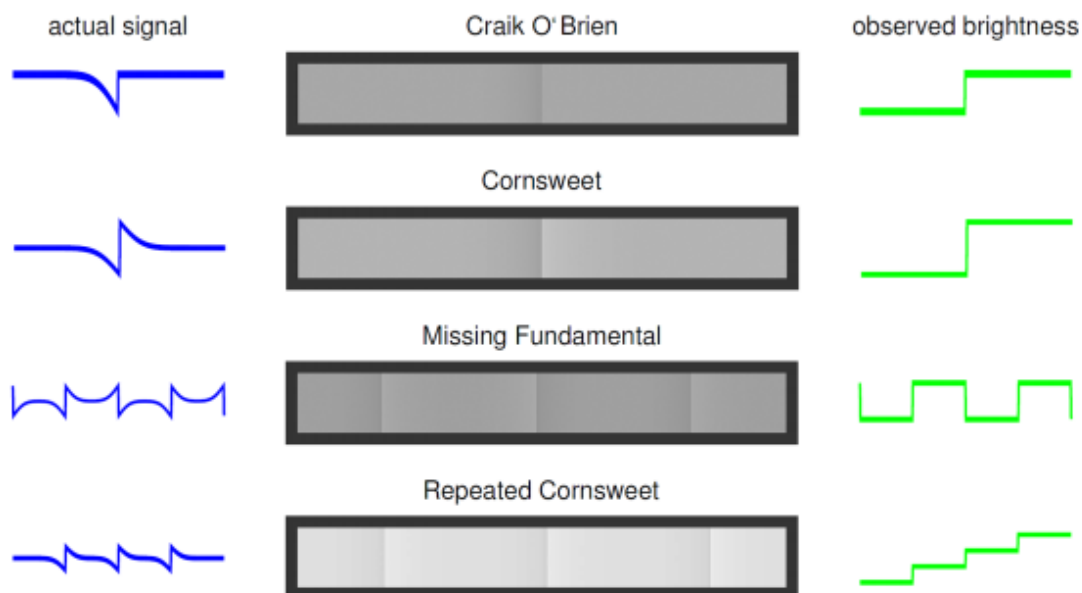
这种方法的主要困难在于以下事实：即使显示器，HVS和色调映射运算符的简化模型也会导致复杂的非线性优化问题，该问题可能表现出局部最小值，或者过于复杂而无法在合理的时间内解决。但是，熟练地解决问题后，可以非常有效地找到解决方案。

这种方法主要解决了场景再现（SRP）运算符（见上“色调映射的目标”）的意图，因为它试图匹配给定视觉指标的外观，而不是通过视觉模型来处理图像。

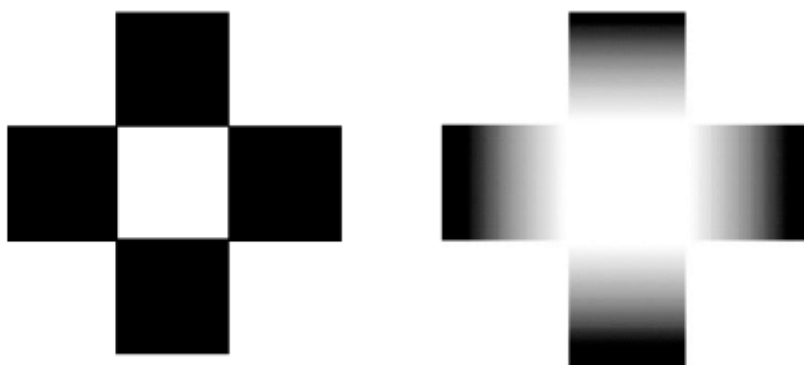
## Perceptual effects for the enhancement of tone-mapped images

(这些感知效果不大理解)

### 【Cornsweet illusion】



### 【Glare illusion】



## 逆色调映射

常用的8比特编码格式的图像若要在HDR显示器上显示则无法提供足够的精度，所以需要从LDR图像恢复HDR信息，这一过程称为逆色调映射(inverse tone mapping)，也被称为相机响应函数的反函数(inversion of complete camera response function)。通过相机响应函数的反函数可以对相机的光学缺陷和传感器的非线性进行补偿，所以逆色调映射的关键就是找到相机响应函数反函数。

在使用多曝光融合技术中，恢复相机响应函数已经有了较为成熟的流程。但仅有单张图片，没有任何拍摄相机信息和参数以及场景信息的情况下，找到合适的相机响应函数就变得十分困难，而这也是大量传统影视资源所面临的情况。所以想要将LDR转换为HDR，就需要逆色调映射技术(inverse tone mapping)。

由于在LDR图像编码存储时会有伽马修正，所以在恢复的第一步就是LDR图像像素值线性化，得到的像素值将是实际场景亮度值的等比例放缩。但也存在两个问题，一是由于有均匀量化过程存在，线性化的过程中会有伪影出现，尤其是在天空等平坦区域，我们称之为带状伪影(banding artifacts)；另外就是在图像高光区域、光源区域和阴影区域的细节恢复，这些细节在LDR图像中是被裁剪的，但在HDR图像中是可见的，所以在高光区域会产生诸如光晕等伪影。所以在恢复的过程中需要对伪影做修正，最后是颜色的修正，这一部分主要是宽色域(WCG)扩展的内容。

## 动态范围恢复

1. 获得逆伽马修正/色调映射函数或者逆相机响应函数
2. 使用逆函数转换所有的LDR像素获得线性像素值，这种像素值和原始场景中的亮度成比例
3. 将得到的像素值扩展到显示设备的最大动态范围能力上（有效扩展对比度）

### LDR像素值线性化

#### 逆伽马修正

Rempel和Masia等人提出简单的用 $\gamma=2.2$ 的逆伽马变换就能得到很好的线性化的像素值，这种变换直接基于伽马修正的反变换。Farid【Hany Farid. Blind inverse gamma correction. IEEE Transactions on Image Processing, pages 1428–1433, 2001.】在此基础上提出在缺少相机标定信息时用单张图片可以估计伽马值。他们的方法是基于这样的观察：伽马校正为图像引入了几个新的谐波，其频率与图像中的原始谐波相关。原始和新引入的谐波的幅度之间也有很强的依赖性。可以看出，频域中的这种更高阶的相关性随着伽马校正的非线性增加而单调地增加。可以使用来自多谱分析的工具来检测这种相关性，并且通过搜索使这种相关最小化的反伽马，可以找到最初应用于图像的实际伽马校正。这称之为盲逆色调映射 (blindly inverse tone mapping)。

#### 逆色调映射曲线

某种角度来讲，逆色调映射可看做色调映射的逆过程，基于此，Banterle等人通过色调映射函数的反函数来对LDR像素值线性化，取得了很好的视觉效果，但是动态范围扩展的幅度受到条带伪影的限制，特别是在明亮的图像区域中sigmoid函数大大压缩了对比度。

#### 逆相机响应曲线

实际上伽马函数只是相机响应函数的粗略近似，在恢复实际亮度时对精度有很大影响。Lin提出了对于单个LDR图像，可以基于物体边缘附近的彩色像素的分布来更精确地重建相机响应曲线。

### 动态范围扩展

#### 线性扩展

Akyuz和Rempel简单的使用线性拉伸就能获得很好的显示效果，Rempel指出将动态范围拉伸到5000:1是在HDR成像质量与可见伪影之间较好的折中。有意思的是，Akyuz在高质量的LDR图像中没有执行线性化的过程，仍然取得较好的HDR恢复图。

【Ahmet Ouz Akyüz, Erik Reinhard, Roland Fleming, Bernhard E. Riecke, and Heinrich H. Bühlhoff. Do HDR displays support LDR content? a psychophysical evaluation. ACM Transactions on Graphics (Proc. SIGGRAPH), 26(3), 2007.】

【Allan G. Rempel, Matthew Trentacoste, Helge Seetzen, H. David Young, Wolfgang Heidrich, Lorne Whitehead, and Greg Ward. Ldr2Hdr: On-the-fly reverse tone mapping of legacy video and photographs. ACM Transactions on Graphics (Proc. SIGGRAPH), 26(3), 2007. Article 39.】

#### 局部变化的亮度扩展(Locally-varying brightness boost)

显然，为了更精确地恢复HDR图像，对LDR图像中不同亮度的区域的扩展程度是不一样的，比如天空、光源相对于墙面就需要更多的扩展，高亮的区域扩展应该高于暗区域。

为了得到质量更好的HDR图像，需要对饱和区域的对比度和亮度同时提升。基于这个原因，饱和像素点周围的区域中至少一个像素通道的像素值需要有额外的增强，需要额外增强区域由一个高斯核对饱和像素点滤波得到。同时YongQing Huo【Yongqing Huo, Fan Yang and Chao Li, "HDR image generation from LDR image with highlight removal," 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Turin, 2015, pp. 1-5.】等人，认为高亮区域在并不占主要位置，并且常见的算法在高亮区域的恢复容易产生伪影，所以在不影响视觉效果的基础上可以通过预处理将高亮区域移除，简化算法流程。



Masia 【Belen Masia, Sandra Agustin, Roland W. Fleming, Olga Sorkine, and Diego Gutierrez. Evaluation of reverse tone mapping through varying exposure conditions. ACM Transactions on Graphics (Proc. SIGGRAPH Asia), 28(5):160:1–160:8, 2009. 】针对高亮区域基于伽马转换的思路解决。

### 基于语义的亮度扩展(Semantically-driven brightness boost)

一些其他方法通过图像中不同场景元素的语义差异，对其进行不同程度的扩展。例如Meylan 【Laurence Meylan, Scott Daly, and Sabine Susstrunk. The reproduction of specular highlights on high dynamic range displays. In Proc. of the 14th Color Imaging Conference, 2006. 】将图像分割为漫反射区域和高亮区域，对不同的区域采用不同的线性扩展系数，并且Meylan在一个心理学实验中观察到在户外场景中，被测试者倾向于将小部分的动态范围分配给高亮区域以获得明亮的整体场景，相反在室内场景中更倾向于向高亮区域分配更大的动态范围。

### 抑制轮廓和量化误差

常见的LDR图像多维8比特编码，在编码后由于量化会造成细节的遗失。有限的像素精度还会导致在平滑梯度区域出现假轮廓（带状伪影），对于彩色通道而言，这通常被称为分层。所有这些效应都会因为LDR到HDR转换的动态范围扩展而强烈加剧。传感器噪声也会变得更加明显。为此，在许多LDR到HDR技术中，在对比度和亮度提升步骤之前执行各种形式的高级滤波。

双边滤波和基于核的滤波是经常采用的滤波技术，因为可以比较好的消除带状伪影和传感器噪声带来的影响，但是滤波后会丢失低振幅和高频率部分的图像细节，比如人的皮肤纹理在滤波后可能会过度平滑导致视觉上不自然。

### 恢复过曝和欠曝区域纹理

传统LDR图像的另一个问题是图像欠曝和过曝区域，其中纹理图案大部分是饱和的，最多只包含场景中的稀疏信息。由于许多不同的场景可能导致在这样的区域中LDR图像的外观相同，这一点即使利用强大的机器学习技术也很难解决。通常认为，要获得较好结果必须要引入人的主观判断。Wang 【Lvdi Wang, Liyi Wei, Kun Zhou, Baining Guo, and Heung-Yeung Shum. High dynamic range image hallucination. In 18th Eurographics Symposium on Rendering, pages 321–326, 2007. 】等人提出可以通过曝光良好的区域作为参考来恢复过曝和欠曝区域的纹理信息。GABRIEL 【GABRIEL EILERTSEN, JOEL KRONANDER, GYORGY DENES, RAFAŁK. MANTIUK, JONAS UNGER, HDR image reconstruction from a single exposure using deep CNNs, ACM Transactions on Graphics, Vol. 36, No. 6, Article 178. Publication date: November 2017. 】等人提出用深度学习的方法，采用VGG网络对过曝区域的信息进行恢复，但对于欠曝区域并不能很好解决。

## HDR图像评价指标

大多数图像质量指标都考虑对某种设备进行质量评估，例如LCD显示屏或打印机。然而，通过计算机图形学方法计算出的物理精确的结果并不与任何具体的设备相关联。它们产生像素包含线性辐射值的图像，而不是显示设备的伽马校正的RGB值。此外，真实场景对应的亮度值可以跨越一个非常大的动态范围，超过了典型显示设备的对比度范围。因此，如何比较用于表示实际场景的这些图像的质量而不是它们经过色调映射(tone mapping)后的显示（即亮度经过调整后的显示）是需要解决的问题。

传统的LDR图像存储的像素值都是经过伽马修正的，在低亮度情况下（CRT显示器80 cd/m<sup>2</sup>）可以认为是感知均衡的，但是当显示亮度变亮时，暗处的失真将变得十分明显，现在的HDR显示器可显示亮度已经达到1000 cd/m<sup>2</sup>。而HDR图像存储的是线性的亮度值或辐射值，而物理上的亮度值之间的差异与实际我们感受到的视觉差异几乎没有关系，人眼对于亮度之间的对比更加敏感而不是绝对的亮度值，这一现象也被称为亮度掩蔽(luminance masking)。所以诸如PSNR之类的评价LDR图像的指标并不直接适用于HDR图像。

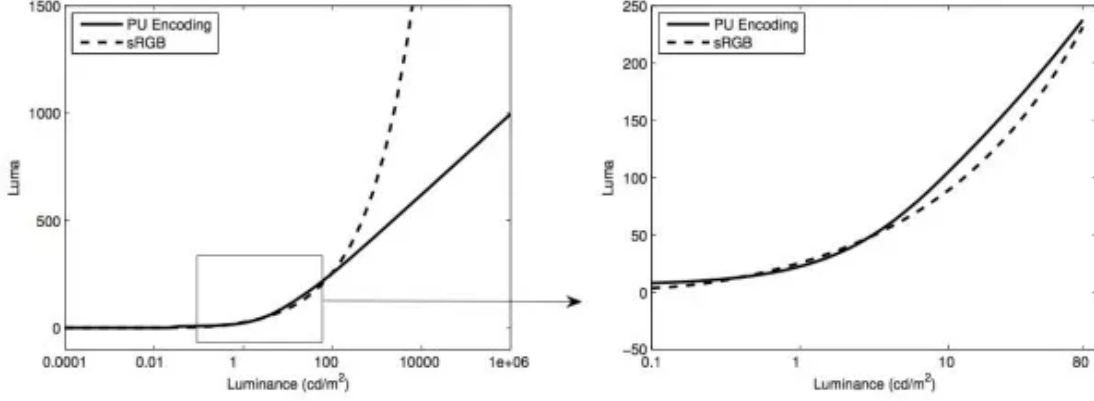
【对HDR图像的评价度量又可以分为对其**绝对亮度**和**相对亮度**的评价】

显示相关的度量是评价图像由显示器显示时显示器发出的绝对亮度，比如，感知均衡编码(perceptually uniform encoding)和HDR-VDP。

**感知均衡编码：**



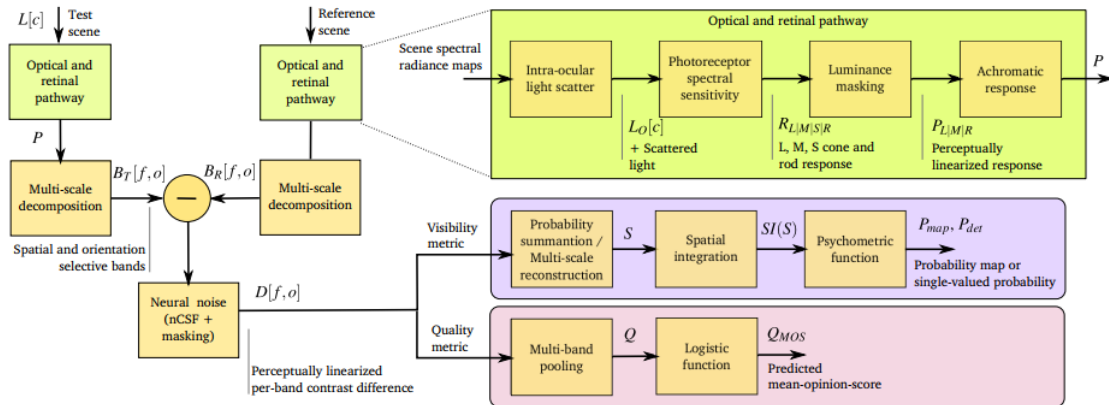
Aydin 【Tunç O. Aydın, Rafał Mantiuk, and Hans-Peter Seidel. Extending quality metrics to full luminance range images. In Proc. of Human Vision and Electronic Imaging, pages 68060B–10, 2008.】等人提出一种简单的亮度编码，从可以使用PSNR和SSIM【Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004.】等LDR的度量对HDR图像评价。通过将物理亮度值近似编码为感知均衡的表示形式，这种转换使用阈值积分法的亮度检测数据，之后又进一步受进一步约束将CRT显示器现实的亮度(0.1-80 cd/m<sup>2</sup>)映射到0-255的范围，以模仿sRGB的非线性，从下图可以看出，感知均衡的编码就是为了使映射后的亮度值与sRGB的相似，从而可以应用传统的度量方式。



## Visual difference predictor for HDR images

感知均匀编码的方式可以解决亮度掩蔽，但无法解决其他亮度相关的影响，比如眼内光的散射或亮度的对比度敏感函数(CSF)的频移。而HDR-VDP【Rafał Mantiuk, Scott Daly, Karol Myszkowski, and Hans-Peter Seidel. Predicting visible differences in high dynamic range images: model and its calibration. Proc. SPIE, 2005】很好的解决了这个问题，HDR-VDP是在传统的VDP基础上修改调整得到的，在HDR-VDP的基础上又调整得到HDR-VDP 2。HDR-VDP 2是对覆盖很广绝对亮度范围的无色图像之间差异的度量指标并可以将其可视化。（没有找到太多HDR-VDP的资料，现在上不了google了）

整个系统将HDR的亮度图或辐射图作为输入，预测其中可能存在差异部分的概率(Pmap和Pdet)以及感知失真水平(Q和QMOS)。一个限制高对比度场景的对比度感知的主要因素是由于光线在人眼的感光部分和视网膜上的散射。HDR-VDP 2将这一部分建模为一个频率空间的滤波器，将其与适当的数据集拟合。



## Luminance-independent Metrics

亮度独立的度量是指将对HDR图像像素值进行操作，并且对于像素值乘以常数后仍然能给出相同的结果。假设观察者对光线的敏感程度服从于韦伯定律，就可以将像素值转换到对数域以方便计算。一个典型的例子就是log-PSNR。

$$\log PSNR = 10 \cdot \log_{10} \frac{\log_{10} L_{\max}}{MSE} \quad (19)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \left[ \log_{10} \left( \widehat{L}_t(i) - 10g_{10} \widehat{L}_r(i) \right) \right] \quad (20)$$

$$\hat{L}_t(i) = \max(L_t(i), L_{\min}) \quad \text{and} \quad \hat{L}_r(i) = \max(L_r(i), L_{\min}) \quad (21)$$

$L_t(i)$  is the luminance of the pixel  $i$  in the test image, and  $L_r(i, c)$  is its counterpart in the reference image.  $L_{\min}$  is the minimum luminance considered to be above the noise level. Without such clamping of the lowest values, the metric introduces very large error for dark and noisy pixels.  $N$  is the total number of pixels in an image, and  $L_{\max}$  is an arbitrary selected peak luminance value. The typical selection of  $L_{\max}$  is 10 000, as few HDR displays exceed this peak luminance level. The value of  $L_{\max}$  must not be selected as the maximum pixel value in an image, as that would make such a metric image-dependent.

### Tone-mapping metrics

色调映射会固有地产生与原始高动态范围参考不同的图像。为了使所得图像适合显示器的可用色域和动态范围，色调映射通常需要压缩对比度并调整亮度。与在高动态范围显示器上看到的原始图像相比，色调映射的图像可能会失去一些质量，但是图像看起来通常非常相似，并且大多数质量指标很难预测质量的下降。史密斯等【K Smith, G Krawczyk, and K Myszkowski. Beyond tone mapping: Enhanced depiction of tone mapped HDR images. Computer Graphics Forum, 25(3):427–438, 2006】提出了第一个度量指标，用于预测由于色调映射引入的局部和全局对比度失真而导致的质量损失。