

论文题目 基于视觉适应机制的高动态图像增强算法研究

专业学位类别 工 程 硕 士

学 号 201322090506

作 者 姓 名 普璇

指 导 教 师 李永杰 教授

分类号_____密级_____

UDC^{注1}_____

学 位 论 文

基于视觉适应机制的高动态图像增强算法研究

(题名和副题名)

普璇

(作者姓名)

指导教师

李永杰

教 授

电子科技大学

成 都

申请学位级别 硕士 专业学位类别 工 程 硕 士

工程领域名称 生物医学工程

提交论文日期 2016.04 论文答辩日期 2016.05

学位授予单位和日期 电子科技大学 2016 年 6 月

答辩委员会主席_____

评阅人_____

注 1：注明《国际十进分类法 UDC》的类号。

Research of High Dynamic Range Image Enhancement Based on Visual Adaptation Mechanism

A Master Thesis Submitted to
University of Electronic Science and Technology of China

Major: **Master of Engineering**

Author: **Xuan Pu**

Supervisor: **Yongjie Li**

School: **School of Life Science and Technology**

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

作者签名：_____ 日期：____年__月__日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

作者签名：_____ 导师签名：_____

日期：____年__月__日

摘 要

随着人们对图像质量的要求不断提高，高动态范围图像（High Dynamic Range Image, HDRI）因其涵盖的场景信息更加丰富，获取成本又日益降低，所以被广泛应用。由于 HDR 图像拥有极大的动态范围，导致其在传统的显示设备上显示效果受到限制，无法完整的展现所包含的数据信息。由此，对 HDR 图像的动态范围进行压缩，并尽可能保留原有的图像信息，使其匹配低动态范围显示器是非常有必要的。人眼可感知的亮度范围十分巨大，从晴天中午的太阳到夜晚的星星都在人眼的感知范围内。本文基于人眼的生理学知识，视网膜的明暗适应机制以及已有的心理物理学实验结果，提出了两个基于人眼视觉适应机制的 HDR 图像动态范围压缩算法。两个算法均是静态的，自适应的局部算法。

算法一主要是利用了 Tumblin 提出的基于视觉特性的色调映射框架，将 HDR 图像映射前场景输入端和映射后显示器输出端引起的人眼响应进行等值匹配，运用一个正向模型和一个逆向模型来计算映射后的图像。我们的算法利用了感光色素的漂白机制来计算感光细胞的响应值，同时还在原有算法框架的基础上引入了色适应变换。

算法二中我们创新性地提出了一个高动态图像增强算法框架，旨在通过模拟真实的视觉适应机制来压缩图像的动态范围。我们参考了人眼在亮度适应过程中的生理机制，通过搭建感光细胞的亮度模型、感光细胞的响应模型、神经节细胞的感受野模型、视杆视锥的信息融合模型以及颜色变换模型来模拟视觉的适应机制，并以此来压缩 HDR 图像的动态范围。

结果表明，我们的两个算法都能够有效的压缩 HDR 图像的动态范围，无论是过亮区域还是过暗区域都能够在处理后很好的显示出信息，并且细节清晰，无色偏和伪影出现。

关键词：图像增强，视觉系统，明暗适应，高动态范围图像，色调映射

ABSTRACT

With our increasing demand of the Image quality, High Dynamic Range Image (High Dynamic Range Image, HDRI) is being widely used due to its richer cover scene information and lower cost. However, the dynamic ranges of most current output devices are too restricted to display the whole information HDR images contain. As a result, it is necessary to compress the dynamic ranges of HDR images as much as possible without losing the original information in order to match traditional displayers. It is known that human visual system is able to perceive a hug range of intensities, from the midday sun to the stars at night. Based on human visual physiology, retinal light and dark adaptation system and existing psychophysics experiments, this dissertation proposes two algorithms to compress the dynamic range of HDR images, both of which are static, local and almost automatic.

The first model is based on an existing framework proposed by Tumblin, in which a forward model and a reverse model could be used to compute the output image in that input image(original scene) and output image(display image)have perceptual match. In our model, the mechanism of photosensitive pigment bleaching is used to compute the response of photoreceptors and a chromatic adaptation model is added up to Tumblin's framework.

We present an innovative arithmetic framework to compress the range of HDR image in our second model. In this model we pay more attention to simulating the physiological process of eyes adaptation. And the major part of our model can be divided into three components, which include photoreceptors luminous model, photoreceptors response model and ganglion cells function. Besides, synthesis of rods and cones response signals as well as color stable are also taken into considerable consideration.

Keywords: image enhancement, human visual system, light and dark adaptation, high dynamic range image, chromatic adaptation

目 录

| | |
|---|----|
| 第一章 绪论 | 1 |
| 1.1 研究背景及意义 | 1 |
| 1.2 相关研究现状 | 1 |
| 1.3 本文研究工作及结构安排 | 3 |
| 第二章 视觉系统的生理基础及理论模型 | 5 |
| 2.1 视觉系统的生理机制 | 5 |
| 2.1.1 视网膜机制 | 5 |
| 2.1.2 视网膜神经节细胞的感受野机制 | 7 |
| 2.2 视觉感知机制 | 8 |
| 2.2.1 视敏锐度 | 8 |
| 2.2.2 视觉的暗适应与明适应机制 | 9 |
| 2.2.3 感光细胞的亮度响应模型 | 11 |
| 2.3 色适应变换 | 13 |
| 2.3.1 色彩学的基础知识 | 13 |
| 2.3.2 色适应变换 | 15 |
| 2.4 本章小结 | 16 |
| 第三章 基于 Tumblin 色调映射框架的高动态图像增强算法 | 18 |
| 3.1 基于视觉特性的 Tumblin 色调映射框架原理及相关研究 | 18 |
| 3.1.1 Tumblin 色调映射框架 | 18 |
| 3.1.2 相关算法 | 19 |
| 3.2 基于 Tumblin 色调映射框架的高动态图像增强算法 | 20 |
| 3.2.1 算法框架 | 20 |
| 3.2.2 算法流程详细介绍 | 21 |
| 3.2.3 颜色适应变换 | 24 |
| 3.2.4 实验结果 | 25 |
| 3.3 本章小结 | 28 |
| 第四章 模拟视觉适应机制的高动态图像增强算法 | 29 |
| 4.1 算法框架 | 29 |
| 4.2 算法流程详细介绍 | 30 |
| 4.2.1 感光细胞亮度计算 | 30 |

| | |
|---------------------------|-----------|
| 4.2.2 感光细胞的亮度响应模型 | 31 |
| 4.2.3 神经节细胞的感受野作用 | 32 |
| 4.2.4 视杆视锥信息融合及颜色变换 | 32 |
| 4.3 结果分析与展示 | 34 |
| 4.3.1 算法过程的中间结果 | 36 |
| 4.3.2 参数的影响 | 36 |
| 4.3.3 结果对比 | 37 |
| 4.4 本章小结 | 45 |
| 第五章 总结与展望 | 47 |
| 5.1 总结 | 47 |
| 5.2 展望 | 48 |
| 致 谢 | 49 |
| 参考文献 | 50 |

第一章 绪论

1.1 研究背景及意义

视觉是人类感知外界信息的主要通道，而图像作为一种有效的信息载体，是人类获取和交换信息的重要来源。因此，图像处理技术在人们的生活中和工作都发挥着重大的作用。图像增强一般处于图像处理的预处理阶段，主要作用是改善图像的质量，有针对性地丰富图像信息，增强可识别性等，以满足后续的处理。

图像动态范围一般指图像亮度的最大值与最小值的比值^[1]。高动态范围图像（High Dynamic Range Image, HDRI）是相对于传统的图像采集和显示设备而言的，HDR 图像的亮度范围跨度更大，涵盖场景信息更丰富。高动态范围图像（后简称高动态图像）能够相对完整的记录真实的场景信息，满足了人们对图像信息和质量日益提升的要求。随着摄像设备的升级和计算机成像算法的进步，我们已经能够很好的获取高动态图像，可以记录亮度在 10^{-6} - 10^8cd/m^2 的场景。然而目前主要使用的显示设备（如 LCD, LED, CRT）的动态范围是非常有限的，通常仅支持 1-2 个数量级，而投影和打印能呈现的动态范围更低。因此，在低动态范围（Low Dynamic Range, LDR）的设备上显示高动态图像，往往会导致严重的图像信息丢失，无法真实完整的再现场景，例如：原本较暗的区域会被压缩成一片漆黑，原本有层次的区域可能会被压缩成同样的亮度，由此造成信息丢失等。

为了能够在低动态范围的显示器上更好的展示高动态图像，我们可以将图像的动态范围进行压缩，并尽可能保留原有的图像信息，而这种为了更好的在低动态范围设备上显示高动态图像的动态范围压缩方法又被称为色调映射（tone mapping）或色调重建（tone reproduction）。本文旨在解决高动态图像的动态范围压缩问题，并由此提出了两个基于视觉生理特性的压缩算法，使得通过算法处理后的图像更贴近人眼直接观看真实场景的感觉。

1.2 相关研究现状

关于色调映射领域的研究工作已经进行了很多年，也有很多的算法发表出来。这里面的大多数算法都是针对图像的亮度通道进行处理，或者是对于图像的 R、G、B 三个通道做相同的处理。根据其处理方式的不同，现有的方法可以简单分为两大类：全局算子，又称为空域不变算法（spatially uniform），例如文献[2-7]都是采用的这种算法。全局算法将图像看做一个整体，对于图像中的每一个像素

都采用相同的变换。第二种是局部算法，又称空域变化算法（spatially varying）。局部算法会参考中心像素及其周围像素的信息，对不同位置的像素采用不同的变换。除了这两种分类以外，色调映射也可以根据其是否引入了时间变量分为静态和动态两种算法。这里我们主要介绍一下全局和局部两种算法的发展历程和研究近况。

全局算法的优势在于其简单易实现，消耗的电脑计算资源少，处理速度快，能够进行实时处理，但同时也存在局部对比度丢失和难以保留细节等问题。1993年 Tumblin 等人^[6]正式提出了针对高动态图像的色调映射问题，并基于 Stevens 等人^[8]的心理物理学数据，实现了利用视觉模型来处理高动态压缩问题，其对场景端和显示端分别采用了两个不同的全局变换。但是该算法只能处理灰度图，同时为了保留亮度损失了一定的对比度^[3]。

最小可觉差（Just Noticeable Difference, JND）是能够引起人觉察到差异感的最小刺激量，在不同的环境中人的最小可觉差是不同的。由此，1996 年 Ferwerda^[9]等人建立了一个基于 TVI（Threshold vs. Intensity, TVI）函数的全局算法。该算法计算出场景观察者在对应的适应状态下的可见阈值，以及估算显示设备观察者适应状态下的可见阈值，将其比值作为缩放因子。将原图按此比例进行缩放，即可得到变换后的图像，计算如公式（1-1）所示：

$$\begin{aligned} L_{da} &= m \times L_{wa} \\ m &= t_{da} / t_{wa} \end{aligned} \quad (1-1)$$

其中 L_{da} 是显示设备观察者的适应亮度， L_{wa} 是输入场景观察者的适应亮度， m 为缩放因子， t_{da} 和 t_{wa} 分别为对应适应亮度 L_{da} 和 L_{wa} 下的可见阈值。和同样基于 TVI 函数的 Ward^[10]算法不同，Ferwerda 不仅考虑了明视情况还加入了暗视条件，但是由于其仍是一个线性模型，能够成功映射到 LDR 设备上的最大动态范围仍然是有限的。

2000 年，Pattanaik^[11]等人提出了一种时变的色调映射算法，该算法在对亮度通道进行动态范围压缩时，利用了视杆和视锥的亮度响应模型，并模拟了随着时间变化的视觉适应过程。该算法的整体框架采用的是 Tumblin 在 1993 年提出的算法框架，并增加了视觉适应模型（Adaptation Model）和视觉展示模型（Appearance Model）来表现视网膜响应，亮度以及颜色。由于该算法是一个全局算法，所以在处理速度上可以达到实时处理，但是在效果上仍然存在对比度低，细节不够清晰等缺陷。

局部算法由于考虑了周边像素的影响，拥有更好的对比度，能够更好的保留细

节，但是在亮度变化较大的边缘容易出现伪影等问题。例如，Pattanaik^[12]等人在 1998 年提出了一个视觉系统适应和空间视觉的综合模型，其对图像中不同的频段运用了不同的增益控制因子，最终导致结果有严重的光晕现象。同时，文献[13, 14]也存在相同的问题。

为解决光晕问题，有不少的研究人员为此努力，其中包括文献[15-18]的工作。Durand^[19]等人于 2002 年提出了一种双边滤波算法，并于 2006 年对其进行了更新改进^[20]。该算法并没有利用人眼的生理特性，它将图像分为了基础层（base layer）和细节层（detail layer），亮度动态范围的压缩主要是在基础层进行，而细节层则很好的保留了图像的细节信息，处理的结果在一定程度上避免了光晕的出现。在本文的第四章中将会展示 Durand 算法的效果图，并与我们的结果进行对比。

基于 Pattanaik^[11]的工作，Ledda^[21]等人于 2004 年提出了一个模拟视网膜明暗适应过程的局部算法。在计算每个像素的适应状态时考虑了周边像素的信息，并采用双边滤波来计算，减少了边界处光晕的出现。但是该算法未考虑处理前后颜色适应的问题。

Meylan^[22]等人于 2006 年提出了一个基于中心外周 Retinex 的高动态图像压缩算法。该算法将通过主成分分析（Principal Component analysis, PCA）将得到的第一主成分作为图像的亮度通道，将第二第三主成分作为色彩通道，动态范围的压缩仅针对亮度通道进行。该算法还有一个创新点，就是采用了一个自适应滤波器（adaptive filter），该滤波器的尺寸可以根据图像的高对比度边缘信息来自动调整，由此也可减少光晕伪迹的出现。但是 Meylan 的算法存在严重的色偏问题，本文将在第四章对其结果进行详细的分析，并与我们的结果进行对比。

1.3 本文研究工作及结构安排

本文的工作旨在利用人眼视觉系统的生理机制来解决高动态图像的动态范围压缩问题。本文尝试采用了两种算法框架：第一种是 Tumblin 于 1993 年提出的基于视觉特性的色调映射框架，也是大多数研究人员经常使用的框架之一，例如 Pattanaik^[11]，Ledda^[21]等。本文在他人已有的工作基础上进行了算法改进，并加入了颜色适应模型等。第二种算法框架式是由我们首次创新性提出，通过学习视觉系统的适应机制，尝试建立并利用视锥视杆细胞的感光模型，对亮度的响应模型以及神经节细胞的感受野模型来更好的模拟视觉适应的生理过程，并利用此模型对图像动态范围进行压缩。

本文的内容结构安排如下：

第一章：介绍了此次研究对象高动态图在应用时存在的问题以及研究的意义。

然后通过对色调映射算法中全局和局部两大类算法已有的一些成果进行举例，介绍了色调映射研究领域研究历程和现状。最后，简要介绍了本文的研究工作以及结构安排。

第二章：介绍了关于视觉系统的生理基础，主要围绕视网膜的生理构造和相关特性。然后介绍了一些经典的视觉感知理论和模型，最后是关于色适应变换的理论原理介绍。

第三章：我们基于已有的研究以及 Tumblin 的色调映射框架，提出了一种改进的高动态图像亮度压缩算法。本章首先介绍了 Tumblin 色调映射框架的原理及发展现状，然后重点介绍了我们算法的框架及具体内容，最后分析展示了算法的结果。

第四章：本章主要介绍了我们全新提出的一种高动态图像压缩算法，该算法旨在通过模拟人眼视觉的适应机制来压缩图像的动态范围。本章首先简单介绍了我们提出的算法框架，然后对于算法中每一步的具体计算模型做了详细介绍，最后对结果进行了展示和分析，并与已有的经典算法做了比较。

第五章：最后一章对我们提出的两个高动态图像压缩算法做了总结，并分析了算法还存在的不足和今后对其进行改进和完善的一些想法。

第二章 视觉系统的生理基础及理论模型

人们在获取信息时，最为重要和基础的感知莫过于视觉，据统计结果表明，有80%~90%的外界信息都是通过人眼传入大脑中。同时，视觉也是非常复杂精密的系统，它的每一个组成部分都是非常重要的，而其中尤为关键的就是视网膜。本章首先介绍一些关于视觉系统的生理基础，主要围绕视网膜的生理构造和相关特性。其次将介绍一些经典的视觉感知理论和模型，最后是关于色适应变换的理论原理介绍。

2.1 视觉系统的生理机制

2.1.1 视网膜机制

外界的信息通过眼睛的光学系统在眼底形成外界物体的物像，视网膜将形成的物像的光能转换成并加工为神经冲动，经由视神经将冲动传入视觉中枢，从而产生视觉^[23]。图2-1展示了成年人眼球的水平切面图。

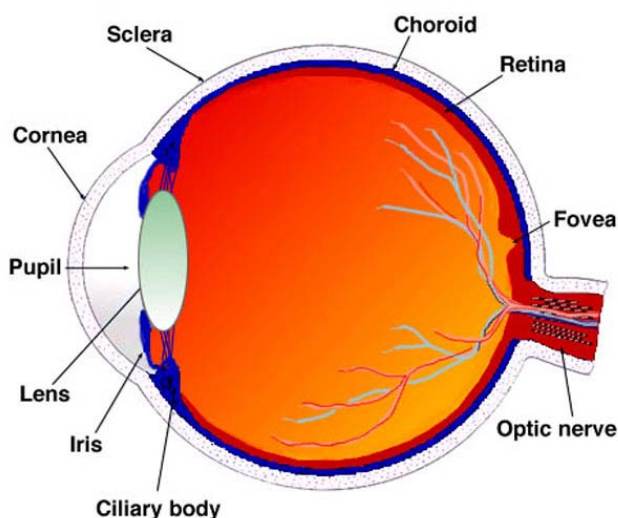


图 2-1 成年人眼球的水平截面图^[24]

视网膜由于和脑在胚胎发育时都是起源于外胚层，且与脑在形态和结构上相似，都具有十分规则的细胞分层结构。同时，视网膜因其能够处理复杂的视觉信息又被称为“外周脑”。视网膜是视觉系统中与感光机制以及视觉处理机制最密切组织。图 2-2^[24]为视网膜的结构图，其主要由 5 种细胞组成，它们被 2 层突触分隔为 3 层结构。最外层是感受器细胞层（Receptor Cell, RC），吸收光能并通过光转

导将其转换为神经电信号。中间层为双极细胞层（Bipolar Cell, BC），接受来自 RC 层传来的信息，处理后经化学突触传递到最内层的神经节细胞层（Ganglion Cell, GC），神经节细胞的轴突形成视神经。这三层细胞中的每一层都包含不止一类细胞，其中还包括了水平细胞，无足细胞，网间细胞等，每层内的细胞又相互联系，形成广泛的联络网^[25]。

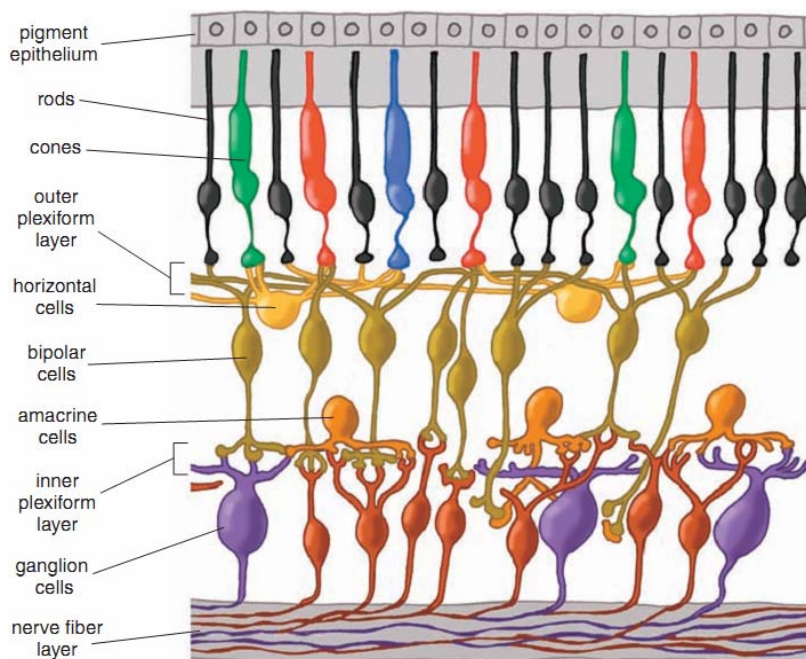
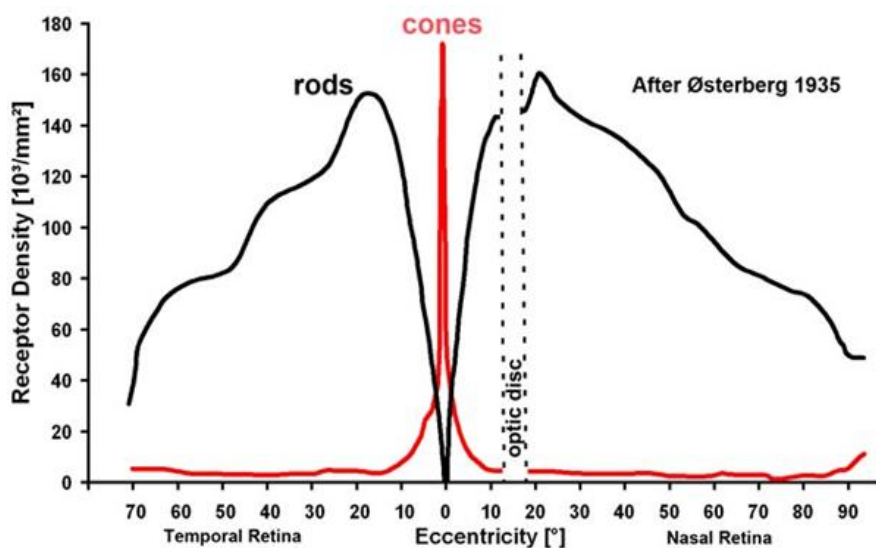


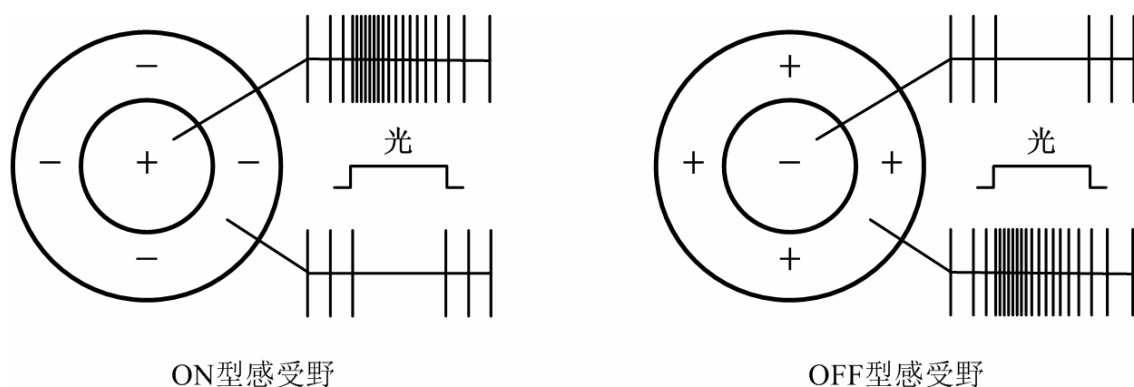
图 2-2 视网膜各层结构示意图^[24]

在视网膜的感受器层有两种感光细胞，分别是视杆细胞（rod cell）和视锥细胞（cone cell），有了它们人眼能够感知的动态范围超过了 10 个亮度等级（ 10^{10} ）。两种细胞在视网膜上的分布有很大的差异，具体可见图 2-3^[26]。视杆细胞主要分布于视网膜的外周区，它们对亮度的响应阈值低，对光线十分敏感，因此能够在黑暗和昏暗（ $10^{-6} \sim 10 \text{ cd/m}^2$ ）的环境下工作，被称为暗视（scotopic）。但是视杆细胞不能分辨颜色信息，且视敏锐度较低。视锥细胞主要分布于中央区，由中央凹向外延伸视锥细胞的密度迅速下降。视锥细胞的亮度敏感度低，主要工作于较亮的环境（ $10^{-1} \sim 10^8 \text{ cd/m}^2$ ），又被称为明视（photopic）。视锥细胞有较高的视敏锐度，能够分辨颜色和细节。当环境的亮度介于 $10^{-1} \sim 10 \text{ cd/m}^2$ 之间时，视锥视杆细胞同时处于工作状态，我们称之为间视（mesopic）。

图 2-3 视网膜上感光细胞在不同视角上的分布情况^[26]

2.1.2 视网膜神经节细胞的感受野机制

1938年Hartline^[27]在蛙单根视神经纤维上记录到了电反应并首次提出了感受野的概念。用一定的时间和空间构型的光去刺激视网膜上的某一区域，从而影响视觉系统中的单细胞的活动，则这个区域就称为该细胞的感受野（receptive field）。Hartline指出蛙的GC的反应可以分为三类：给光型（on型），即GC细胞在给光或光强增加时电位发放频率增加。撤光型（off型），即GC在撤光或光强骤降时细胞发放频率增加，而给光时无反应或频率减小。给光-撤光型（on-off型），即GC在给光或撤光时都能引起细胞发放频率增加。1953年Kuffler^[28]在给猫做实验室时发现猫的GC感受野在空间上呈现同心圆拮抗模式，即感受野的中心与外周在功能上相互拮抗。由此感受野可分为两类：给光-中心型（on-center）和撤光型（off-center），如图2-4所示。

图 2-4 ON 型和 OFF 型神经节细胞感受野的反应形式^[24]。左图为 ON 型感受野，右图为 OFF 型感受野，图中+代表兴奋区，-代表抑制区。

用单独的光点刺激ON型感受野的中心区域，GC发放频率增加，用光点刺激感受野的周边区域，细胞受到抑制，发放频率降低，甚至低于GC细胞的自发频率。OFF型感受野则刚好相反，用光点刺激其中心区域，GC发放频率降低，撤光时频率增加，对其周边区域给光，GC发放频率增加。

Rodieck^[29]于1965年提出了一个同心圆拮抗式感受野的数学模型，又称高斯差模型（difference of the Gaussians, DOG），具体见图2-5。它是由一个中心兴奋而外周抑制的同心圆组成，且中心和外周都具有高斯分布的性质。中心机制的峰敏值更高，周边机制尺度更大但峰值更低，且两个高斯方向相反，是相减抑制的关系，具体公式如下：

$$S(x) = k_c \times \exp(-x^2 / 2r_c^2) - k_s \times \exp(-x^2 / 2r_s^2) \quad (2-1)$$

其中， k_c 和 k_s 分别代表感受野中心区和周边区的最大敏感度。 r_c 和 r_s 分别代表感受野中心区和周边区的敏感度从最大值降至1/e时所对应的半径。

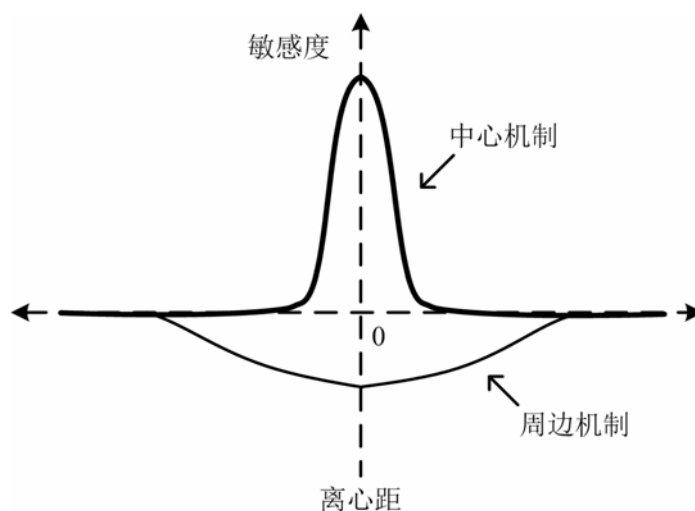


图 2-5 DOG 感受野模型^[24]

2.2 视觉感知机制

2.2.1 视敏锐度

视敏锐度（Visual Acuity，俗称视力）是指人眼辨别物体形态细节的能力。一般来说人眼能察觉的物体细节越微小表示视敏锐度越大。在日常生活中，医生常用含有E字母的视力表来测试我们的视敏锐度。背景亮度对视敏度有很大的影响，在较暗的环境中，人们对物体的细节和颜色辨别能力较弱，随着亮度的提升，人们辨别物体细节特征的能力变强，也能够识别场景中的颜色。图2-6^[9]给出了由实验得出的视敏锐度和背景亮度之间的关系，其中背景亮度涵盖了从白天到黑夜的

动态范围。从图像可以看出，在白天亮度为 $3\log\text{cd}/\text{m}^2$ 时，人眼的是视敏锐度为 $50\text{cycles}/\text{degree}$ ，而当亮度降为 $-3.3\log\text{cd}/\text{m}^2$ 时，视敏锐度也下降到 $2\text{cycles}/\text{degree}$ 。图中的曲线还可以用来预测在不同的亮度状态下场景细节的可视度^[9]。

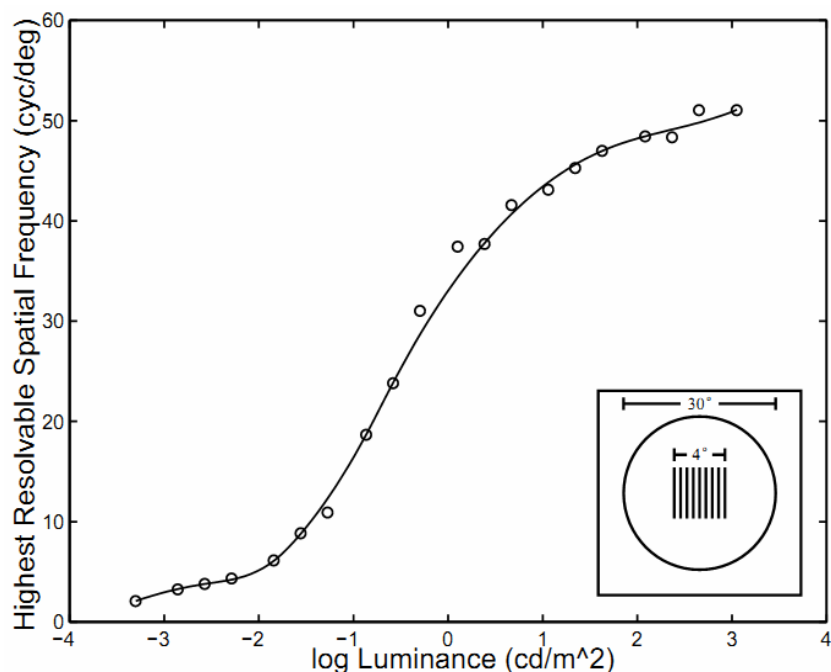


图 2-6 视敏度与背景亮度的关系^[9]

2.2.2 视觉的暗适应与明适应机制

在人们的现实生活中，真实场景的动态范围往往是十分巨大的，从黑夜的星辰到晴朗的白天，亮度的差值高达10的10次方数量级以上。人眼之所以能够对动态范围如此巨大的场景产生响应，很大程度上是由视觉的适应机制决定的。当人们从明亮的户外进入昏暗的电影院时，起初几乎看不见任何物体，但随着时间的推移视觉感知慢慢恢复，可以看见室内的东西，我们将这种视觉逐渐恢复的过程称为暗适应，其主要依赖于视杆细胞的作用。相对的，当我们从黑暗的环境突然进入到光亮的环境时（例如从漆黑的房间走到明亮的室外），在最初的一瞬间会觉得炫目刺眼，无法看清周围的物体，而很快我们会恢复视觉，这样的适应过程就称为明适应，这个过程主要受视锥细胞影响。而引起这种现象的原因主要是感光细胞难以承受如此巨大的亮度变化，当收到的光刺激超过细胞最大阈值时便无法再产生响应。

视觉适应系统是通过协调瞳孔，视杆视锥系统，光化学反应以及神经机制来适应当前的环境^[30]。瞳孔通过调整直径的大小来控制光线的进入量，但是起到的作用是比较小的。感光细胞在接受光线作用时，其内部的感光色素会进行分解和

合成。在明视的状态下，感光色素的漂白（分解）速度大于合成速度，因此感光细胞能够接受的光子数量变少，对光的敏感程度降低。而在暗视环境中，感光色素的合成速度远大于漂白速度，因此对光的敏感度增强。

2.2.2.1 暗适应过程及特性

通过用微弱的闪光刺激暗室里的人，测量记录其暗适应过程中的阈值，由此可获得暗适应曲线。实验使用一个大面积的紫色闪光灯刺激视网膜可得到如图2-7所示的视觉阈值变化过程。我们可以将此暗适应曲线看作两个阶段，在第一阶段中，主要由视锥细胞主导，表现了明视系统在暗视环境中的功能变化。在最初的3~4分钟内，阈值大幅度下降至原来的1/100左右，然后达到一个平稳状态。在经过5~8分钟后，视杆细胞（暗视系统）开始发挥主导作用，阈值出现第二次快速大幅下降，直到暗适应后的30分钟左右下降才减缓。而视觉适应系统真正要达到最低阈值（也叫绝对阈值）大概是在暗适应1小时以后。

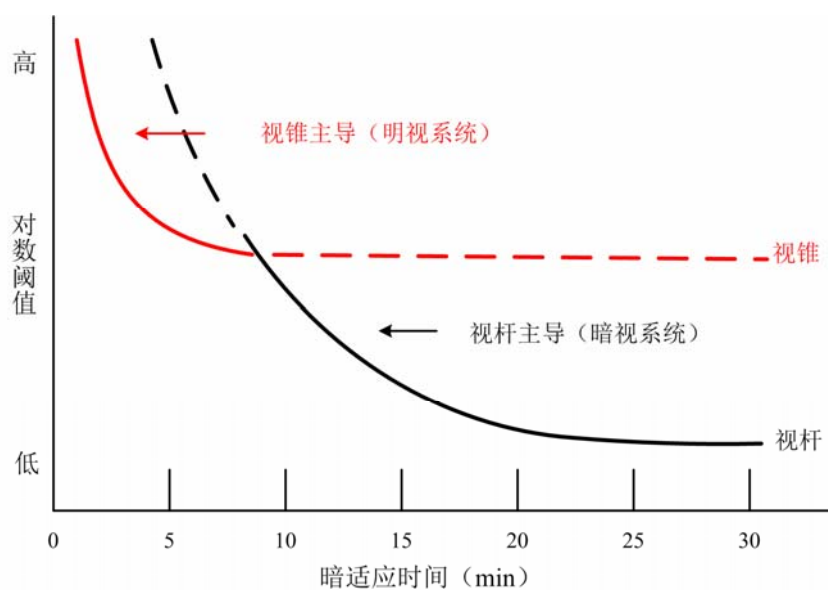


图 2-7 人眼的暗适应曲线

2.2.2.2 明适应过程及特性

明适应可以通过测量增阈值（increment thresholds）来研究，即找出能够引起视觉感知的最小亮度值。实验的大致流程如下：测试对象被要求注视一个背景亮度为 I_a 的大屏幕，直到眼睛完全适应背景为止。然后再在屏幕上呈现一个亮度为 $I_a + \Delta I$ 的刺激光点， ΔI 逐渐增加直到测试对象辨认出测试光点为止。研究人员将能够被测试人员辨认出的最小 ΔI 值以及对应的背景亮度 I_a 记录下来。这个经典的实验也叫做阈值-强度实验，由此得到的阈值与背景强度的关系曲线叫做TVI（threshold

versus intensity) 曲线, 图2-8是视杆和视锥的TVI曲线。从图中可以看出随着背景亮度的增加, 能够引起视觉感知的阈值也逐渐增加 (曲线上升)。我们可以将曲线分为三个阶段, 在亮度值较低的时候, 背景变化对阈值的影响较小 (曲线平缓)。随着背景亮度的上升, 背景与阈值的比值基本保持不变, 即曲线的斜率恒定, 这就是著名的韦伯-菲纳关系 (Weber-Fechner): $\Delta I/I = \text{常数}$ 。然而, 当背景亮度上升到 $2 \log \text{cd/m}^2$ 以上时, 微小的背景亮度增加都会导致视杆阈值的陡增, 由此表明此时的视杆系统已达到饱和状态, 难以识别出背景亮度的变化^[31]。

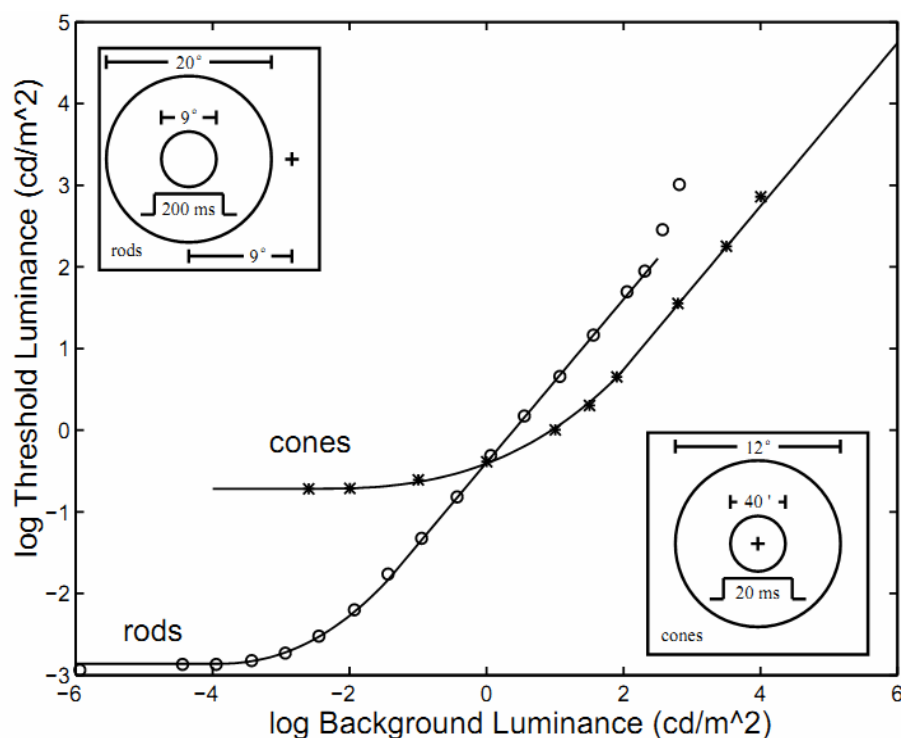


图 2-8 视杆视锥阈值与背景亮度的关系^[9]

2.2.3 感光细胞的亮度响应模型

感光细胞会将接收到的光能转换为神经响应, 人们通过细胞间的记录实验发现视锥和视杆细胞的响应特性遵循以下规则, 即不论背景亮度的大小如何, 感光细胞仅对范围很小的强度有响应, 大约为 $3 \log \text{cd/m}^2$ 的亮度范围。实验将已适应暗视环境的视杆视锥细胞暴露在不同的光强下, 测试其神经响应值, 得出了视杆和视锥细胞的响应曲线, 如图 2-9 所示^[32], 其中左边曲线表示视杆的响应, 右边为视锥的响应。

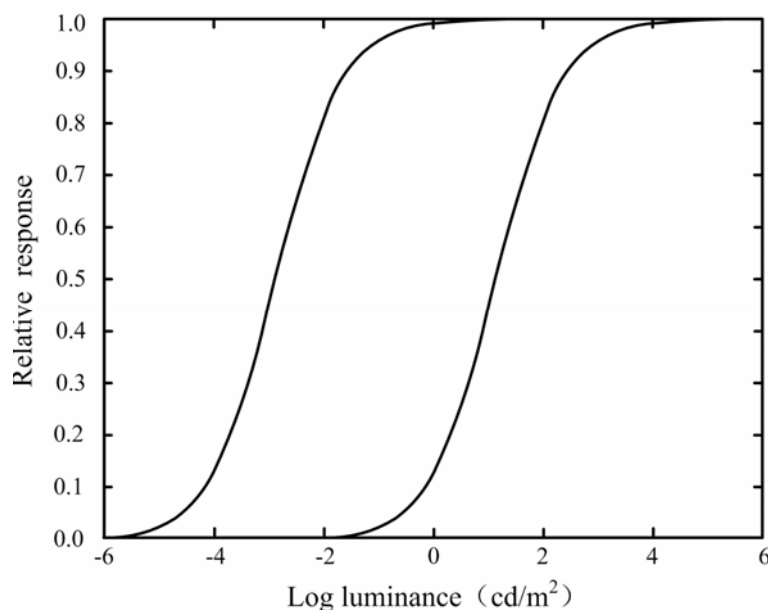


图 2-9 暗视环境下视杆视锥的亮度响应曲线^[32]。左边为视杆的响应曲线，右边为视锥的响应曲线。

如图 2-9 所示，视杆视锥的响应可以用一条 S 型曲线表示，我们可以通过以下公式来进行模拟：

$$\frac{R}{R_{\max}} = \frac{I^n}{I^n + \sigma^n} \quad (2-2)$$

该公式是 Naka 和 Rushton^[33]于 1996 年提出用以描述鱼的 S 电位，曾多次出现在心理物理学实验^[34]和神经检测实验中^[35]，并被发现可以用来表示视锥细胞^[36, 37]和视杆细胞^[38, 39]的响应。这里的 R 是感光细胞的响应 ($0 < R < R_{\max}$)， R_{\max} 是最大响应值， I 是当前光强。 σ 是一个半饱和常数，当 $\sigma = I$ 时， R 为最大响应值的一半。 n 是一个灵敏度控制指数，通常取值在 0.7 到 1 之间^[40, 41]。当 n 取 1 时，公式 (2-2) 又被称为米氏方程 (Michaelis-Menten equation)。 σ 的取值决定了 S 曲线的水平位置，可以简单的通过取两个大小不同的 σ 值来分别表示视杆和视锥响应，得出类似图 2-9 的结果。

从图 2-9 我们可以看出，如果将暗视环境中的感光细胞直接暴露在高亮度下，细胞的响应将达到最大值，细胞会出现饱和现象，此时无论亮度如何增加感光细胞都无法再察觉。这也和我们平时的生活经验相符，当我们的突然被曝光在亮度大于先前亮度 100 倍以上的环境中时，我们会觉得刺眼眩晕。当然，这种感觉不会持续太久，在很短的时间内，人眼的视觉系统就会适应新的环境，我们的视觉也会恢复到正常状态。

2.3 色适应变换

2.3.1 色彩学的基础知识

色彩并不是真实存在的实体，而是人对光的一种主观感觉。不同波长的光进入人眼后能够引起不同颜色的感觉，光照射到物体表面后被吸收和反射的光的波长是不同的，从而我们接收到的来自物体的光也就不同，因此物体有了不同的颜色。

在很早以前科学家就发现，光的颜色可以用波长来量化。人的视觉系统能够感应的波长大约在380nm~780nm之间，也就是常说的可见光，波长由长及短颜色的变化次序为：红、橙、黄、绿、蓝、靛、紫。视网膜中视杆和视锥细胞对不同波长的光吸收差异很大，具体可见图2-10所示。在视网膜内只有一种视杆细胞，其主要对蓝-绿光比较敏感，峰值波段在496nm左右。因为通过调节单色光的强度，总能够将视杆内的色素分子漂白到同一程度^[23]，因此当夜间只有视杆系统作用时，我们无法识别物体的颜色，只能区别亮度。在视网膜内有三种不同的视锥细胞，分别含有三种不同的感光色素。我们将这三种视锥细胞分别称为L、M、S型视锥细胞，它们对应的最敏感波长依次为长波、中波、短波。L和M型细胞能吸收的波长范围有很大的重叠，仅在峰值上有一点差异。其中L型细胞对红光比较敏感，峰值在559nm左右。M型细胞对绿光较敏感，峰值在531nm左右。S型细胞则对波长较短的紫-蓝光比较敏感，峰值约在419nm^[32]。当三种视锥细胞接收到光时，其中不同波段的光不能够仅通过调节光强而将三种感光色素漂白到同一程度，也就是说不同波段的波长对三种细胞的刺激是不同的。因此三种视锥细胞受光刺激后能够产生三种不同程度的兴奋，这些兴奋信息经过组合后传至大脑皮层，产生色觉。

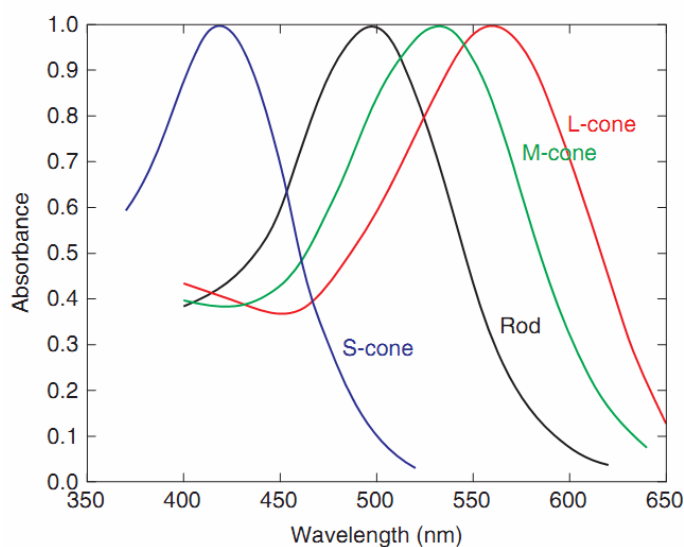


图 2-10 感光细胞对不同波长光的吸收光谱图^[42]

任何光引起的人对颜色和亮度的主观感觉，都可以由三种以内的适当波长的光通过调节光强混合而成，由此达到相同的颜色感知。人们通过颜色匹配实验确认了红（R）、绿（G）、蓝（B）为三种基本色（三原色），它们使得三种视锥的感光色素都有可能被漂白。任何亮度和颜色的光都可以通过调整RGB的比值混合而成，图2-11给出了具体的RGB色彩匹配函数^[43]。

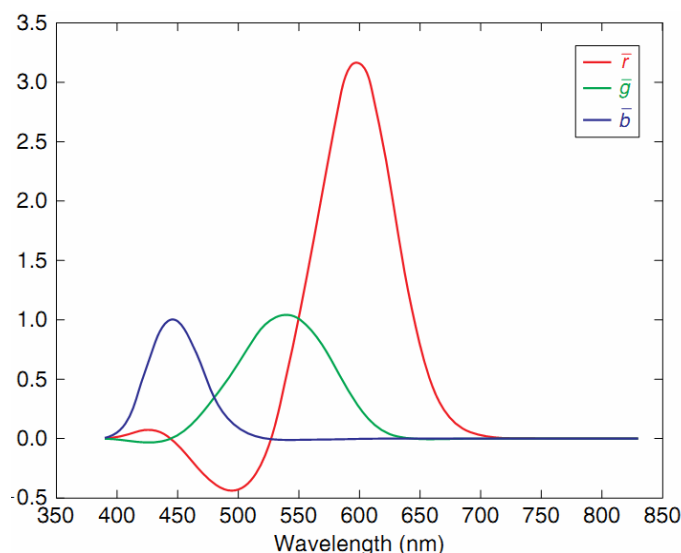


图 2-11 RGB 色彩匹配函数^[43]

虽然 RGB 色彩系统可以混合出任何颜色的可见光，但是有时为了得到某种特定光，R 的取值可能为负值。为了让色彩空间的三色值能够一直保持为正值，国际照明委员会（CIE）引入了一种新的色彩匹配函数，命名为 CIE XYZ，如图 2-12 所示。

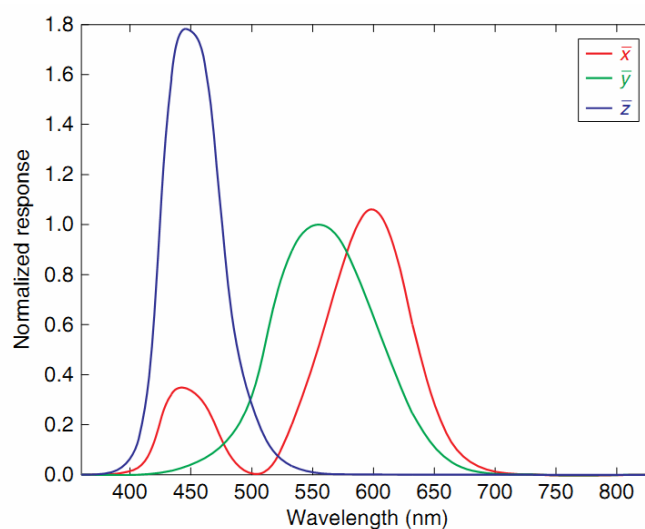


图 2-12 XYZ 色彩匹配函数^[32]

一般大家将 XYZ 中的 Y 值作为人眼对亮度的响应。RGB 和 XYZ 可以通过矩阵来相互转化，如公式（2-3）所示，该转换矩阵是由国际电信联盟无线电通信组（ITU-R）提出的。要注意的是，这里的 RGB 系统和图像处理、显示器等用的 RGB 空间是不同，其与任何的设备无关。

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2-3)$$

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

2.3.2 色适应变换

通过上一小节的介绍我们可以知道，人对物体颜色的感知主要是三种视锥细胞对光的响应的集合作用于大脑皮层后产生的一种主观感觉。那么，当外界光源发生变化时，物体表面反射的光也会随之变动，人感受到物体的颜色似乎也会变化。然而，真实的情况却不是这样，即使光源发生了变化，人眼感知的颜色依然会保持不变^[44]。例如，人们在日光下观看一张白纸，能感知到纸是白色的，当将这张纸放到白炽灯下观察时，虽然刚开始也会认为纸有些发黄，但在适应了白炽灯的环境后仍然感觉纸是白色的，这种现象被称为色适应（chromatic adaptation）。这种在感知物体表面颜色时不受外界光源亮度影响的能力叫做人眼的颜色恒常性（color constancy）^[32]。

在1902年，von Kries提出了一种假设，他认为人眼的色适应主要是由三种视锥细胞调节^[45]，色适应的过程就是红，绿，蓝三种视锥细胞各自对亮度的适应过程。Von Kries提出的这条简单的假设，成为现代色适应模型的基础，现有的很多色适应变化模型都是在此基础上发展起来的。色适应的模型就是将刺激转换到视锥的响应空间，然后根据当前亮度和期望的结果来分别调节红绿蓝三种成分。

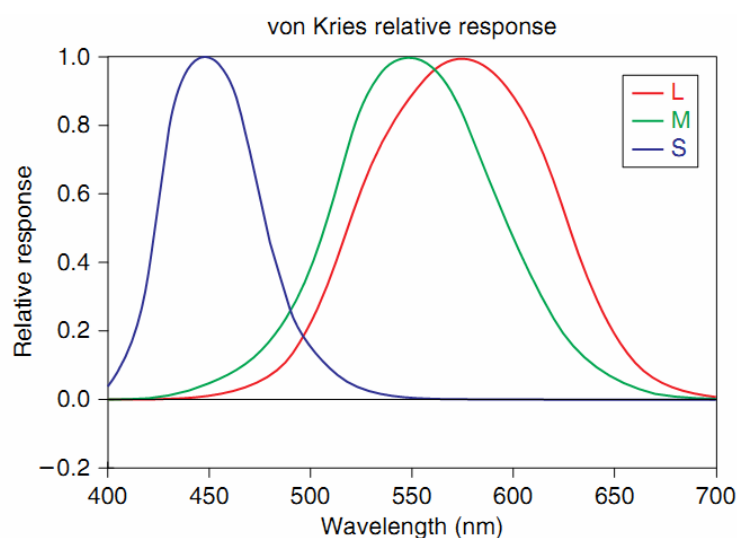
色适应变换过程可分为两个主要步骤（流程见图2-13），首先是将原条件下的 $X_s Y_s Z_s$ 三刺激转换到LMS视锥空间，这里需要利用色适应转换矩阵M，也是整个模型中最重要的部分。第二步则是根据源条件白点和期望条件白点的关系，将源条件下的 $L_s M_s S_s$ 响应转换为期望条件下的 $L_d M_d S_d$ 。最后利用M的逆矩阵将LMS空间还原到XYZ空间，得到色适应后的结果。



图 2-13 色适应变换流程

将 XYZ 空间转换到视锥的 LMS 空间采用的是一个 3×3 的转换矩阵 M ，它是整个模型中最核心的部分。目前主要的转换矩阵有三种，分别是 von Kries 提出的 $M_{\text{von Kries}}$ ，Bradford 提出的 $M_{\text{Bradford}}^{[46]}$ ，以及用于 CIECAM02 色貌模型的 M_{CAT02} ，具体参数见公式 (2-4)，其中 von Kries 的 LMS 空间函数如图 2-13 所示。

$$\begin{aligned}
 M_{\text{von Kries}} &= \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \\
 M_{\text{Bradford}} &= \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix} \\
 M_{\text{CAT02}} &= \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix}
 \end{aligned} \tag{2-4}$$


 图 2-13 von Kries 的色适应变换响应函数^[47]

2.4 本章小结

本章主要介绍了和视觉系统相关的生理基础以及理论模型。首先介绍了视网

膜的生理构造以及神经节细胞的感受野机制。然后介绍了视觉感知相关的理论模型，重点介绍了暗适应和明适应的一些物理特性，以及明暗适应机制的响应模型。最后介绍了色彩学的基础知识，以及人眼的色适应现象和图像处理中色适应模型的基本原理。本章中介绍的基础知识和理论模型是本文算法提出的前提，有了这些理论作为支撑，我们才能够学习并将其运用起来，创新更好的解决方案。

第三章 基于 Tumblin 色调映射框架的高动态图像增强算法

在第二章中我们详细的介绍了视网膜的生理机制以及明暗适应相关的数学模型。我们基于这些已有的研究以及Tumblin的色调映射框架，提出了一种改进的高动态范围图像的亮度压缩算法。在本章中，首先要介绍Tumblin色调映射框架的基本原理和已有的相关研究。然后重点介绍自己提出的改进后的算法以及结果的展示和分析。

3.1 基于视觉特性的 Tumblin 色调映射框架原理及相关研究

我们在第一章中就简要介绍过色调映射技术目前的发展状况，并将现有的工作分为了两大类，即全局算法和局部算法。全局算法简单且消耗计算资源少，运算速度快，适合进行实时处理，但在细节保留和局部对比度压缩上存在问题，容易造成图像的信息丢失。局部算法在这方面就做的比较好，能够保留清晰的细节，同时由于考虑了周边像素对中间像素的影响，拥有更好的对比度，但在亮度变化比较大的边界上容易出现伪影。针对高动态图像色调映射技术的研究在很早以前就开始了，并于1993年由Tumblin等人^[6]正式提出。在后续色调映射技术的研究发展中，有很多研究人员都采用了Tumblin提出的色调映射框架（例如[6,8]），接下来我们就介绍一下这个由Tumblin提出的算法框架的原理，以及基于此框架的相关研究工作。

3.1.1 Tumblin 色调映射框架

场景所呈现出来的外在表现主要取决于两个因素，场景的亮度级别以及对比度的范围^[48]。这些都体现在我们的日常生活中，例如，在晴天的时候场景的颜色会显得更加鲜艳，场景中明暗处的差异也更加明显。在夜间的时候，即使颜色丰富的场景也会变的灰暗。而在月光下，我们看到的场景会有一点偏蓝色。正是由于这些现象，想要使得处理后的场景匹配低动态的显示设备，仅对亮度水平和对比度范围进行简单的比例缩放或者压缩，并不能有效地重建出满足人视觉感知的场景。因此，Tumblin和Rushmeier建议利用视觉模型来解决这个问题，他们提出了如图3-1所示^[32]的算法框架。

从图 3-1 中我们可以看出，Tumblin 的算法框架可以分为两个部分，框架图的上面一排为正向处理，下面一排为逆向处理。在正向处理中，输入的是相机采集

到的场景亮度，通过一个和场景属性相关的正向模型后得到一个中间值，将中间值输入到和显示设备属性相关的逆向模型中，得到最终在设备上显示的图像。该算法的主要目的是想达到真实观察场景和经过色调映射后在显示器上的 HDR 图像能够在人的视觉感知上相匹配。这个算法框架的具体应用可以通过下面 Pattanaik^[11]的工作来进一步了解。

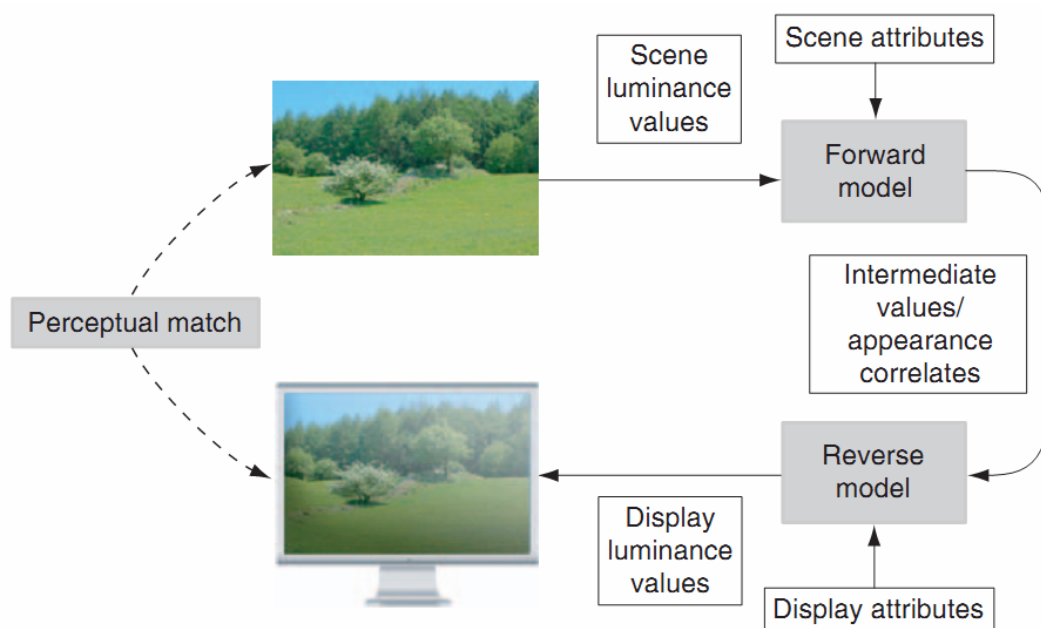


图 3-1 色调映射算法的概要图。最终的目标是要达到真实观察的场景和经过色调映射后在显示器上的 HDR 图像能够在人的视觉感知上相匹配。

3.1.2 相关算法

2000年，Pattanaik^[11]等人提出了一种时变的色调映射算法，采用的是Tumblin提出的算法框架，并增加了视觉适应模型（Adaptation Model）和视觉展示模型（Appearance Model）来表现视网膜响应、亮度以及颜色。其中视觉适应模型是Hunt等人^[49]提出的一个色觉静态模型的简化，Pattanaik在此基础上考虑了时间的变化，增加了一个时变的成分。该算法在对亮度通道进行动态范围压缩时，利用了视杆和视锥的亮度响应模型，也就是在第二章中介绍过的Naka-Rushton亮度响应模型。在Pattanaik的算法中，场景中的每个像素对应的视杆视锥亮度响应的具体计算公式如（3-1）所示：

$$R_{rod} = B_{rod} \frac{L_{rod}^n}{L_{rod}^n + \sigma_{rod}^n}$$

$$R_{cone} = B_{cone} \frac{L_{cone}^n}{L_{cone}^n + \sigma_{cone}^n} \quad (3-1)$$

这个公式和Naka-Rushton公式非常的类似，只是对每一个像素都分别计算了其视杆和视锥的响应值。其中， L 是输入亮度， B 是Hunt模型中感光色素漂白参数。 σ_{rod} 和 σ_{cone} 分别表示视杆和视锥的半饱和参数。 L 、 B 以及 σ 的取值都和全局亮度 A 相关。在Hunt的模型中， L 、 B 和 σ 的取值都是针对静态的图像，而Pattanaik对全局亮度 A 进行了改进，使得其能够随着时间而变化，具体的公式及参数取值可以参看文献[6]。虽然Pattanaik提出的时变算法考虑了视觉机制的影响，但其为了降低运算时的资源消耗，提高运算速度，采用的是全局亮度来进行计算，这并不符合人眼观看场景时的真实情况。

2004年Ledda^[21]等人在Pattanaik工作的基础上提出了一个相似的局部算法。其在计算每个像素的适应状态时考虑了周边像素的信息，采用了双边滤波来计算，以此避免伪影的产生，但是算法没有考虑处理前后颜色适应的问题。

3.2 基于 Tumblin 色调映射框架的高动态图像增强算法

我们基于Tumblin的色调映射框架提出了一个静态的高动态图像增强算法。这个算法参考了Pattanaik^[11]和Ledda^[21]的工作，并对其中未能考虑到的问题进行了改进，并增加了颜色适应模型。下面将详细介绍算法的流程以及最终的结果。

3.2.1 算法框架

我们的算法框架如图 3-2 所示，首先针对输入的场景计算视杆和视锥细胞的适应亮度，这里采用的是双边滤波算法，通过周边像素来计算中间像素的适应亮度。然后利用亮度响应模型分别计算视杆和视锥对亮度的响应，采用的是 Naka-Rushton 亮度响应公式。同时，我们假定输入场景和显示场景对视觉引起的响应值相等，由此将高动态的图像映射到低动态的显示设备上。最后通过色适应变换，矫正输出图像的颜色，得到最终的结果。

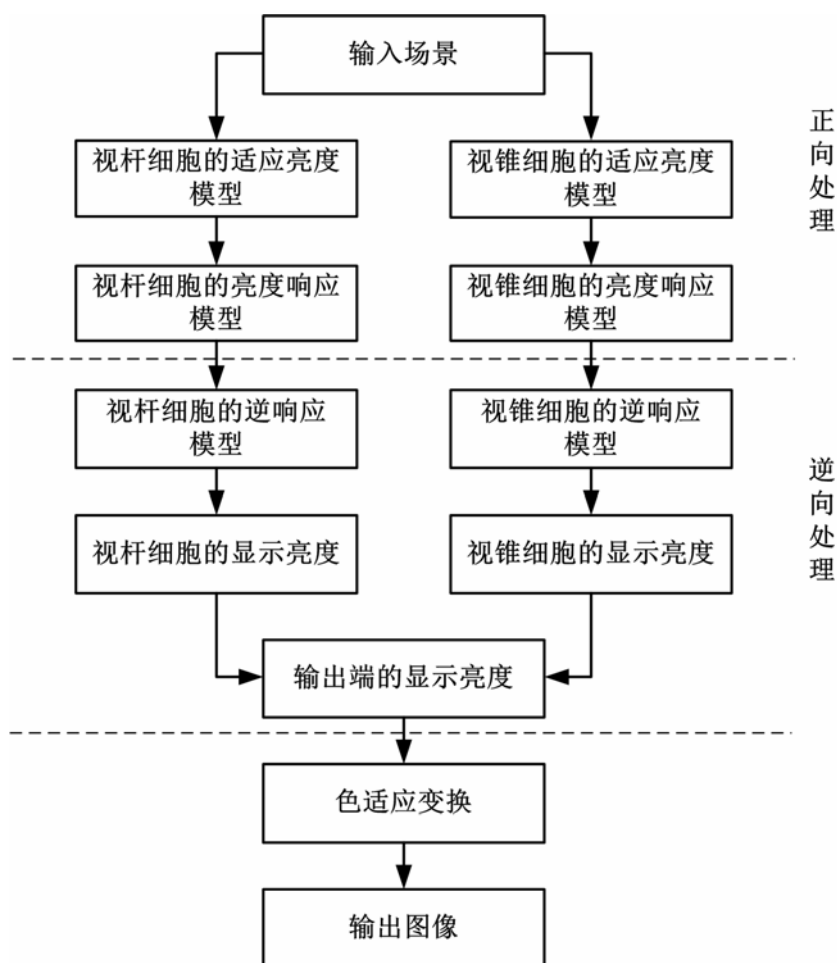


图 3-2 算法流程图

3.2.2 算法流程详细介绍

本节中，我们将按照图3-2给出的算法流程，详细的介绍每个模型的工作原理及计算方式。考虑到色适应变换并没有包括在Tumblin的色调映射框架中，因此我们将其放在下一小节中单独介绍。

3.2.2.1 适应亮度模型

在动态范围较大的场景中，人眼不仅可以看清亮处的物体，也能够看清暗处的物体，这是因为人眼具有局部适应的特性。我们可以推测，视网膜中的感光细胞的适应状态并不仅与自身的感受亮度相关，也与周边细胞的感受亮度有关系。因此，我们采用双边滤波来计算当前像素点周边的平均亮度，该滤波器常用于边缘保护，可以避免伪影。双边滤波器是一个非线性滤波器，它的输出是输入的加权平均值。每个像素点的适应亮度 L_a 的具体计算方式如公式（3-2）所示：

$$L_a(v) = \frac{\sum_{u \in s(v)} f(u-v)g(I_u - I_v)I_u}{k(v)} \quad (3-2)$$

$$k(v) = \sum_{u \in s(v)} f(u-v)g(I_u - I_v)$$

其中， u 是中心 v 的邻域像素，用 $s(v)$ 表示。 $k(v)$ 是一个归一化因子， $f(\cdot)$ 是一个空
间距离函数， $g(\cdot)$ 是灰度值距离函数。

3.2.2.2 亮度响应模型

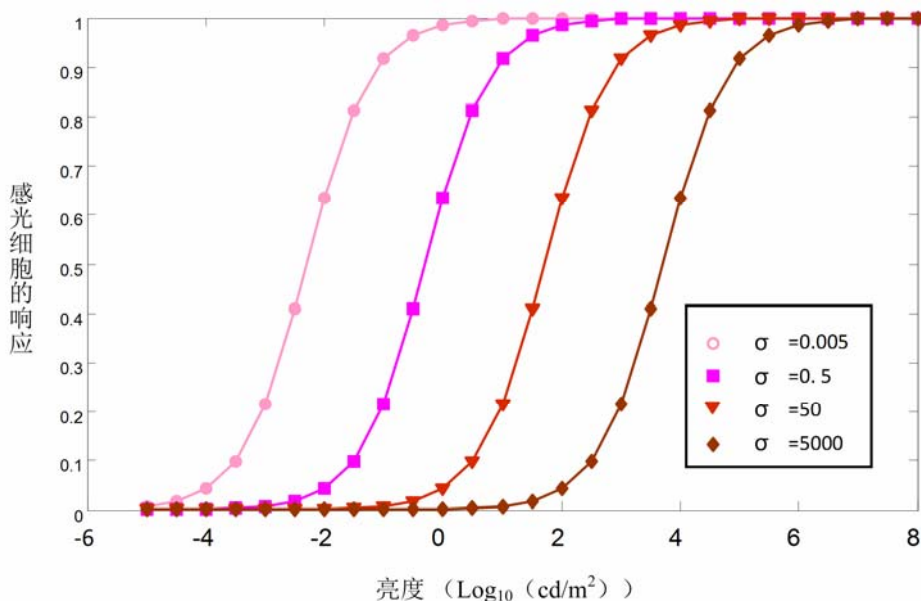
人眼之所以能够感知巨大的亮度范围，主要原因在于其不断的调整响应范围，
适应当前的环境。我们采用Naka和Rushton提出的感光细胞对亮度的响应公式来作
为我们的亮度响应模型，具体的表达方式如公式（3-3）所示：

$$R_{rod}(x, y) = \frac{L_{in}^n(x, y)}{L_{in}^n(x, y) + \sigma_{rod}^n(x, y)} \quad (3-3)$$

$$R_{cone}(x, y) = \frac{L_{in}^n(x, y)}{L_{in}^n(x, y) + \sigma_{cone}^n(x, y)}$$

其中 R_{rod} 和 R_{cone} 分别代表视杆和视锥对亮度的响应值， L_{in} 是当前像素点的亮度，
 σ_{rod} 和 σ_{cone} 分别是视杆和视锥的半饱和参数。 n 是一个灵敏度控制参数，一般取值
在0.7到1之间，通过计算实验比较后我们将 n 取0.8，且对视杆和视锥无差异。

亮度响应模型的关键在于半饱和参数 σ 的取值，图3-3给出了不同 σ 取值下响应
曲线的变化。横坐标代表当前点的亮度，纵坐标表示感光细胞对应的响应值，四
种颜色的曲线分别对应四种不同适应状态下 σ 的取值，从0.005到5000。由于此处
仅考虑 σ 的影响，所以每一条曲线的 n 值都取0.8。从图中我们可以看出， σ 的取值
主要决定了S型响应曲线的水平位置，从而直接影响了像素点在曲线上对应的响应
值。同时，我们可以用这个图来解释视觉的明暗适应过程。例如，当我们从明亮
的室外走入黑暗的房间时，在开始的一段时间，我们无法看清周边的环境。造成
这样的原因在于，当我们进入暗室时， σ 的值仍处于高亮度适应状态，而场景中
真实的亮度是极低的，因此在曲线上对应的响应也是极低的。然而在一段时间之
后，适应状态的亮度下降， σ 的值也变小，对应的响应值逐步增加，这就是暗适
应的过程，同理可解释明适应的过程。由此我们也可以看出 σ 的值与当前的适应
状态有关。

图 3-3 不同 σ 取值下对应的亮度响应曲线

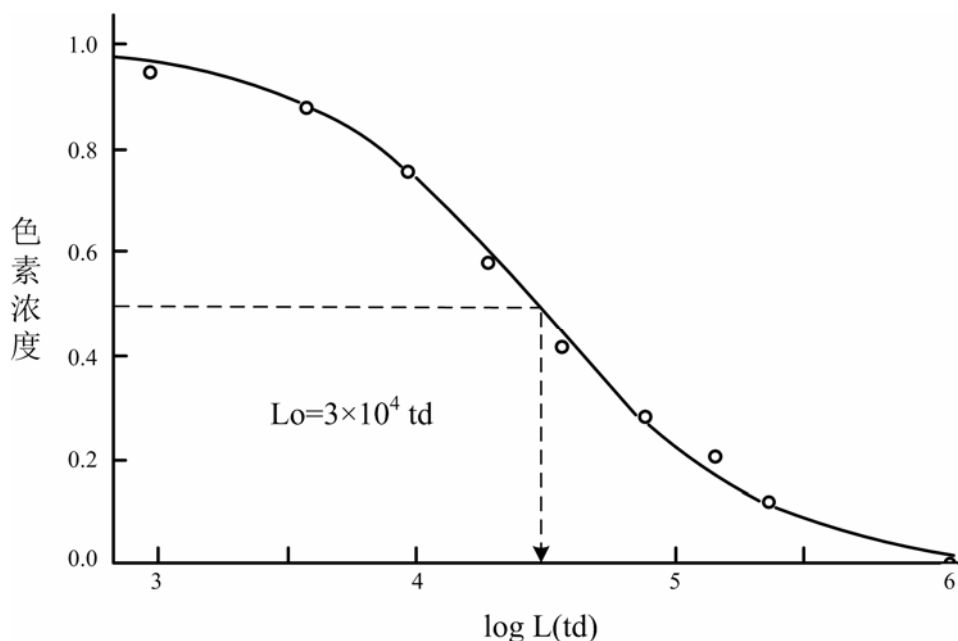
在第二章中我们介绍过，感光细胞受到光线的刺激时，其内部的感光色素会进行分解和合成。在明视的状态下，感光色素的漂白速度大于合成速度，因此能够接受的光子数量变少，对光的敏感程度降低。而在暗视环境中，感光色素的合成速度远大于漂白速度，对光的敏感度增强。由此我们可以看出，人眼对当前亮度的适应与感光色素的浓度有很大的关系。未漂白的色素浓度越高适应状态越低，而半饱和响应参数 σ 的值又与适应状态成正比，因此我们可以将感光色素浓度与参数 σ 的关系定义如下：

$$P(x, y) \times \sigma(x, y) = C \quad (3-4)$$

其中 P 是当前未漂白色素的浓度， C 为一个常数，表示色素浓度 P 与参数 σ 成反比关系。同时，我们采用 Rushton^[50] 给出的色素浓度与当前亮度的关系来求解 P 的值，具体公式如下所示：

$$P(x, y) = \frac{L_o}{L_o + L_a} \quad (3-5)$$

其中， L_o 是将色素漂白为一半浓度的亮度值， L_a 是当前的适应亮度。在这里我们将视杆和视锥的 L_o 取不同的两个常数，分别为 2×10^3 和 2×10^4 。图3-4给出了不同亮度下色素浓度的变化情况，该图是基于 M.Hollins^[51] 的实验结果得出的，实验中的 $L_o = 3 \times 10^4$ td。

图 3-4 不同亮度下的色素浓度^[51]

3.2.2.3 逆变换模型

在Tumblin的色调映射框架中，假定了HDR图在输入端和映射后的显示端所引起人眼的感知是相同的，也就是映射前后引起的视网膜的响应是相同的，即 $R_{out} = R_{in}$ 。由此，我们可以通过变换公式（2-1）来求解输出端的图像亮度，对视杆输出亮度 L_{rodo} 与视锥输出亮度 L_{coneo} 的求解分别如下所示：

$$\begin{aligned} L_{rodo}(x, y) &= \sigma_{out} \left(\frac{R_{rod}(x, y)}{1 - R_{rod}(x, y)} \right)^{1/n} \\ L_{coneo}(x, y) &= \sigma_{out} \left(\frac{R_{cone}(x, y)}{1 - R_{cone}(x, y)} \right)^{1/n} \end{aligned} \quad (3-6)$$

在公式（3-6）中， σ_{out} 是输出端的半饱和响应参数，为了简化模型，我们不再对每个输出点计算其 σ_{out} 的值，而是采用和 Pattanaik^[11]类似方式，用一个全局亮度 La_{out} 来作为输出端的适应亮度。我们令 $La_{out} = 30 \text{ cd/m}^2$ ，然后通过公式（3-4）和（3-5）可以求出 σ_{out} 对应的视杆和视锥值。在求得输出端视杆亮度 L_{rodo} 和视锥亮度 L_{coneo} 之后，我们对其赋予不同的权重，将结果融合在一起。这里的权重是一个 sigmoid 函数，具体的表现形式可参看下一章 4.2.4.1 中对融合因子 w 的求解。

3.2.3 颜色适应变换

颜色适应的模型我们在第二章中已经简要的介绍过，这里采用的色适应变换

流程如图3-5所示：



图 3-5 颜色适应变换流程

首先是将色调映射之后的 X_{out} Y_{out} Z_{out} 三刺激转换到LMS视锥空间，这里我们采用的是von Kries提出的转换矩阵，具体转换如公式（3-7）所示：

$$\begin{bmatrix} L_{out} \\ M_{out} \\ S_{out} \end{bmatrix} = M_{\text{von Kries}} \begin{bmatrix} X_{out} \\ Y_{out} \\ Z_{out} \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} X_{out} \\ Y_{out} \\ Z_{out} \end{bmatrix} \quad (3-7)$$

然后我们根据色调映射后输出场景的白点和期望（原始输入）白点的关系，将输出场景下的 $L_{out}M_{out}S_{out}$ 响应转换为期望条件下的 $L_D M_D S_D$ 。最后利用 $M_{\text{von Kries}}$ 的逆矩阵将LMS空间还原到XYZ空间，得到色适应后的结果。整个变换过程可以通过一个对角矩来分别调整XYZ三刺激值达到目的，如公式（3-8）所示：

$$\begin{bmatrix} X_D \\ Y_D \\ Z_D \end{bmatrix} = M_{\text{von Kries}}^{-1} \begin{bmatrix} \frac{L_{whiteD}}{L_{whiteout}} \\ \frac{M_{whiteD}}{M_{whiteout}} \\ \frac{S_{whiteD}}{S_{whiteout}} \end{bmatrix} M_{\text{von Kries}} \begin{bmatrix} X_{out} \\ Y_{out} \\ Z_{out} \end{bmatrix} \quad (3-8)$$

其中， (X_D, Y_D, Z_D) 是色适应之后显示端对应的XYZ空间值。 $L_{whiteout}$ 、 $M_{whiteout}$ 、 $S_{whiteout}$ 是色适应前对应的白点， L_{whiteD} 、 M_{whiteD} 、 S_{whiteD} 是期望条件对应的白点。

3.2.4 实验结果

在本小节中，我们将展示算法处理的结果，并对结果进行分析。我们的算法是一个完全自适应的模型，因此对于展示的所有结果，我们在处理时采用的参数取值都是一样的，具体的取值参看3.2.2小节和3.2.3小节。为了更好的展示处理结果，我们首先处理了一幅室外场景图和一幅室内场景图，并放大展示了细节部分。同时，我们也在图3-8中展示了我们的结果与其他研究者的7种算法结果的对比。这里需要说明，我们所处理的图像均为.hdr格式（一种高动态范围图像格式）。

图3-6展示了室外场景的处理结果，左边为全景图，右边为细节图，是左边图像中用白色矩阵圈出部分的放大图。图3-6（a）-（c）分别是原图，色调映射之后的效果，色适应变换的结果。从图3-6（a）和3-6（b）的对比我们可以看出，算法

能够有效的压缩图像的动态范围，让原本无法展现出来的信息也能够在低动态的显示器上展现出来。从细节图3-6（d）和图3-6（e）的对比可以看出，算法处理后图像的边界和细节清晰明显。从图3-6（b）和图3-6（c）的比较则显示了色适应变换的作用，变换前图3-6（b）的颜色饱和度明显偏低，变换后的图像在视觉感知上更贴近真实的场景。



图 3-6 室外场景处理结果



图 3-7 室内场景处理结果

图 3-7 展示了室内场景“Hotel Room”的处理结果，其中图 3-7 (a) - (c) 为全景图，图 3-7 (d) - (f) 是对应的细节放大图，我们用白色矩阵将放大部分圈了出来。结果再次证明了我们的算法能够有效地压缩图像的动态范围，窗外的高亮部分和室内的暗区处理后都能够清晰的显示，且处理结果没有光晕出现。同时，图 3-7 (b) 和图 3-7 (c) 的对比也显示了色适应变换后图像的颜色饱和度更好，能够更好的匹配人眼视觉的感知。

为了更好的比较我们的算法与其他算法的处理效果，图 3-8 展示了一幅室内图

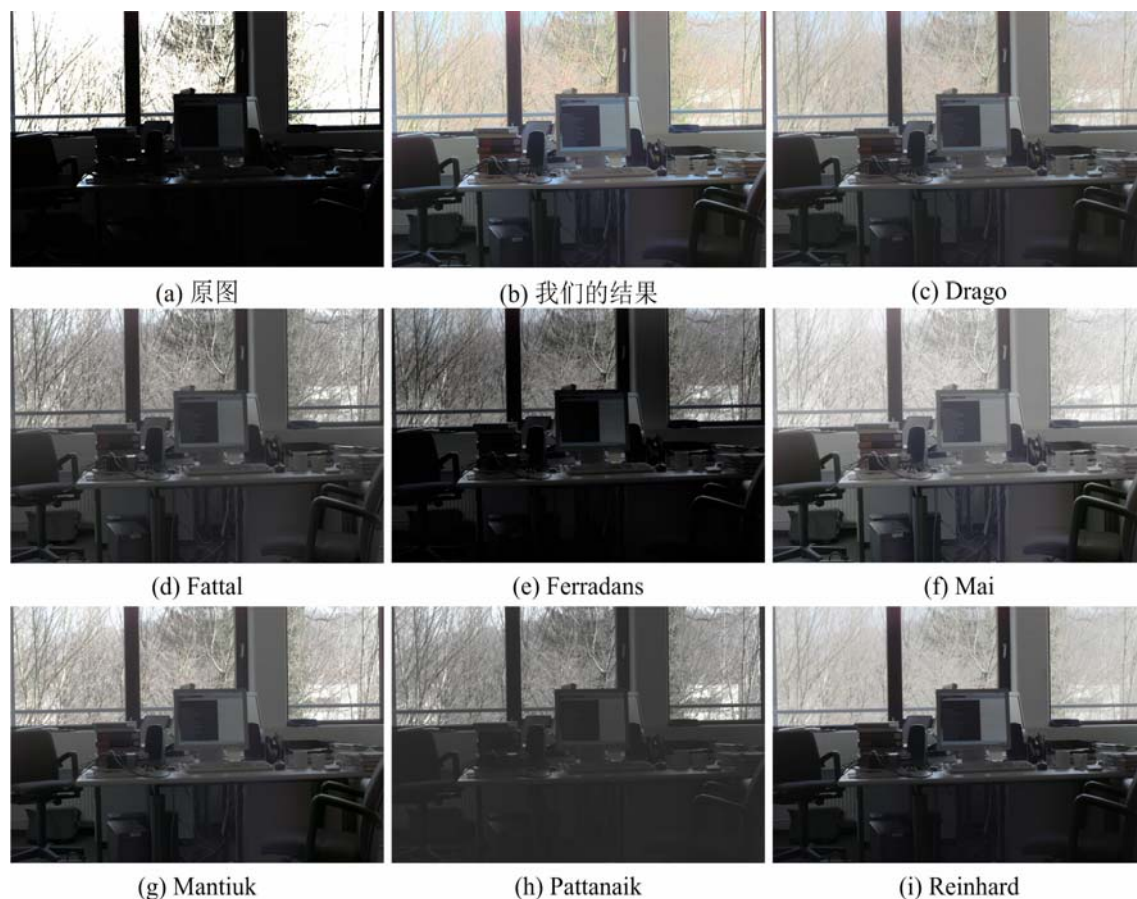


图 3-8 室内场景处理结果对比

像用 8 种不同算法处理后的结果。其中，图 3-8 (a) - (i) 分别为原图，我们处理的结果、Drago^[7]的结果、Fattal^[16]的结果、Ferradans^[52]的结果、Mai^[53]的结果、Mantiuk^[54]的结果、Pattanaik^[11]的结果以及 Reinhard^[18]的结果。以上这些研究人员的结果图均来自 Max Planck Institut Informatik (MPII) (<http://resources.mpi-inf.mpg.de/tmo/NewExperiment/TmoOverview.html>)。通过比较可以看出，我们的算法在图像的动态范围压缩和颜色适应方面都有很好的表现。

3.3 本章小结

在这一章中，主要介绍了我们基于Tumblin的色调映射框架提出的一种新的HDR图像压缩算法。我们的算法利用了感光色素的漂白机制来计算感光细胞的响应值，并在原框架正向模型和逆向模型的基础上，还考虑了视觉对颜色的适应，增加了色适应变换模型。从最后展示的结果可以看出，我们的算法能够有效的压缩图像的动态范围，在细节和边界显示上清晰明显，处理结果无光晕现象，色适应变换后的图像饱和度更贴近真实的场景。

第四章 模拟视觉适应机制的高动态图像增强算法

人眼能够承受场景的亮度范围是极广的，从日光下 10^8cd/m^2 到夜晚 10^{-6}cd/m^2 的亮度范围。其实人类视觉系统的细胞响应范围是有限的，单个神经元的响应只有大约 $1.5\log$ 单元^[55]，但是通过视觉系统的适应机制，改变细胞的灵敏度来对不同的亮度级别做出响应。我们可以将人眼的这种视觉适应机制看做是一个完美的色调映射模型，让人们能够在高动态范围的真实场景中同时看清过亮和过暗的区域。在第三章中，我们基于 Tumblin 提出的色调映射框架建立了一个高动态图像增强算法，在其他章节中我们也介绍过一些采用视觉机制来建立色调映射模型的方法，但是这些方法中大多将算法分为两个过程，即正向处理过程和逆向处理过程。然而，通过第二章的学习我们可以知道，人眼的视觉机制并没有逆向处理这样的过程，所以，这些算法都没有模拟出人眼真实的视觉适应机制。

在本章中，我们提出了一个全新的高动态图像增强算法框架，旨在通过模拟真实的视觉适应机制来压缩图像的动态范围，使得处理后的结果更加贴近人眼所见到的场景。我们首先介绍算法的框架，然后详细介绍算法流程中的各个模型，最后展示处理后的结果图，并分析算法中部分参数的作用。同时还会将我们的算法和已有的几个经典算法进行对比和分析。

4.1 算法框架

在我们的算法中，对图像动态范围的压缩主要是在亮度通道，并且整个压缩过程都是分为视杆和视锥两个平行通道进行。算法的流程见图 4-1，算法大致可以分为三个部分，由三个主要模型构成：

感光细胞的亮度模型：光线进入人眼后，感光细胞便能够接收来自光线的刺激。亮度模型将输入图像（即场景）转换为视杆细胞和视锥细胞的感受亮度。

感光细胞的响应模型：感光细胞受到光刺激后便会产生神经电信号。因此在第二步中，我们将视杆和视锥细胞的感受亮度分别转换为神经响应值。

神经节细胞的感受野模型：视杆视锥细胞产生的神经电信号最终将传入视网膜的神经节细胞中，在这里我们主要考虑了神经节细胞的感受野机制。

算法除了这三个重要的部分外，我们还考虑了视杆和视锥两个通道处理后的信息融合，以及保持颜色稳定的问题。这里需要特别说明的是，视觉的生理机制是非常复杂的，各细胞和层级之间有着错综复杂的联系，我们没有完全模拟所有生理机制，而是重点模拟了视觉适应过程中的生理物理特性。这样做的原因主要

是考虑到算法的简洁实用性，我们希望用尽量简化的模型来模拟视觉的适应过程。

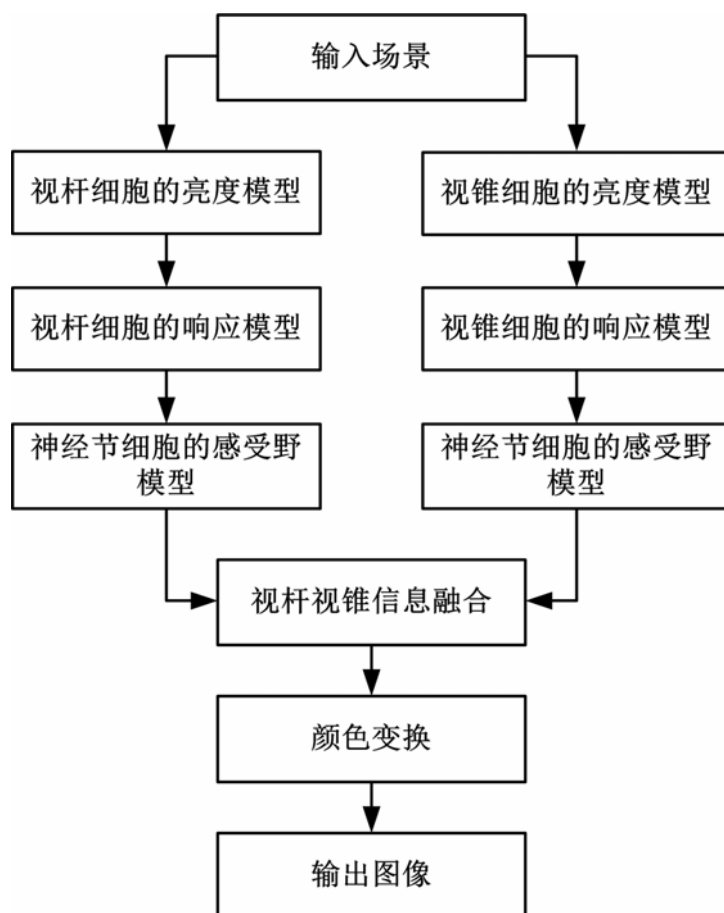


图 4-1 算法流程图

4.2 算法流程详细介绍

在本小节中，我们将根据图 4-1 中的算法框架，按照流程依次详细地介绍各部分的工作原理。

4.2.1 感光细胞亮度计算

我们的算法在对图像动态范围进行压缩时，所有的处理都是在亮度通道上进行的。在输入彩色的RGB图像后，我们需要将其转换为视觉系统中视杆和视锥细胞所感知的亮度。在研究中，通常用XYZ空间中的Y通道来代表视锥所感知的亮度，也称为CIE standard $Y^{[11]}$ 。同时，我们参考文献^[56]，用CIE暗视亮度函数 $V'(\lambda)$ 来表示视杆的感知亮度。具体的转换计算公式如下所示：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4-1)$$

$$L_{cone}(x, y) = \text{CIE standard } Y = 0.25R + 0.67G + 0.065B \quad (4-2)$$

$$L_{rod}(x, y) = V(\lambda) = -0.702X + 1.039Y + 0.433Z$$

公式(4-1)代表RGB空间转换到XYZ空间,其中所求Y代表了视锥(明视)所感知的亮度。这里用的转换矩阵和我们在第二章中给出的RGB到XYZ的转换矩阵有所不同,主要原因是第二章中是光学系统上的RGB颜色系统与XYZ颜色系统的转换,与任何设备无关。而这里的RGB是图像领域中的RGB空间,不是同一个概念。公式(4-2)中, L_{rod} 和 L_{cone} 分别表示视杆和视锥的感知亮度。

4.2.2 感光细胞的亮度响应模型

在前面的章节中我们就介绍过感光细胞的亮度响应模型,在这里我们依然选择Naka-Rushton公式来计算视杆和视锥细胞的亮度响应,具体表现形式如下:

$$R_{rod}(x, y) = \frac{L_{rod}^n(x, y)}{L_{rod}^n(x, y) + \sigma_{rod}^n(x, y)} \quad (4-3)$$

$$R_{cone}(x, y) = \frac{L_{cone}^n(x, y)}{L_{cone}^n(x, y) + \sigma_{cone}^n(x, y)}$$

其中 R_{rod} 和 R_{cone} 分别代表视杆和视锥对亮度的响应值, L_{rod} 和 L_{cone} 是在上一小节求出的视杆和视锥的亮度, σ_{rod} 和 σ_{cone} 分别是视杆和视锥的半饱和参数。 n 是一个灵敏度控制参数,我们对其的取值与之前相同,通过计算实验比较后视杆视锥的 n 值都取0.8。对于图像中的每一个像素点,其对应的半饱和参数以及响应值都是不同的,与当前像素点的感受亮度有关,由此可见,我们的算法是一个局部算法。

在上一章中我们便提到参数 σ 的重要性,在这里对于参数 σ 的计算方式,我们采用了Xie和Stockham^[57]于1989年提出的 σ 与背景亮度之间的经验关系:

$$\sigma = I_a^\alpha \times \beta \quad (4-4)$$

其中 I 为当前背景亮度, α 和 β 为两个常数, α 的取值为0.69。在我们的算法中,用 L_{rod} 和 L_{cone} 来代替当前的背景亮度 I ,具体的计算公式如下:

$$\sigma_{rod} = L_{rod}^\alpha(x, y) \times \beta_{rod}$$

$$\sigma_{cone} = L_{cone}^\alpha(x, y) \times \beta_{cone} \quad (4-5)$$

其中 α 的取值与文献[36]中相同，即 0.69。 β 的取值对视杆和视锥存在差异，其中视杆的 β 取值为 2，视锥的取值为 4。

4.2.3 神经节细胞的感受野作用

根据前面的计算，我们已经可以得出每个像素点的视杆和视锥的响应值。在视觉系统中，这些响应信号将通过双极细胞层传入到神经节细胞层。为了简化算法模型，我们在此仅考虑神经节细胞的作用。我们用神经节细胞经典感受野的数学模型，即高斯差模型（difference of the Gaussians, DOG）^[29]，来模拟感受野的作用。由于在第二章中已经介绍过 DOG 模型的基本原理，这里就不再赘述。在我们的算法中，感受野模型的具体表现形式如下：

$$\begin{aligned} G_{rod}(x, y) &= R_{rod}(x, y) \otimes (k_c \times \exp(-\frac{x^2 + y^2}{2r_c^2}) - k_s \times \exp(-\frac{x^2 + y^2}{2r_s^2})) \\ G_{cone}(x, y) &= R_{cone}(x, y) \otimes (k_c \times \exp(-\frac{x^2 + y^2}{2r_c^2}) - k_s \times \exp(-\frac{x^2 + y^2}{2r_s^2})) \end{aligned} \quad (4-6)$$

在公式（4-6）中， G_{rod} 和 G_{cone} 表示神经节细胞的响应输出。 k_c 和 k_s 分别表示感受野中心和外周的最大灵敏度（也就是峰值）， k_c 和 k_s 的取值是相对，在 k_c 保持不变的情况下， k_s 的取值越大，周边对中心的抑制作用越强。 r_c 和 r_s 分别表示感受野中心和外周的高斯标准差（也就是尺度），在运用时，中心的尺度 r_c 总是小于外周的尺度 r_s 。

4.2.4 视杆视锥信息融合及颜色变换

4.2.4.1 视杆视锥信息融合

视觉的适应机制有明视应和暗适应两种，其中明适应主要指人眼对明亮环境的适应，而暗适应主要是指对昏暗或者黑暗环境的适应，这两种适应机制分别依赖于视锥细胞和视杆细胞的作用。也就是说，在明亮的条件下，视锥系统起着主导作用，在黑暗的条件下，视杆系统占主导地位。因此，在对视杆信息和视锥信息进行融合的时候，我们也从这个角度进行了考虑，并采用一个 sigmoid 函数作为权重因子 w ，如公式（4-7）所示：

$$w(x, y) = \frac{1}{(1 - m) + L(x, y)^{-0.1}} \quad (4-7)$$

其中 m 是 $L(x, y)^{-0.1}$ 的最小值， $w(x, y)$ 是信息融合能够合时视锥的权重因子，其随着亮度 $L(x, y)$ 的增加而增加，基本模拟了视觉适应过程中明暗状态下的特性，即当

亮度越高的时候视锥起的作用越大，当亮度越低的时候，视杆的作用越大。亮度 L 与权重 w 的关系曲线见图 4-2。

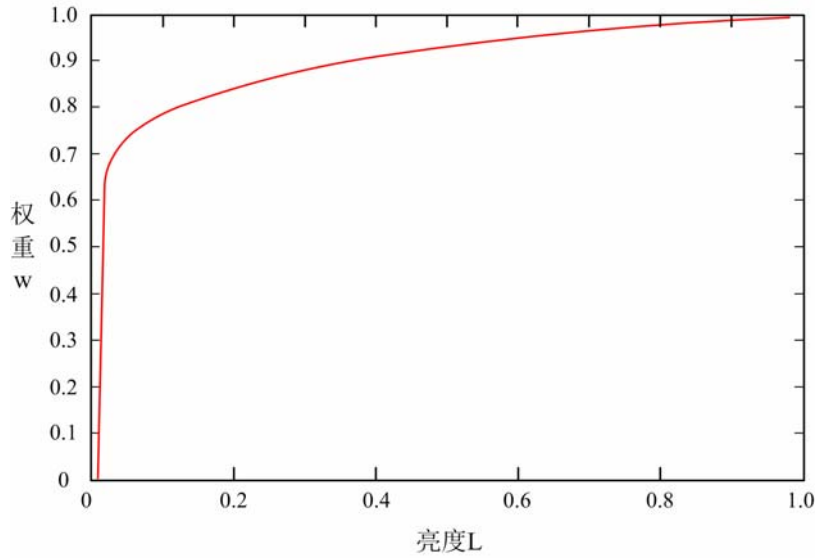


图 4-2 权重 w 与亮度的关系

视杆和视锥信息融合的具体计算公式如下所示：

$$L_{out}(x, y) = w(x, y) \times G_{cone}(x, y) + (1 - w(x, y)) \times G_{rod}(x, y) \quad (4-8)$$

其中 G_{rod} 和 G_{cone} 是神经节细胞感受野作用后的输出， L_{out} 是动态范围压缩后的亮度输出。

4.2.4.2 颜色变换

我们之前的所有处理都是针对图像的亮度通道进行的，在最后一步，我们需要将图像的颜色还原。为了将图像处理的色偏达到最小，我们保持了图像在处理前后 R、G、B 三个通道之间比例的不变^[58]，将输入图像的 R、G、B 三个通道同时乘以处理后的亮度 L_{out} 与处理前亮度 L_{in} 的比值。因此输出图像的 RGB 可以通过公式 (4-9) 来计算：

$$\begin{aligned} R_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{R_{in}(x, y)}{L_{in}(x, y)} \right) \\ G_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{G_{in}(x, y)}{L_{in}(x, y)} \right) \\ B_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{B_{in}(x, y)}{L_{in}(x, y)} \right) \end{aligned} \quad (4-9)$$

其中, R_{in} 、 G_{in} 、 B_{in} 分别表示输入图像对应的 RGB 三通道值, R_{out} 、 G_{out} 、 B_{out} 分别表示最终的输出图像对应的 RGB 三通道值。 L_{in} 是输入图像对应的亮度, 因为人眼对亮度的感知主要和 XYZ 空间中的 Y 分量有关, 所以在这里我们用 Y 分量来代替 L_{in} 的值。

当我们用公式 (4-9) 来还原颜色信息时, 对于有的图像会出现饱和度过高的情况。可能是因为原图某些区域 R、G、B 中的某通道占比较高, 在亮度升高以后, 其等比例放大后的数值较其他通道增大过多, 因此导致图像出现过饱和现象。为了解决此问题, 我们在原来的公式中引入了一个饱和度控制参数 s , 具体的表现形式如下:

$$\begin{aligned} R_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{R_{in}(x, y)}{L_{in}(x, y)} \right)^s \\ G_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{G_{in}(x, y)}{L_{in}(x, y)} \right)^s \\ B_{out}(x, y) &= L_{out}(x, y) \times \left(\frac{B_{in}(x, y)}{L_{in}(x, y)} \right)^s \end{aligned} \quad (4-10)$$

其中, s 作为一个饱和度调整参数, 取值在 0 到 1 之间^[32]。当 $s=1$ 时, 公式 (4-10) 就变成了公式 (4-9), 此时输出图像的 RGB 比值与原图相同, 图像的饱和度达到最高, 我们称之为全饱和。当 s 逐渐变小时, 颜色还原后的图像的饱和度也逐渐下降。在我们的算法中 s 的取值为 0.8。关于参数 s 的作用效果, 我们会在后面结果分析的时候进行展示。

4.3 结果分析与展示

在本小节中, 我们分析和展示算法的结果, 其中包括算法处理过程中的中间结果, 一些参数对结果的影响以及我们与其他算法的结果对比, 同时也会分析我们算法的优势和存在的问题。

对于在本节中出现的所有结果图, 其在算法处理中所取的参数值有以下说明: 我们的算法是一个高度自适应的算法, 因此在对图像进行处理时不用对参数进行调整, 所有处理都采用同样的默认参数值, 这些参数在各模型中的具体取值请参看上一小节 (4.2 节)。由于在 4.3.2 节中需要展示各参数变化对结果的影响, 因此会取不同参数值做比较, 如无特殊说明, 图像都是按照默认取值进行处理。

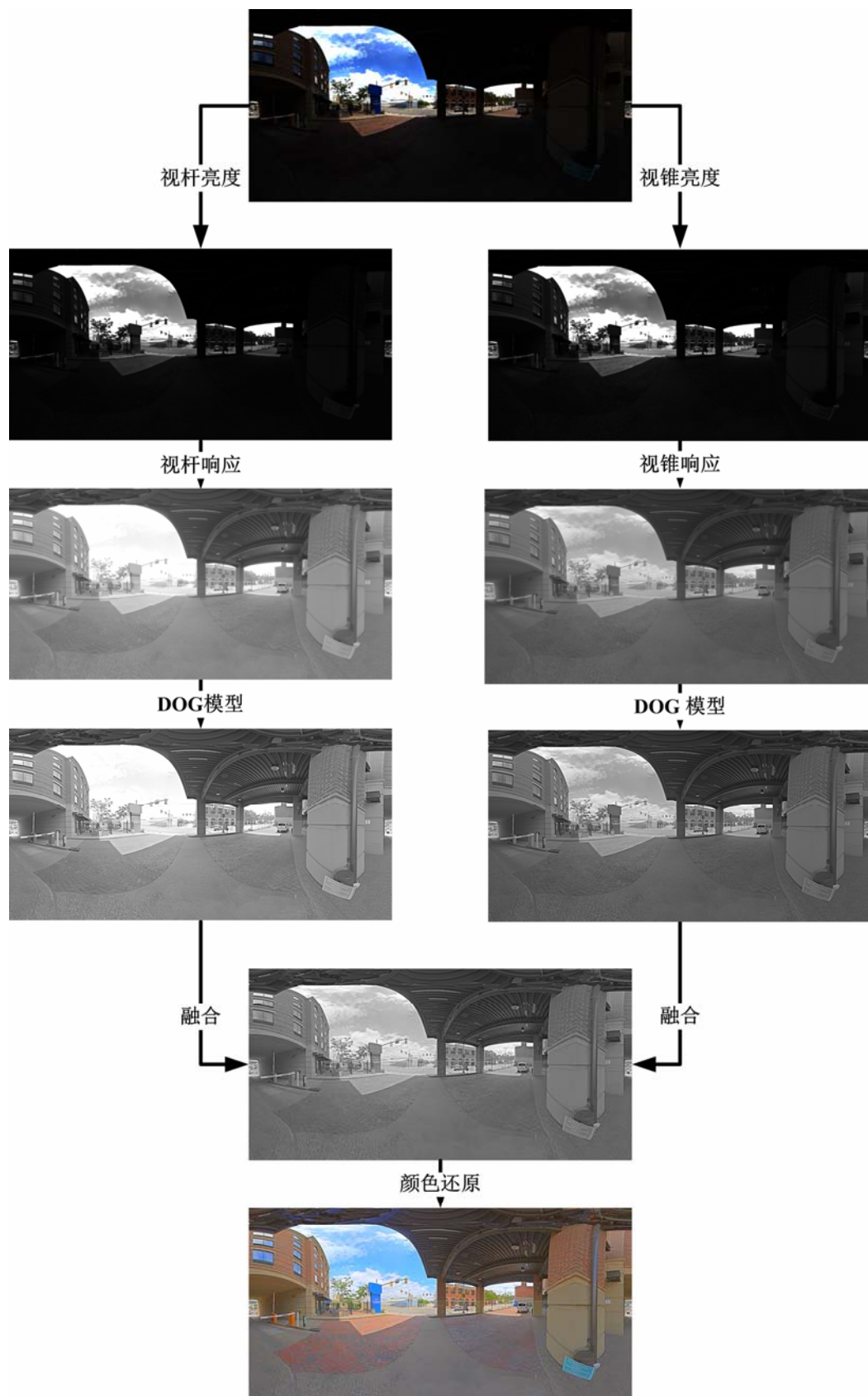


图 4-3 算法过程的中间结果图

4.3.1 算法过程的中间结果

为了更清楚的展示算法每一步的作用效果，我们选取了一张在暗处拍摄的室外高动态图像，图 4-3 展示了我们的算法在处理该图的过程中每一步得出的中间结果。其中左边为视杆通道的处理，右边为视锥通道的处理。流程图从上到下依次为原始图像（输入图像）、视杆和视锥的亮度图、视杆和视锥的响应图、经过感受野作用后的视杆和视锥信号、视杆和视锥信号融合后的图像、颜色还原后的图像（输出图像）。

从视杆和视锥的响应图可以看出，视杆和视锥系统都能在一定程度上对图像的亮度进行压缩。其中视杆系统对过暗区域的处理更为明显，能够更好的提升暗区的亮度，如图中靠近下面的地面部分。而视锥系统则对过亮区域有更好的效果，能够更有效的压缩过亮区域的亮度，如图中白云的部分。同时，由于 DOG 模型的作用，图像里的边界更加的清晰，拥有更多的细节信息。我们采用的融合权重因子 w 能够很好的发挥视杆和视锥系统各自的优点，使得融合之后的图像不论是在过暗还是过亮的区域都能够有很好的表现。

4.3.2 参数的影响

在上一节中我们便说明了 σ 取值对视杆和视锥亮度响应的重要影响，在这里我们将形象的展示不同 σ 取值对效果的影响。根据我们的对 σ 的计算方式可知，通过调整 β 和 α 值的大小可以改变 σ 的取值。为了简化影响因子的数量，我们在此处仅通过改变视锥的 β 值来观察不同 σ 取值下对结果的影响，视杆的 β 值和 α 的值仍保持不变，具体结果如图 4-4 所示。图 4-4 从左至右视锥系统中参数 β 的取值分别为 3、4、5。从结果中我们可以看出，随着 β 的增加，图像中过亮区域的亮度压缩更加明显。这是由于 β 越大， σ 的值越大，响应曲线右移，对亮度的压缩力度越大。

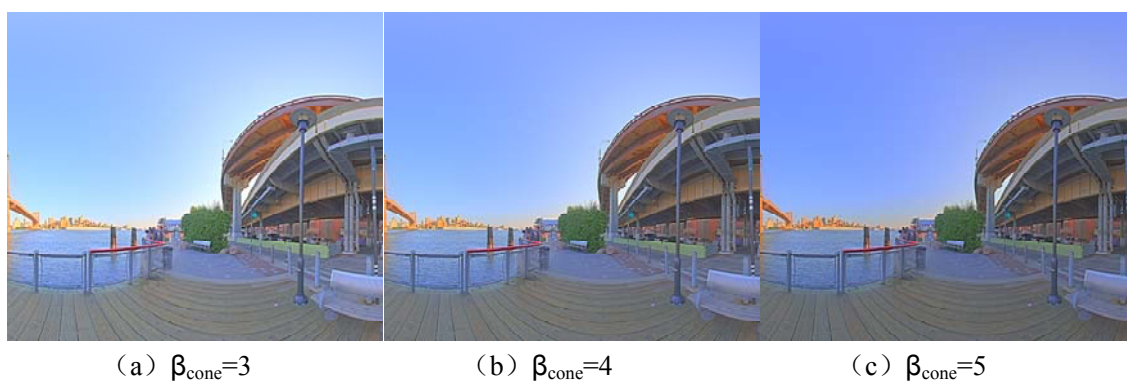


图 4-4 响应参数 σ 取不同值的结果

在上一节的颜色变换中我们加入了一个饱和度控制参数 s ，以此来解决部分图像处理后饱和度过高的问题。在图4-5中，我们展示了不同 s 取值下对应的颜色变换后的结果。图中从左至右参数 s 的取值分别为1、0.8、0.6，依次减小。从结果中我们可以看出，随着参数 s 的减小，图像的饱和度逐渐降低。当 s 取值为1时，还原后的图像饱和度最高，我们可以称之为全饱和。我们从图像的真实感和整体效果来考虑，在处理时 s 的取值一般为0.8。



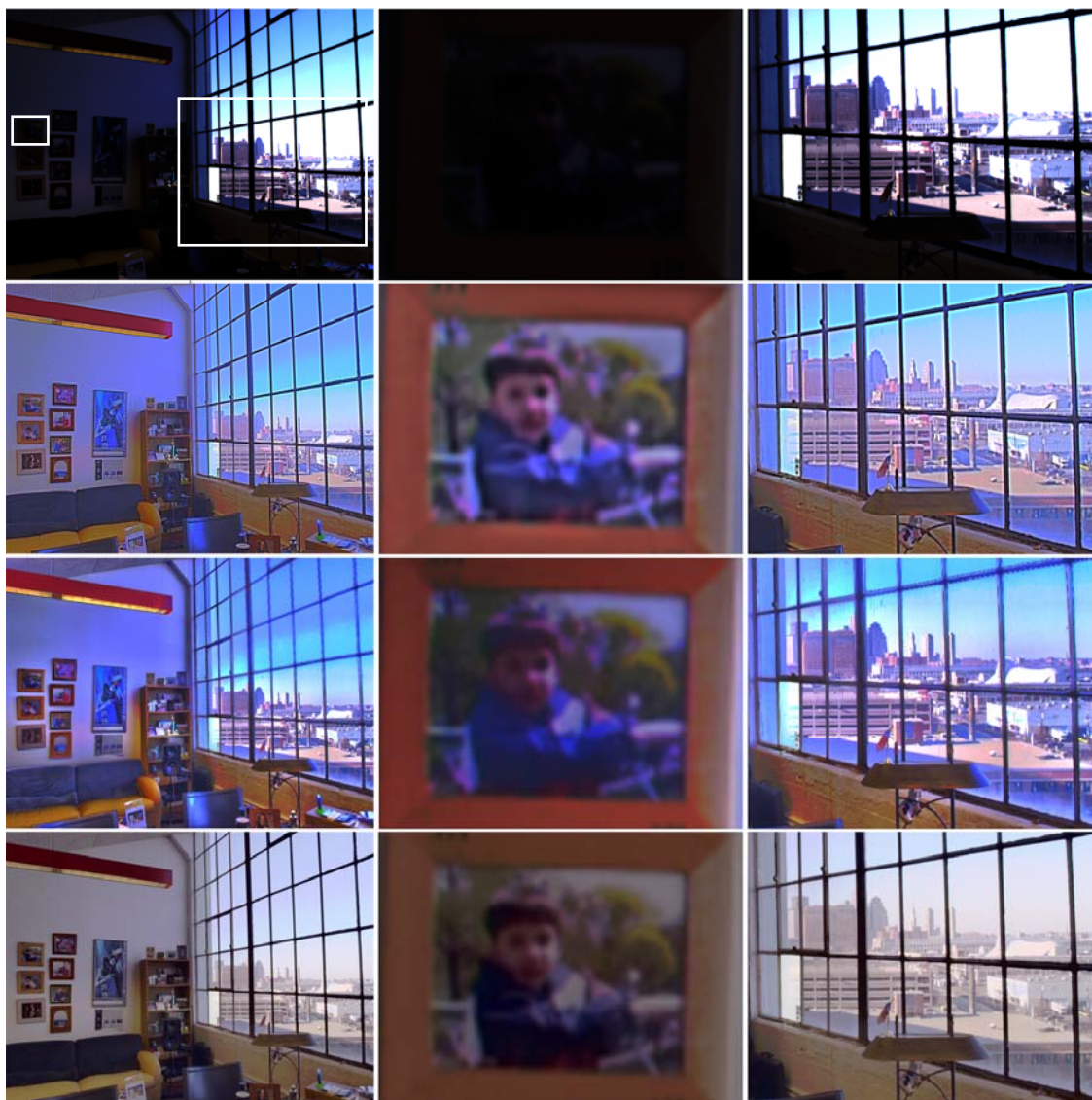
图 4-5 饱和参数 s 取不同值的结果

4.3.3 结果对比

在这一小节中，我们会将自己算法与其他研究人员提出的动态范围压缩算法进行对比。在这里，我们选取了三个经典的算法作为主要的比较对象，它们分别是：Durand的算法^[20]、Meylan的算法^[22]以及Pattanaik的算法^[11]。其中，Durand和Meylan的算法是局部的，Pattanaik的算法是全局的。Durand的算法中利用了双边滤波来处理图像，并将原图分为了基础层和细节层，整个算法并没有模拟人眼的视觉系统。Meylan和Pattanaik的算法都利用了人的视觉系统机制，其中Meylan的算法是一个Retinex模型。除了这三种算法之外，我们也与

这里我们对被比较的三种算法的处理结果来源做一个说明，Durand和Meylan的结果都是通过作者官方提供的程序代码进行运算得到的，他们的算法中都有非常多的可调节参数，我们一般采用作者的推荐值或者程序的默认值。Pattanaik的处

理结果是来自pfstools (http://pfstools.sourceforge.net/tmo_gallery/)。



(a) 处理结果的全景和细节图



(b) 原场景在不同曝光条件下的图像

图4-6 图像“office”处理结果对比。(a) 从上至下分别是原图、我们的结果、Durand的结果、Meylan的结果。(b) 原场景在不同曝光条件下的图像。

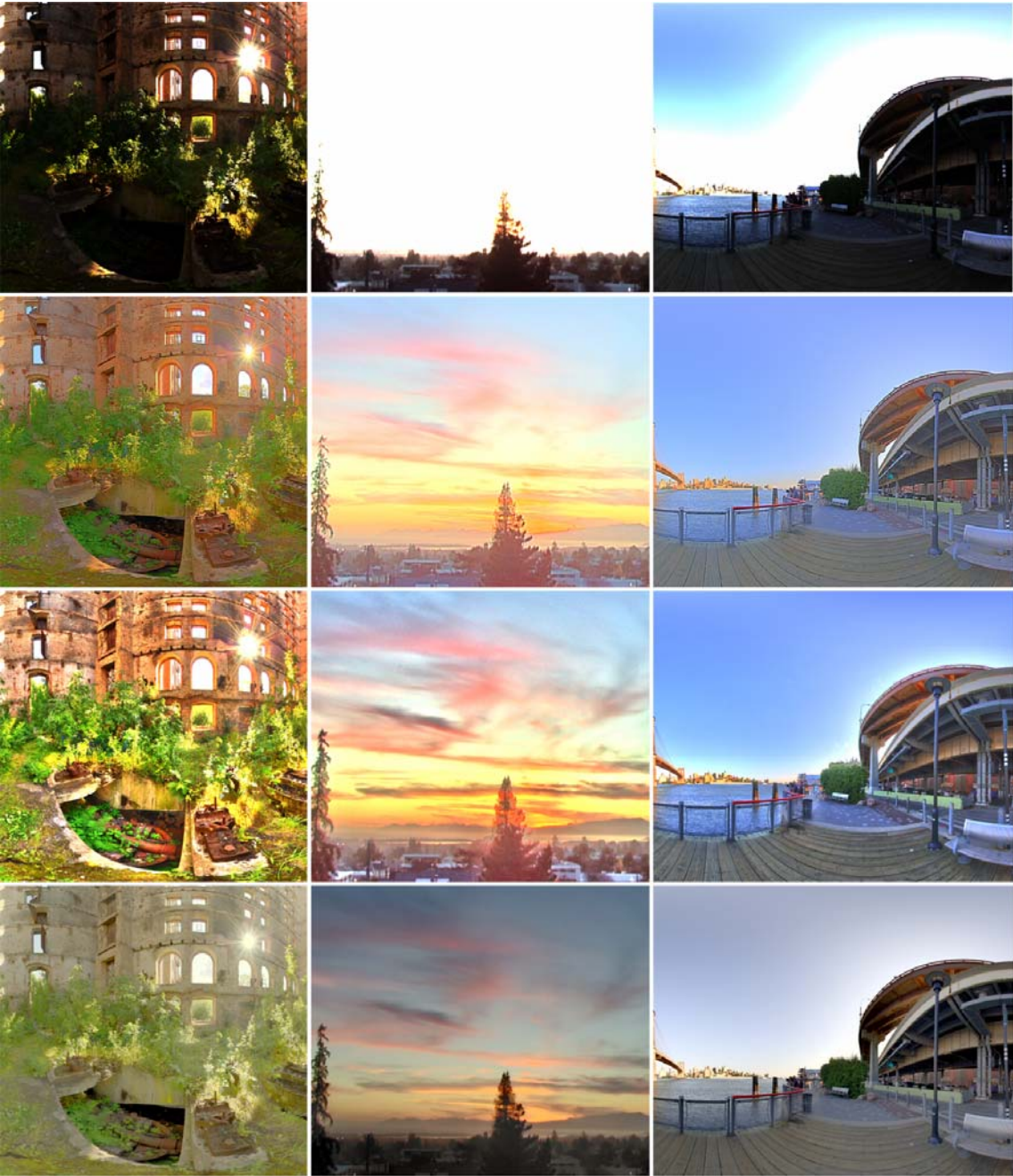
在做结果比较的时候，我们先和两个局部算法进行对比，然后再与 Pattanaik 的全局算法进行对比。在色调映射领域的研究中，对于处理结果的评价大多数研究人员仍以主观判断为主，因为人眼即使在没有参照标准的情况下，也可以辨别出图像质量的高低。因此我们选用了 8 个有差异化的场景进行比较，希望可以给出更加全面和客观的结果对比展示。

图4-6展示了我们与Durand和Meylan结果的对比，选取的是一张室内的HDR图像“office”。图4-6(a)共有4排组合而成，从上至下分别是原图、我们的结果、Durand的结果、Meylan的结果。图4-6(b)是原场景在不同曝光条件下得到的图像。为了比较对细节的处理情况，我们从图中选取了两个细节区域进行放大，也就是原图中用个白色矩形框圈出的部分。放大后的细节如图4-6(a)的第2列和第3列所示，分别为暗区和亮区的细节图。从全景图的处理结果来看，三种算法都能够在一定程度上压缩图像的动态范围，但是Meylan的算法存在比较严重的色偏现象。我们可以从图4-6(b)在不同曝光下的图像中看出，室内的墙是微微发蓝的，在我们和Durand的结果中墙面的颜色和真实情况基本一致，而Meylan处理后的墙面呈现一种微黄的感觉。对于细节处理的效果，我们和Meylan的算法能够展示出更多的暗区细节，Durand的算法相对较弱，亮度较低。而对亮区的处理，Meylan算法的压缩力度不够，难以看清细节。我们的算法处理后可以很清楚的看出图中远处的建筑，细节显示也比Durand的结果清晰。

为了更好的展示算法的成果，我们又选取了三幅室外图（如图4-7）以及两幅室内图（如图4-8）进行结果对比。在图4-7中从上至下分别是原图、我们的结果、Durand的结果、Meylan的结果。从对比结果中可以看出，我们的算法在压缩过亮区域时比其他两个算法表现更好，例如场景一中的太阳以及场景三中过度曝光的区域。Durand的算法在视觉感知（看上去的视觉效果）上优于我们的算法，但是有时候会出现信息丢失的情况，比如Durand处理后场景一中的太阳、云彩以及草丛都出现了信息丢失。

在图 4-8(a)中，最左边的四幅图从上至下从左至右分别是原图、我们的结果、Durand的结果、Meylan的结果，右边的三幅图从上至下分别是我们、Durand以及Meylan结果的细节放大图。虽然我们没有图“实验室”在正常曝光下的颜色，但是通过三种算法的对比依然可以发现，Meylan算法的色偏比较严重，从右侧的放大图可以看出，黄色的灯直接被处理成了白色。同时，Durand的算法在处理边界时出现了变形，比如在右侧的放大图中，Durand处理后的方形灯的边界出现了

波浪和圆弧角。图 4-8 (b) 是对图“厨房”的处理，从左至右从上至下依次为原图、我们的结果、Durand 的结果、Meylan 的结果。



(a) 场景一 (b) 场景二 (c) 场景三

图 4-7 室外场景处理结果对比。从上至下分别是原图、我们的结果、Durand 的结果、Meylan 的结果。



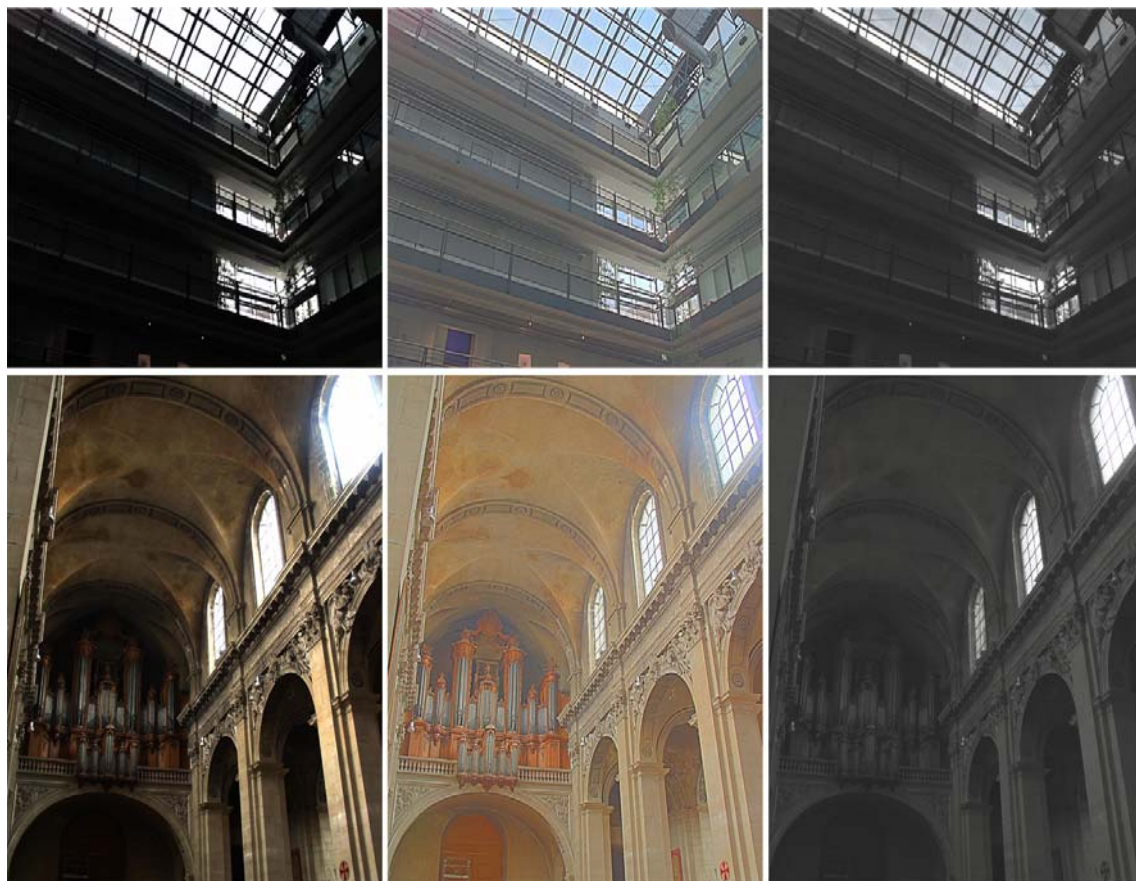
(a) 实验室



(b) 厨房

图 4-8 室内场景处理结果对比。(a) 左边四幅图从上至下从左至右分别是原图、我们的结果、Durand 的结果、Meylan 的结果，右边三幅图从上至下分别是我们、Durand 以及 Meylan 结果的细节放大图。(b) 从上至下从左至右分别是原图、我们、Durand 以及 Meylan 的结果。

图 4-9 是我们与 Pattanaik 提出的全局算法进行的对比, 从结果中我们可以看出, 全局算法无论是在细节还是在整体的动态范围压缩上都不如局部算法效果好。同时, Pattanaik 的结果图饱和度极低, 几乎和灰度图一样。



(a) 原图

(b) 我们的结果

(c) Pattanaik 的结果

图 4-9 与 Pattanaik 的全局算法结果对比

无参考 (no reference) 图像的质量评价在图像分析领域一直是个难题, 因为需要在不参考“原始”的高质量图的情况下, 用一个模型来对失真图像进行评价是非常难的。模糊检测累积概率法 (Cumulative Probability of Blur Detection, CPBD)^[59]是无参考评价中公认比较优越的方法, 其主要是针对图像的模糊度进行评价, 同时也可认为是对图像的锐利程度做评价。CPBD 是基于边缘检测的方法, 该方法将图像分块并将其分为边缘块和平滑块, 根据边缘块的亮度、宽度结合可觉察模糊 (Just noticeable blur, JBN) 得到每个点的模糊概率, 最后统计累加模糊概率得出模糊度。这里我们采用 CPBD 来评估我们、Durand 以及 Meylan 的处理结果, 没有与 Pattanaik 结果进行量化比较的原因在于没有相关的代码, 难以找到和其他算法处理相同图像的结果, 同时 Pattanaik 是全局算法, 效果比较差, 我们的主要

比较对象仍为局部算法。

表格 4-1 是我们、Durand 和 Meylan 结果的 CPBD 评价分数，表格中给出了 6 幅图像的评价分数，从左至右图像的顺序和之前图 4-6 到图 4-8 中图像出现的顺序一致。CPBD 得分越高表示图像边缘保护的更好，图像锐化程度更高。从表中可以看出，我们的算法评分中有 5 幅图都是最高分（除了图 4-7（b）），由此可见我们的算法对图像边缘保护更好，细节更清晰。

表 4-1 不同算法的 CPBD 评价分数

| 图像 | Office | 图 4-7(a) | 图 4-7(b) | 图 4-7(c) | 实验室 | 厨房 |
|--------|--------|----------|----------|----------|-------|-------|
| 我们的算法 | 0.657 | 0.703 | 0.652 | 0.727 | 0.283 | 0.741 |
| Durand | 0.534 | 0.651 | 0.700 | 0.642 | 0.210 | 0.560 |
| Meylan | 0.409 | 0.652 | 0.485 | 0.633 | 0.245 | 0.555 |

为了更好的展示算法的定量分析评价，我们采用了和文献[60]相同的评价方式，通过计算处理后图像的信息熵（entropy）来评价图像所含信息的丰富程度，得分越高表示图像所含信息越丰富。在此次定量分析中，我们新选取了 5 个场景进行处理，分别是 Napa Valley、Hotel Room、Atrium Morning、Atrium Night 和 Memorial。其中 Napa Valley 和 Hotel Room 的处理结果如图 4-10 和图 4-11 所示，其他处理结果考虑到篇幅问题不再放入文中。在此次定性评价中，我们与其他 8 种算法做了比较，分别是 Ashikmin^[61]、Drago^[7]、Durand^[19]、Fatta^[16]l、Reinhard^[18]、Retinex^[62]、Tumblin^[5]以及 Ward^[10]。这些算法的结果均来自 Max Planck Institut Informatik (MPII)(<http://resources.mpi-inf.mpg.de/tmo/NewExperiment/TmoOverview.html>)。

表 4-2 MPII 上算法的 entropy 评估

| | Napa Valley | Hotel Room | Atrium Morning | Atrium Night | Memorial |
|----------|-------------|------------|----------------|--------------|----------|
| 我们的算法 | 7.941 | 7.267 | 7.830 | 8.309 | 8.030 |
| Ashikmin | 7.416 | 6.966 | 7.778 | 7.972 | 7.769 |
| Drago | 7.381 | 6.756 | 7.421 | 7.907 | 7.753 |
| Durand | 7.182 | 6.695 | 7.046 | 7.424 | 7.291 |
| Fattal | 7.600 | 7.445 | 8.031 | 7.455 | 8.504 |
| Reinhard | 7.232 | 6.817 | 7.375 | 7.860 | 7.837 |
| Retinex | 7.200 | 7.169 | 7.356 | 8.330 | 8.187 |
| Tumblin | 7.016 | 6.391 | 6.995 | 7.491 | 7.668 |
| Ward | 7.416 | 6.694 | 7.251 | 7.915 | 7.783 |

表格 4-2 展示了以上 8 种算法以及我们的算法对 5 个场景处理后所求得的 entropy 值。除了 Memorial 场景我们的得分排在第三名外，其余场景我们的得分均在前两名。由此可见，我们算法处理后的图像所含信息丰富，对图像的动态范围压缩较好。



图 4-10 图像“Napa Valley”的结果对比

通过主观比较和定量分析我们在此对结果的对比做一个总结。首先，局部算法无论是在整体亮度压缩还是局部细节保留上都比全局算法更好，且 Pattanaik 的算法处理后图像的颜色饱和度过低。其次，对于局部算法容易出现伪影的情况，我们、Durand 以及 Meylan 的算法都在这方面表现较好，处理结果中没有明显的伪影出现。Durand 的结果能够给人较好的视觉感知，但是针对某些图像会出现信息丢失和边界变形，如图 4-7 (a) 和图 4-8 (a)。Meylan 的算法有明显的色偏问题，



图 4-11 图像“Hotel Room”的结果对比

这几乎体现在所有的结果图中，而当其所处理的图像有明显的过暗区和过亮区时（如图 4-6），对于亮区的压缩力度不足，呈现出的信息层次不分明，能够显示出来的细节较少，不如另外两个算法。我们的算法能够很好的压缩图像的整体动态范围，不论是过暗区还是过亮区处理后都可以显示出隐藏的信息。定量评价的结果表明，我们的算法能够很好的处理细节，处理后的结果边界清晰，无伪影出现，在 CPBD 的评价得分最高。处理结果的 entropy 分数也表明我们的结果所含信息丰富，能够有效地压缩图像的动态范围。颜色方面，我们处理后的图像无明显的色偏，人眼在观看上的感知和真实场景相近。

4.4 本章小结

在本章中，我们提出了一个全新的高动态图像压缩算法，旨在通过模拟人眼

视觉的适应机制来压缩图像的动态范围。我们在本章中首先介绍了算法的大体框架，然后对于算法中具体每一步的计算模型进行了详细的介绍，最后对算法的结果进行了分析和比较，包括模型中参数的作用影响以及我们与其他算法结果的对比。通过以上内容的介绍，可以看出我们的算法是一个静态的、局部的、自适应的模型，能够很好的压缩图像的动态范围，保留图像的边缘，展现更多的细节信息，处理后的颜色贴近现实无色偏。

第五章 总结与展望

本文的工作已基本介绍完毕，我们在此对全文的内容做一个总结，并分析和讨论我们的算法存在的一些问题和对今后改进完善的一些想法。

5.1 总结

相较于一般的图像，HDR图所包含的场景信息更加丰富，因此在各领域中的应用越来越广泛。然而我们现在主流的显示成像设备都是低动态的，因此在显示HDR图像时无法完整的展示其中的信息。为了解决这个矛盾，我们基于人眼视觉的适应机制提出了两个高动态图像增强算法，旨在保留原图信息的同时更有效地压缩图像的动态范围。我们的算法是以人眼的生理机制、视网膜上感光细胞的明暗适应机制以及已有的心理物理学实验为基础而构建的，这些支撑我们算法的理论和模型都在第二章中进行了介绍。虽然我们提出的两个算法都是基于视觉的适应机制，都是为了压缩图像的动态范围，但是它们之间还是有很大的差异。

我们提出的第一个算法采用了色调映射（TOM）领域较为通用的算法框架，是由Tumblin在1993年提出的。我们的算法将图像的输入端和输出端看做两个场景，并假设这两个场景引起的人的视觉响应是相同的。通过一个正向模型，可以求解出输入场景引起的视觉响应，在此我们考虑了视觉适应过程中感光色素的漂白机制。然后通过一个逆向模型，求出响应值对应的输出端的图像。我们在Tumblin的算法框架上，还增加了一个色适应模型，因为在之前的正逆向变换中都未能考虑颜色的适应，导致处理后的图像饱和度较低。色适应变换考虑了视觉的颜色适应，通过将RGB颜色空间转换到LMS空间，并在LMS空间对处理前后的图像进行匹配来保证输出后的图像在颜色感知上更贴近原场景。算法的具体流程和处理结果我们在第三章中进行了详细的介绍和展示，结果表明我们的算法可以有效地压缩图像的动态范围，且边界和细节清晰，无光晕的出现，色适应变换后的结果颜色贴近真实场景，视觉感知更好。

在第二个算法中，我们创新性地提出了一个新的算法框架，旨在更真实的模拟视觉适应机制。之前的算法由于采用了逆向模型，明显不符合视觉适应生理机制。我们新提出的算法框架完全参考了视觉适应的过程，并搭建了各部分的模型，主要包括了感光细胞的亮度模型、感光细胞的响应模型、神经节细胞的感受野模型，此外我们还考虑了视杆和视锥信息的融合以及颜色变换等。算法的流程和结果分析比较我们在第四章中有详细的介绍。实验的结果证明了我们的算法能够有

效的压缩图像的动态范围，处理后图像细节清晰无色偏。在与其他作者的结果对比时可以看出，我们的算法无论是图中的过暗区还是过亮区，都能够在处理后清楚的看出细节信息，而与我们进行对比的算法有时无法有效的压缩过亮区域。最后我们还采用了CPBD评价方式来对结果进行打分，最终得分也在一定程度上显示了我们的算法在细节上优于其他对比算法。

5.2 展望

虽然我们提出的两个算法都能够取得较好的结果，但是在具体的模型上还是有可以再改进的地方。目前我们的算法都是局部的算法，相对于全局算法而言，局部算法更贴近人眼在对局部场景进行适应时的特性，但是我们也不能忽略整体环境对局部适应的影响。因此，为了使模型更加完善，在求解感光细胞的响应时，我们可以考虑在其中加入一个全局变量，以此来代替整体环境对局部响应的影响。同时，我们的算法是静态的模型，只能用于处理单幅图像，为了增加算法的适用场景（比如自动驾驶时对周边环境的实时处理），我们将会考虑在其中增加时间参数，将其用于处理视频文件等，以此来拓展算法的应用范围。

致 谢

三年的研究生学习生涯即将过去，在此期间有很多老师，同学，朋友，家人都给予了我学习和生活的帮助。正是有了他们的支持才让我的研究生生活有了特殊的意义也有了更多的收获。

在这里首先要感谢我的导师李永杰教授，感谢李老师在这三年期间对我科研上的帮助和教导，带领我在计算机视觉的研究领域中学习和前行。帮助我寻找研究的方向，在我研究遇到困难时给我建议和指引。同时，李老师在工作中负责的态度，学术研究中严谨的作风和刻苦专研的精神也将继续影响着我今后的学习和工作。另外，我还要感谢李朝义院士在我研究过程中给予的关注和研究方向上的指导，还有教研室组会上颜红梅老师和王玲老师给我研究报告的提出的宝贵建议。

这里要特别感谢杨开富师兄，研究中遇到的很多问题都是在他的帮助和分析讨论中找到解决的方法，同时杨师兄在编程方面也给我提供了很多帮助。感谢张显石师兄和李惠在学习资料和图像上的分享，以及在讨论中的收获。感谢 301 教研室的所有同学，因为有了你们我的研究生生活更加丰富多彩。感谢张琼副书记和卢竞老师以及学院的所有老师，在我的生活和课程学习中给予的帮助。

最后，我要感谢我的家人，是他们在我的背后默默的支持着我的学习和生活，给予我关爱，给了我努力向前的动力。

参考文献

- [1] K. Devlin. A Review of Tone Reproduction Techniques. Department of Computer Science University of Bristol[J], 2002.
- [2] G. Ward. A contrast-based scalefactor for luminance display. Graphics gems IV[J], 1994: 415-421.
- [3] G. W. Larson, H. Rushmeier, C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. Visualization and Computer Graphics, IEEE Transactions on[J], 1997, 3(4): 291-306.
- [4] C. Schlick, Quantization techniques for visualization of high dynamic range pictures, in *Photorealistic Rendering Techniques*: Springer, 1995, 7-20.
- [5] J. Tumblin, J. K. Hodgins, B. K. Guenter. Two methods for display of high contrast images. ACM Transactions on Graphics (TOG)[J], 1999, 18(1): 56-94.
- [6] J. Tumblin, H. Rushmeier. Tone reproduction for realistic images. Computer Graphics & Applications IEEE[J], 1993, 13(6): 42-48.
- [7] F. Drago, K. Myszkowski, T. Annen, *et al.* Adaptive logarithmic mapping for displaying high contrast scenes[C], Computer Graphics Forum, 2003. 419-426.
- [8] J. C. Stevens, S. S. Stevens. Brightness function: effects of adaptation. Journal of the Optical Society of America[J], 1963, 53(53): 375-85.
- [9] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, *et al.* A model of visual adaptation for realistic image synthesis[C], Conference on Computer Graphics & Interactive Techniques, 1996. 249-258.
- [10] G. Ward. A Contrast-based Scale Factor for Luminance Display. Graphics Gems IV[J], 1994.
- [11] S. N. Pattanaik, J. Tumblin. Time-Dependent Visual Adaptation For Fast Realistic Image Display[C], Conference on Computer Graphics, 2000. 47-54.
- [12] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, *et al.* A multiscale model of adaptation and spatial vision for realistic image display. Proc.acm Siggraph'98 Acm Trans.graph[J], 1998: 287--298.
- [13] K. Chiu, M. Herf, P. Shirley, *et al.* Spatially nonuniform scaling functions for high contrast images[C], Graphics Interface, 1993. 245-245.
- [14] D. J. Jobson, Z.-u. Rahman, G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. Image Processing, IEEE Transactions on[J], 1997, 6(7): 965-976.
- [15] B. M. Oh, M. Chen, J. Dorsey, *et al.* Image-based modeling and photo editing[C], Proceedings of the 28th annual conference on Computer graphics and interactive techniques, 2001. 433-442.

- [16] R. Fattal, D. Lischinski, M. Werman. Gradient domain high dynamic range compression[C], ACM Transactions on Graphics (TOG), 2002. 249-256.
- [17] S. Pattanaik, H. Yee. Adaptive gain control for high dynamic range image display[C], Proceedings of the 18th spring conference on Computer graphics, 2002. 83-87.
- [18] E. Reinhard, M. Stark, P. Shirley, *et al.* Photographic tone reproduction for digital images[C], ACM Transactions on Graphics (TOG), 2002. 267-276.
- [19] F. Durand, J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. Acm Transactions on Graphics[J], 2002, 21(3): 257-266.
- [20] S. Paris, F. Durand, A fast approximation of the bilateral filter using a signal processing approach, in *Computer Vision—ECCV 2006*: Springer, 2006, 568-580.
- [21] P. Ledda, L. P. Santos, A. Chalmers. A local model of eye adaptation for high dynamic range images[C], IN IN PROCEEDINGS OF ACM AFRIGRAPH '04, 2004. 151--160.
- [22] M. Laurence, S. Sabine. High dynamic range image rendering with a Retinex-based adaptive filter. Image Processing IEEE Transactions on[J], 2006, 15(9): 2820 - 2830.
- [23] 寿天德, 视觉信息处理的脑机制[M]: 中国科学技术大学出版社, 2010.
- [24] H. Kolb. How the Retina Works. American Scientist[J], 2003, 91(1): 488-494.
- [25] E. R. Kandel, J. H. Schwartz, T. M. Jessell, Principles of neural science[M] vol. 4: McGraw-hill New York, 2000.
- [26] Osterberg, Gustav, Osterberg, *et al.* Topography of the layer of rods and cones in the human retina. Journal of the American Medical Association[J], 1937, 108(3): 232-232.
- [27] H. K. Hartline. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. American Journal of Physiology[J], 1938, 121(2): 400-415.
- [28] S. W. Kuffler. Discharge patterns and functional organization of mammalian retina. Journal of Neurophysiology[J], 1953, 16(1): 37-68.
- [29] R. W. Rodieck. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Research[J], 1966, 5(11): 583-601.
- [30] J. E. Dowling, The retina: an approachable part of the brain[M]: Harvard University Press, 1987.
- [31] S. Shlaer. THE RELATION BETWEEN VISUAL ACUITY AND ILLUMINATION. Journal of General Physiology[J], 1937, 21(2): 972.
- [32] E. Reinhard, W. Heidrich, P. Debevec, *et al.*, High dynamic range imaging: acquisition, display, and image-based lighting[M]: Morgan Kaufmann, 2010.
- [33] K. I. Naka, W. A. H. Rushton. S-potentials from colour units in the retina of fish (Cyprinidae). Journal of Physiology[J], 1966, 185(3): 536-55.

- [34] E. H. Adelson. Saturation and adaptation in the rod system. *Vision Research*[J], 1982, 22(10): 1299-312.
- [35] J. E. Dowling, *The retina : an approachable part of the brain*[M]: Belknap Press of Harvard University Press, 1987.
- [36] I. Iroto, M. Hirano, H. Tomita. Electromyographic investigation of human vocal cord paralysis. *Annals of Otolaryngology & Laryngology*[J], 1968, 77(2): 296-304.
- [37] D. A. Baylor, M. G. F. Fuortes. Electrical responses of single cones in the retina of the turtle. *Journal of Physiology*[J], 1970, 207(1): 77-92.
- [38] J. E. Dowling, H. Ripps, . Adaptation in skate photoreceptors. *Journal of General Physiology*[J], 1972, 60(6): 60--698.
- [39] R. D. Penn, W. A. Hagins. Kinetics of the Photocurrent of Retinal Rods. *Biophysical Journal*[J], 1972, 12(8): 1073-94.
- [40] R. M. Boynton, D. N. Whitten. Visual adaptation in monkey cones: recordings of late receptor potentials. *Science*[J], 1970, 170(3965): 1423-1426.
- [41] R. A. Normann, F. S. Werblin. Control of retinal sensitivity I. Light and dark adaptation of vertebrate rods and cones. *The Journal of general physiology*[J], 1974, 63(1): 37-61.
- [42] H. J. Dartnall, J. K. Bowmaker, J. D. Mollon. Human Visual Pigments: Microspectrophotometric Results from the Eyes of Seven Persons. *Proceedings of the Royal Society of London*[J], 1983, 220(1218): 115-30.
- [43] W. S. Stiles, J. M. Burch. N.P.L. Colour-matching Investigation: Final Report (1958). *Optica Acta*[J], 2010, January 1959(1): 1-26.
- [44] Kravetz, Alan. A Field Guide to Digital Color. *Campmadron Org*[J], 2013, 31(31): 436-437.
- [45] D. L. Macadam, *Chromatic Adaptation*[M]: Springer Berlin Heidelberg, 1985.
- [46] N. Katoh, K. Nakabayashi. Applying mixed adaptation to various chromatic adaptation transformation (CAT) models[C], *PICS*, 2001. 299-305.
- [47] E. Suman, ., S. Singh, ., S. Kotian, . Color Ratios and Chromatic Adaptation[C], *Conference on Colour in Graphics*, 2002. 7-10(4).
- [48] M. D. Fairchild. Color Appearance Models. *Journal of Electronic Imaging*[J], 2014, 23(3): 97-100(4).
- [49] R. W. G. Hunt, A. L. King. The Reproduction of Colour. *American Journal of Physics*[J], 1969, 37(1): 844.
- [50] W. A. H. Rushton. KINETICS OF CONE PIGMENTS MEASURED OBJECTIVELY ON THE LIVING HUMAN FOVEA. *Annals of the New York Academy of Sciences*[J], 1958, 74(2): 291-304.

- [51] M. Hollins, M. Alpern. Dark adaptation and visual pigment regeneration in human cones. *Journal of General Physiology*[J], 1973, 62(4): 430-47.
- [52] S. Ferradans, M. Bertalmio, E. Provenzi, *et al.* An analysis of visual adaptation and contrast perception for tone mapping. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*[J], 2011, 33(10): 2002-2012.
- [53] Z. Mai, H. Mansour, R. Mantiuk, *et al.* Optimizing a tone curve for backward-compatible high dynamic range image and video compression. *Image Processing, IEEE Transactions on*[J], 2011, 20(6): 1558-1571.
- [54] R. Mantiuk, S. Daly, L. Kerofsky. Display adaptive tone mapping. *ACM Transactions on Graphics (TOG)*[J], 2008, 27(3): 68.
- [55] L. Spillmann, J. S. Werner, *Visual perception: The neurophysiological foundations*[M]: Elsevier, 2012.
- [56] G. Wyszecki, W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae. Journal of Modern Optics*[J], 2000, 15(2): 197.
- [57] Z. Xie, T. G. Stockham. Toward the unification of three visual laws and two visual models in brightness perception. *IEEE Transactions on Systems Man & Cybernetics*[J], 1989, 19(2): 379-387.
- [58] R. Hall, *Illumination and Color in Computer Generated Imagery*[M]: Springer New York, 1989.
- [59] N. D. Narvekar, L. J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*[J], 2011, 20(9): 2678-2683.
- [60] A. T. Celebi, R. Duvar, O. Urhan. Fuzzy fusion based high dynamic range imaging using adaptive histogram separation. *Consumer Electronics, IEEE Transactions on*[J], 2015, 61(1): 119-127.
- [61] M. Ashikhmin. A tone mapping algorithm for high contrast images[C], *Proceedings of the 13th Eurographics workshop on Rendering*, 2002. 145-156.
- [62] F. Drago, W. L. Martens, K. Myszkowski, *et al.* Design of a tone mapping operator for high-dynamic range images based upon psychophysical evaluation and preference mapping[C], *Electronic Imaging 2003*, 2003. 321-331.