

高动态范围图像和色阶映射算子

杨克虎¹ 姬靖¹ 郭建军² 郁文生¹

摘要 图像传感器动态响应范围的局限使其在捕捉高动态范围场景时力不从心, 为了捕捉高动态范围图像 (High dynamic range image, HDRI), 近年来出现了许多新型传感器和新方法, 本文将简要介绍这些研究进展; 同样由于动态响应范围的局限, 显示设备也不能胜任 HDRI 的显示, 必须利用色阶映射算子 (Tone mapping operator, TMO) 将图像的动态范围进行合理的压缩, TMO 最终决定了图像显示的质量, 本文将众多的 TMO 归纳为全局算子和局部算子并进行了详细论述。

关键词 高动态范围图像, 色阶映射算子, 人类视觉系统
中图分类号 TP751

High Dynamic Range Images and Tone Mapping Operator

YANG Ke-Hu¹ JI Jing¹ GUO Jian-Jun² YU Wen-Sheng¹

Abstract The existing digital image capture device is not competent for the task of recording the vast-varying luminance range in the real world. In this paper, we first introduced the developments of high dynamic range image (HDRI) capture technology and the HDRI encoding format. To display such an HDRI in the traditional display device (CRT, LCD or printer), the dynamic range mismatch between image and display must be overcome by a tone mapping operator (TMO). An ideal TMO must keep a consistent perceptual experience between the displayed image and the real world scene. In recent years, many TMOs have been proposed, we summarized them into two types (global and local) and gave a detailed explanation.

Key words High dynamic range image (HDRI), tone mapping operator (TMO), human visual system (HVS)

光学数字成像过程是将真实场景的光辐射值通过图像传感器 (CCD 或 CMOS) 转化为电信号, 并以数字图像的方式保存下来, 而图像显示的目的是通过显示媒质 (CRT 或 LCD 等) 重现一幅数字图像所描述的真实场景, 两者的终极目标是使用户获得与直接观察真实场景相同的感官体验。长期以来, 图像传感器的动态响应范围都很小 (一般 CCD 传感器的动态范围不超过 1 000 : 1), 但真实环境的亮度值动态变化范围非常广, 夜晚星光照射下场景的平均亮度大概为 0.001 cd/m², 而白天阳光照射下的场景则达到了 100 000 cd/m², 可见平常大部分场景的亮度范围都远远超过了图像传感器的动态范围。如果用传统的图像传感器去拍摄高动态范围场景, 得到的图像将会是令人失望的, 这些图像要么明亮部分曝光过度丢失了太多的细节, 要么阴暗部分曝光

不足而无法分辨。为了捕捉到高动态范围场景的每一个细节, 已有不少新型的传感器和新途径被提出, 本文第 1 节将着重介绍这些进展。

高动态范围图像 (High dynamic range image, HDRI) 通过显示设备显示时, 同样遇到了动态范围不匹配的问题。目前主流的 CRT 显示器所能产生的亮度范围大约是 1 cd/m² ~ 100 cd/m², 虽然动态范围很窄, 但是已经足以应付大多数低动态范围图像的显示要求, 并且效果也基本令人满意。而相对于高动态范围图像的亮度范围 (0.001 cd/m² ~ 100 000 cd/m²), 显示器所能产生的亮度范围显得过于狭窄, 如果将图像的动态范围线性压缩到显示器的响应范围来显示, 所得到的观赏效果跟原始场景相去甚远。为了解决真实场景和显示设备动态范围不匹配的矛盾, 近年来国外的许多学者提出了各种各样所谓的色阶映射算子 (Tone mapping operator, TMO), 这些内容将在本文的第 2 节详细介绍。

1 高动态范围图像

1.1 高动态范围图像的获取

高动态范围图像的获取一般有两种方式: 1) 用图像传感器捕获的方式; 2) 图像合成的方式。在计算机图形学领域, 整体光照度方法 (辐射度方法, 光线跟踪算法等) 大多计算的是场景的真实辐射值, 而不是最终的显示亮度值, 因而动态范围会非常宽广。

收稿日期 2008-02-21 收修改稿日期 2008-08-26
Received February 21, 2008; in revised form August 26, 2008
国家自然科学基金 (60572056, 60874010, 60334020, 60621001), 中国科学院海外杰出学者基金资助
Supported by National Natural Science Foundation of China (60572056, 60874010, 60334020, 60621001) and the Overseas Outstanding Young Researcher Foundation of Chinese Academy of Sciences

1. 中国科学院自动化研究所复杂系统与智能科学重点实验室 北京 100190 2. 宏易未来 (北京) 科技有限公司 北京 100085
1. Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. HPI Innovation (Beijing) Co., Ltd., Beijing 100085

DOI: 10.3724/SP.J.1004.2009.00113

下面主要讨论如何通过摄像手段来捕获高动态范围图像。

1.1.1 多次曝光图像序列

Mann^[1] 和 Debevec^[2] 意识到虽然目前的传感器动态范围有限, 但是通过对同一个场景使用不同的曝光量拍摄多幅图像, 高曝光图像采集场景中暗的部分, 低曝光图像采集场景中亮的部分, 再通过特定的算法就可以从这一系列图像中大致恢复出真实场景的动态范围. 在这个过程中, 首先要恢复底片或数码相机的响应曲线, 这是因为底片的光响应曲线是非线性的, 而数码相机的感光器件在动态范围内的光响应虽然是线性的, 但通常得到的图像却是传感器原始输出图像的非线性变换. 不过现在的大多数数码相机都能够提供原始像素格式 (RAW pixel) 的图像, Madden^[3] 就是通过 RAW pixel 图像直接恢复出了原始的高动态范围场景. 假定一幅图像 I_1 , 其中等亮度的细节能很好地分辨, 而明亮部分曝光过度呈现全白, 阴暗部分曝光不足无法分辨, 适当增加和减小曝光量再各拍一幅图像 I_2 和 I_3 . 对于 I_2 , I_1 中曝光不足的部分得到了明显的改善; 同样对于 I_3 , I_1 中曝光过度的部分细节也能够分辨. 而在传感器线性响应范围内, 图像像素点的亮度与曝光量和该点的场景辐射强度成线性关系, 对于同一场景点, 辐射强度是不变的, 那么根据曝光量综合 I_1 , I_2 和 I_3 的可分辨区域, 进行简单的线性运算就能恢复出整个场景的细节. 由于这种方法需要拍摄多幅图像, 成像速度很慢, 所以只适合于拍摄静态的场景.

1.1.2 光束分离

Aggarwal^[4] 在 2001 年提出了光束分离的方法, 并且研制出了原理性样机, 这种方法的原理和上述的多次曝光图像序列基本一致, 也是采用对同一场景的多个不同曝光的图像来恢复原始场景. 不过前者采集的过程有先有后, 是一个时间序列; 而后者由于采用了特定的光路, 使所有的图像能够一次性获得, 从而大大提高了成像速度, 可以用于动态场景的拍摄. 如图 1 所示, 使用了一个 6 面的分光棱镜将光线分成 6 束, 每个分光面对应一个传感器, 这样就可以通过设定每个传感器的曝光时间或者采用非均匀的分光方式得到一组曝光量各不相同的图像.

1.1.3 高动态范围图像传感器

常见的 8 bit 灰度图像的动态范围只有 48 db, 这一方面是因为传感器本身的动态范围低, 另一方面是为了降低成本, A/D 转换器的分辨力也只有 8 位. 近年来半导体技术的发展使图像传感器动态范围有了很大的提高, 比如 Dlasa 公司的高性能 CCD^[5], 线性动态响应范围达到了 72 db (4000 : 1), Fairchild 公司的科学级 CCD^[6] 的动态范围达到了

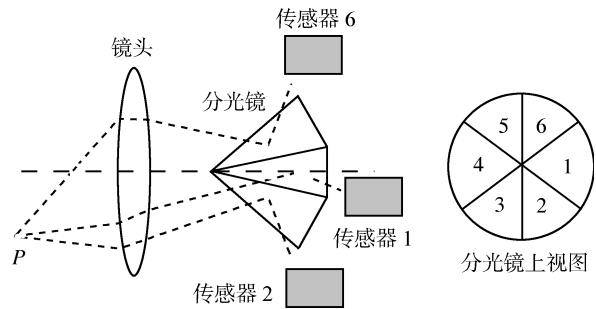


图 1 光束分离示意图

Fig. 1 Sketch of beam splitter

83 db (14000 : 1), 而 Hamamatsu 公司的 S10140/10141^[7] 系列, 标称的动态范围达到了 90 db (30000 : 1). 但是这仍然不足以捕捉真正的高动态范围图像, 于是一些新型的传感器应运而生.

人类视觉系统 (Human visual system, HVS) 有一整套完整高效的自动调节机制, 瞳孔会自动根据光线的强弱变大或变小以调节通过的光通量, 此外视网膜能根据投射到该区域的辐射强度自动调节视锥细胞的敏感度. 受到这一机理的启发, Chen 和 Ginosar 开发了一种 Adaptive SensitivityTM CCD 图像传感器^[8], 这种传感器每个像素的灵敏度都是可以控制的 (通过控制曝光时间), 并且研制出了 9×16 的试验性芯片.

Nayar 和 Mitsunaga 提出了一种空域变曝光像素 (Spatially varying pixel exposures, SVE)^[9] 的方法, 如图 2 所示的像素阵列中有灰度不同的四种方格 (e_0, e_1, e_2, e_3), 分别代表四种不同敏感度的像素, 越亮的表示敏感度越高 ($e_0 < e_1 < e_2 < e_3$). 当某个像素饱和时, 可能它邻近的某个像素没有饱和, 反之, 当某个像素输出为零时, 它邻近的某个像素可能输出为非零. 像素敏感度在空间域变化的同时对动态范围进行了采样, 这样就可以利用邻近像素的信息估算出当前像素的实际值, 从而得到真实场景的高动态范围图像.

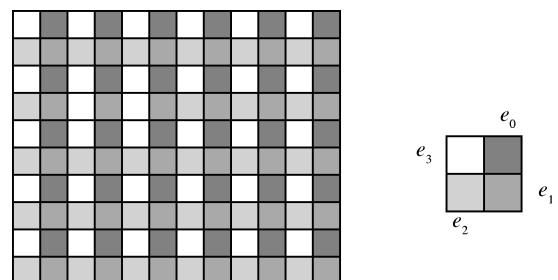


图 2 SVE 像素阵列

Fig. 2 Pixels array of SVE CCD

近年来, 数字图像传感器市场已经不再是 CCD 一枝独秀了, CMOS 传感器由于成本较低也被大量用在了对图像质量要求不高的摄像头、手机拍照等场合, 吸引了越来越多的半导体厂商投身到高性能 CMOS 图像传感器的开发中来. HDRC[®] 是 Omron 公司最新研发的动态范围高达 120 db (10⁶ : 1) 的 CMOS 图像传感器, 它甚至能够同时记录下发光白炽灯的灯丝和灯泡玻璃罩上的黑字. 有关 HDRC[®] 的更多信息可以访问相关的网站¹.

1.2 高动态范围图像的编码格式

高动态范围图像由于具有极广的灰阶和色域, 传统的 24-bit RGB 图像编码格式已经不足以描述如此之多的信息量, 简单地用浮点数来表示又需要很大的存储空间, 因此急需开发一种简单高效的编码方式. 不同的研究领域为了特定的应用需求采用了不同的格式, 如计算机图形学领域的 RGBE 和 OpenEXR 格式, 摄影领域的 RAW 格式, 医学成像领域的 DICOM 格式, 电影摄影学领域的 DPX 格式. 即使是同一种格式 (例如 RAW 格式), 不同厂家的标准也不太一样, 这样不利于信息的交流, 给使用者造成了很大的不便. 最具代表性的编码格式有如下几种^[10]:

- 1) Pixar 33-bit log-encoded TIFF;
- 2) RGBE and XYZE radiance 32-bit;
- 3) IEEE 96-bit TIFF & portable FloatMap;
- 4) 16-bit/sample TIFF;
- 5) LogLuv TIFF;
- 6) ILM 48-bit OpenEXR format.

2 色阶映射算子

首先给出亮度和视觉亮度的定义以及一些符号的含义, 如表 1 所示.

考虑到人眼对不同频率的光的敏感度不一样, 亮度由光谱辐射强度 $P(\lambda)$ 用光谱光效应函数加权后积分得到, 计算公式为

$$L = \int_{\lambda_1}^{\lambda_2} P(\lambda) \cdot V(\lambda) d\lambda \quad (1)$$

其中 $V(\lambda)$ 是光谱光效应函数, 是人眼对特定频率光的归一化敏感程度.

视觉亮度不是个物理量, 而是光刺激人眼产生的主观感觉, 它不能由仪器测出, 只能通过心理物理学实验来间接得到. 众所周知, 人眼在不同的环境下对同样强度的光刺激可能产生不同的主观感觉, 另外在室内和室外观察同一台显示器得到的视觉效果也不一样, 这说明视觉亮度还与环境光强有关, 而环

表 1 符号和定义

Table 1 Symbols and definitions

符号	定义
L_w	场景亮度值
L_d	显示设备亮度值
L_{wmin}	场景最小亮度值
L_{wmax}	场景最大亮度值
L_{dmin}	显示设备最小亮度值
L_{dmax}	显示设备最大亮度值
B_w	场景的视觉亮度
B_d	显示设备的视觉亮度
L_a	人眼适应水平
L_{wa}	场景的适应水平
L_{da}	显示设备的适应水平
(i, j)	图像像素坐标

境光强又在一定程度上决定了人眼的适应水平. 迄今为止, 有多种视觉亮度的近似数学模型被提出, 比如被广泛采用的 Stevens 模型^[11-12].

本文前面论述了高动态范围图像的获取和存储, 而只有将获得的图像通过显示媒质 (显示器, 打印机等) 表现出来, 才能给人直观的体验. 在这个过程中, 同样遇到了动态范围不匹配的问题. 现有的显示媒质的动态范围太小, 但是由于人眼的自动调节机制, 以及对绝对亮度的敏感性大大低于对相对亮度的敏感性, 因此就有可能在动态范围较低的显示设备上大致还原出真实的高动态范围场景, 而这需要有一个映射将场景的亮度值映射到显示设备的亮度值, 这就是我们要讨论的 TMO. 可以从数学上定义 TMO 如下

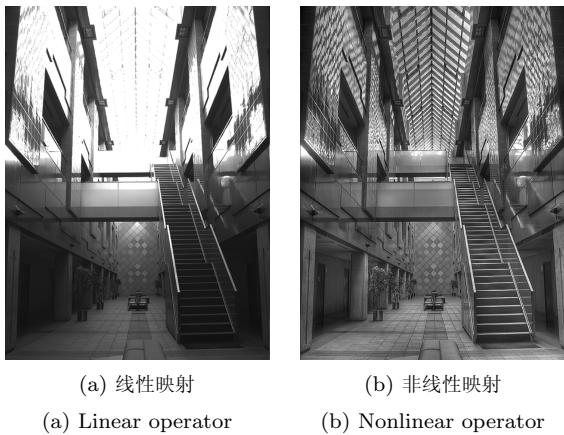
$$T: L_d = T(L_w) \quad L_d \in D, L_w \in W \quad (2)$$

式 (2) 中 T 就是 TMO, D 是显示设备的亮度值集合, W 是世界场景亮度值的集合, L_d 和 L_w 分别是显示设备和场景的亮度值.

一个最简单的方式是使用线性映射, 即将场景的最大和最小亮度分别映射到显示设备的最大和最小亮度, 中间亮度都按照这个关系线性映射到这个范围之内. 这种方法计算量小, 在场景的动态范围与显示设备相仿的时候一般能得到令人满意的效果. 但是对于高动态范围图像, 则会丢失很多高亮和阴暗的场景细节. 另外由于这种方式完全没有考虑到场景的绝对辐射强度, 有可能导致同一个场景在白天日光照射下和夜晚微弱灯光照射下分别映射得到相同的图像, 丢失了场景的整体亮度印象信息. 图 3(a) 和 3(b) (见下页) 分别是采用线性映射和某种 TMO 映射方式得到的显示图像, 可以明显地看出, 图 3(b) 中的图像无论是从细节还是从整体的视觉

¹<http://www.hdr.com/>

效果来看都比图 3(a) 有很大的提高. 早在 1984 年 Miller 等^[13] 就利用 Stevens^[11-12] 通过实验得到的心理物理学数据来匹配真实场景和显示图像的视觉亮度, 随后 1985 年 Upstill 在他的博士论文^[14] 中再次强调了利用一个显性的感知模型来实现色阶重建的必要性. 但是真正将色阶重建问题规范正式提出的则是 Tumblin 和 Rushmeier 于 1993 年发表的论文^[15], 随后许多学者在这方面做了大量的工作, 提出了各种不同的 TMO, 所有这些 TMO 都是基于图 4 这样的基本框架, 即使真实场景的观察者和显示设备的观察者都能获得相同的视觉感受.



(a) 线性映射 (b) 非线性映射
(a) Linear operator (b) Nonlinear operator

图 3 线性映射和非线性映射的效果比较

Fig. 3 Comparing the mapping results of linear and nonlinear operators

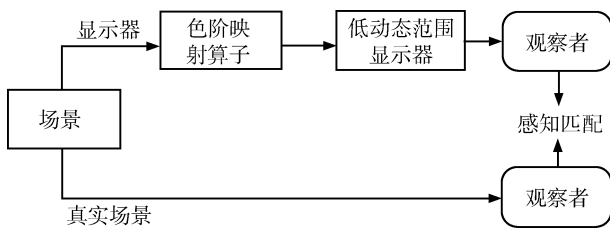


图 4 色阶映射算子的一般模型

Fig. 4 A common model of tone mapping operator

在图 4 中, 最终目的是获得感知匹配, 这个感知包括哪些内容? Ferwerda 等根据心理物理学的实验数据提出了人类视觉系统的计算模型^[16], 认为除了亮度、可见性、色彩恒常性等基本信息外, 还应包括明视觉和暗视觉、亮度分辨力、视觉敏锐度(即通常所说的视力, 也称为对空间频率的敏感性)等视觉特性. 对于序列图像的显示, 还必须考虑到人眼的明适应和暗适应过程. 对于静态图像的显示, 所用的 TMO 大体上可分为两类: 1) 全局(Global) 算子或空间均匀(Spatially uniform) 算子; 2) 局部(Local) 算子或空间非均匀(Spatially non-uniform) 算子.

TMO 是高动态范围图像图形处理流程的最后

一个环节, 它直接决定了观察者的主观体验, 同时它也是一个高度交叉的研究领域, 涵盖了图像处理、视觉生理学、色彩学、心理物理学等多个领域, 吸引了众多的学者、研究机构 (Microsoft Research Asia, MPI²等) 和图形硬件厂商 (Nvidia, ATI, Intel 等) 的关注. 近年来国际上关于这方面的研究越来越多, 很多高质量的工作在 ACM 的 SIGGRAPH 年会和 *Transactions on Graphic* 上发表, 特别是在 2002 年的 SIGGRAPH 年会上同时出现了三篇该领域的论文. 在 IEEE 和 SPIE 的各种期刊和会议上也有很多这方面的高质量论文发表, 但是在国内还是很少见到关于这方面的研究工作.

2.1 全局算子

对于式 (2) 中定义的非线性映射 T , 如果 T 只依赖于当前像素的亮度值, 而与当前像素的坐标无关, 那么这样的算子就是全局算子. 相同场景亮度值经 T 映射后相等, 不会因为二者所在的位置不同而不同.

2.1.1 亮度匹配

Tumblin 和 Rushmeier^[15] 根据 Stevens 的亮度感知模型^[11-12], 着眼于保持图像的整体亮度印象, 匹配真实场景和显示图像的视觉亮度 $B_w = B_d$, 其中, $B_w = 10^{\beta_w} L_w^{\alpha_w}$, $B_d = 10^{\beta_d} L_d^{\alpha_d}$, 解得

$$L_d = L_w^{\frac{\alpha_w}{\alpha_d}} 10^{\frac{\beta_w - \beta_d}{\alpha_d}} \quad (3)$$

式中真实场景和显示设备的指数 α, β 都可以通过各自的适应水平 L_a 得出. 但是根据 Stevens 的模型, L_a 是关于场景 (或显示图像) 亮度值的函数, 对于真实场景, L_a 容易求得, 但对于显示图像, 由于亮度本身就是待求量, 所以无从知道 L_a . 为了解决这个问题, 可以近似认为显示图像的 L_a 是一个常数, 因为大多数显示设备的动态范围都很小, 所以 L_a 的变化不会太大, 具体计算公式为 $L_a = L_{\max} / \sqrt{C_{\max}}$, 其中 L_{\max} 是显示设备的最大亮度, C_{\max} 是显示设备的对比度. 由于 CRT 亮度和电压的非线性关系, 得到的图像在显示的时候还需通过 gamma 校正才能获得最终的控制电压值^[15]. 这种方法所依赖的模型本身是基于简单场景建立的, 计算量很小, 但是并不能很好地适用于高度复杂的场景. 关于 Stevens 的视觉亮度模型, 见式 (4) ~ (7).

$$B = 10^\beta \cdot L^\alpha \quad (4)$$

$$\alpha = 0.4 \lg(L_a) + 2.92 \quad (5)$$

$$\beta = -0.4(\lg(L_a))^2 - 2.584 \lg(L_a) + 2.0208 \quad (6)$$

$$L_a = 6.9183 \cdot 10^{E(\lg(L))} \quad (7)$$

²Max-Planck-Institut Informatik, Germany

其中, B , L 分别是视觉亮度和亮度, L_a 是人眼适应水平, E 是数学期望.

2.1.2 最小可辨差

由于人眼对环境亮度的适应性, 导致了人眼在一定的背景亮度下的最小可辨差 (Just noticeable difference, JND) 并不是恒定的, 而是背景亮度越低, 最小可辨差就越小, 二者之间存在一定的关系, Blackwell^[17] 根据试验心理学数据得出了如下数学模型

$$\Delta L(L_a) = 0.0594 \cdot (1.219 + L_a^{0.4})^{2.5} \quad (8)$$

即在一定的适应水平 L_a 下, 人眼的最小可辨差为 ΔL . 为了使真实场景中人眼所能分辨的亮度差在显示图像中也能分辨出来, Ward^[18] 认为真实场景和显示图像的亮度值应该保持一个恒定的比例, 即 $L_d = mL_w$. 如果 m 太小, L_w 被压缩到一个很小的范围, 就有可能导致原来可分辨的细节由于过度压缩而不可分辨; 如果 m 太大又可能使压缩后的动态范围超出显示设备的动态范围. 为了保证可分辨力的一致性, m 必须满足

$$\Delta L(L_{da}) = m \cdot \Delta L(L_{wa}) \quad (9)$$

其中, L_{da} 和 L_{wa} 分别是人眼对显示设备和真实场景的适应水平. 从而可以得到

$$m = \left(\frac{1.219 + L_{da}^{0.4}}{1.219 + L_{wa}^{0.4}} \right)^{2.5} \quad (10)$$

Ward 建议 L_{da} 取为显示设备的最大亮度值的一半. 而对于 L_{wa} , 在没有其他信息的情况下, Ward 简单地取其为场景的亮度平均值. 由于人类视觉系统的高度复杂性, 现有的任何模型都只能是在一定程度上去逼近真实的视觉响应特性. 随后也有其他不同的模型被采用, 如 Ferwerda 等^[16] 所采用的模型考虑了杆状和锥状视锥细胞不同的敏感特性, 比 Ward 所采用的模型更为精确.

2.1.3 直方图调整 (Histogram adjustment)

Ward 等提出了一种基于直方图调整技术的 TMO^[19], 这种方法与图像增强领域的直方图均衡化方法相类似, 都是调整图像的灰度级分布. 二者最主要的区别是, 图像增强的目的是将图像的对比度和可见性最大化, 但并不考虑最大化以后的主观感知体验是否和真实情况相符; 而 TMO 则是为了最大限度地保持主观感知体验的一致性, 是为了重现原始的对比度和可见性, 而不是将它最大化. Ward 认为所有采用单一适应水平的 TMO 存在两个不足之

处: 1) 高动态范围图像的直方图大多呈现一种比较稀疏的分布, 在很多灰度级上都只有极少数的像素分布, 而采用单一适应水平的映射方法并不能有效地降低这个稀疏程度, 也就不能充分利用显示设备的动态范围. 如果人眼对绝对亮度敏感的话, 这样做似乎并无不妥, 但实际上人眼对绝对亮度并不敏感, 对于亮和暗的区域, 人眼一般只能识别一个亮一个暗, 并不能给出一个确定的绝对亮度差值, 所以完全可以压缩直方图中稀疏的部分以降低图像的动态范围. 2) 视锥细胞的适应水平是随着观察内容的变化而改变的, 大约每 1° 立体角的视场范围内就有一个适应水平. 因此采用单一的适应水平并不能很好地模拟人类的视觉适应特性. 文献 [19] 首先由原始图像确定一幅包含了图像各个部分适应水平的下采样图像, 把整个视场划分为若干个 1° 的立体角, 对每一个立体角计算一个适应水平, 然后求出这幅下采样图像累积分布函数. 最初为了得到一个比较直观的结果, 使用了直接基于累积分布函数的直方图均衡化的方法, 但输出图像的直方图在稀疏部分被压缩的同时也放大了密集部分的对比度而产生了明显的失真. 为了解决这个问题, 考虑到线性映射算子能较好地适用于低动态范围图像, 文献 [19] 认为图像的任何区域经过映射后的对比度都不应该超过由线性算子映射后的对比度, 引入了所谓的“线性封顶 (Linear ceiling)”的方法, 输出图像的整体效果有了明显的改善. 进一步地, 根据人类视觉系统的对比度敏感特性 (最小可辨差), 文献 [19] 提出了另一种基于视觉感知特性的“封顶”方法. 随后考虑了更多的视觉特性因素, 如闪光的模拟、色彩的感知以及视觉敏锐度等, 并且将它们集成到同一个映射算子中. 从以上的讨论可以看出, 直方图调整的方法实际上已经不再是一个严格意义上的全局算子了, 它已经考虑到了空间位置移动的影响, 但是习惯上仍然将它当成一种全局算子.

全局算子是最早被提出用来解决场景和显示设备动态范围不匹配问题的, 虽然这种方法所采用的模型一般比较简单, 不能很好地模拟真实的视觉响应特性, 但是由于相对较低的计算复杂度, 在一些场合中还是得到了大量的应用. 目前, 还有一些其他的全局算子被提出来, 如 Scheel 等^[20] 提出的一种交互式的纹理贴图方法, Cohen 等^[21] 利用两幅单独的纹理贴图, 通过动态调节曝光水平并将两幅纹理贴图合成在一起的方法来显示图像. 随着对图像显示逼真度的要求越来越高, 全局算子在处理一些场景内容高度复杂或极高动态范围的图像时并不能给出令人满意的结果, 因而越来越多的学者更加深入地研究了人类视觉系统的生理特性, 在此基础上提出了众多的局部算子.

2.2 局部算子

局部算子和全局算子的最大区别在于后者只依赖于像素的亮度值, 而不关心像素在图像中的位置以及它邻近像素的影响; 前者则考虑到了这种影响, 使得算子在空间上是变化的, 可以把相同亮度不同位置的像素映射到不同的亮度.

2.2.1 非均匀比例函数

Chiu 等^[22] 在压缩一幅高动态范围的合成图像时, 发现不论采用哪种全局算子都不能得到令人满意的效果, 他意识到人类视觉系统本身就是空间非均匀的, 用局部算子描述这样的生理特性可能比较合适, 而且画家和摄影师也经常使用这种局部的方法来记录真实场景 (如在相纸的不同地方采用不同的曝光时间), 因而提出了如下的映射模型

$$L_d(i, j) = m(i, j)L_w(i, j) \quad (11)$$

式中 $m(i, j)$ 是像素 (i, j) 处的比例因子. 首先可以得出 $m(i, j)$ 的一些限制条件, 因为 $0 \leq L_d(i, j) \leq 1$, 所以 $0 \leq m(i, j) \leq 1/L_w(i, j)$. 根据 Land^[23] 和 Marr^[24] 的研究, 人眼会或多或少地忽略相对比较平缓的空间亮度变化, 这就意味着只要 $m(i, j)$ 有相对较小的梯度, 那么映射后的图像就会有相当逼真的显示效果. 另外一些视觉研究者指出, 人眼在图像某部分的适应水平与图像的低频成分有关系^[25], 所以很自然地联想到可以用低通滤波后的图像来代表这个适应水平, 从而可以得到局部算子

$$m(i, j) = \frac{1}{k \cdot L_{blur}(i, j)} \quad (12)$$

式中 L_{blur} 由 L_w 经过低通滤波得到, k 是比例系数. 这种方法的一个缺点是会产生对比度反转, 在高亮部分的周围生成黑色的环绕带, 严重影响了视觉效果, 作者认为产生这种情况的一个原因是没有考虑到人眼在高亮度刺激下会产生的浸润效应.

2.2.2 有理分式映射

为了构造一个普遍适用而且计算复杂度较低的颜色阶映射算子, Schlick^[26] 分析了计算机图形学领域中几种比较常用的算子: 线性伽玛映射、箝位伽玛映射、对数映射和指数函数映射. 虽然利用这些算子通常也能得到令人满意的结果, 但是没有一个好的方法来自动获得它们各自的参数, 而且这些方法都要进行指数或对数运算, 计算量较大. 为了克服这些缺点, Schlick 提出了有理分式映射

$$L_d = \frac{p \cdot L_w}{p \cdot L_w - L_w + L_{wmax}}, \quad p \in [1, \infty) \quad (13)$$

其中 p 是非线性参数, 决定了映射的非线性程度, 虽然这种映射和前述几种映射作用于图像的效果大同

小异, 但是由于避免了指数或对数运算, 计算复杂度大大降低, 而且还可以自动地获得控制参数 p , 这是前述几种方法所不具备的. 参数 p 的选取必须同时考虑到显示设备和人类视觉系统的非线性响应以及环境、观察距离等多方面因素的影响. 那么如何自动地获取参数 p ? Schlick 认为某些因素 (环境的亮度、观察距离等) 变化时, 改变的只是人眼所能分辨的最低亮度, 场景中的最低亮度值只能是映射到显示设备的最低亮度值, 即

$$L_d(L_{wmin}) = L_{dmin} = \frac{M}{N} \quad (14)$$

其中, M 是人眼所能觉察到的显示设备的最低灰度级, N 是显示设备的最高灰度级. 根据式 (13), 得

$$p = \frac{M \cdot L_{wmax} - M \cdot L_{wmin}}{N \cdot L_{wmin} - M \cdot L_{wmin}} \quad (15)$$

而 M, N 都是可以事先确定的, 所以参数 p 就可以随着输入图像的改变而自动调节了. 但是对同一幅输入图像 p 是固定的, 这样得到的算子实际上还是一个全局算子. 为了模拟人的视觉特性, 就必须使 p 随着位置的改变而改变. Schlick 根据各个像素点附近的适应水平 L_a 来确定 p , 使具有较小的 L_a 的像素看起来适当地亮一些, 而对于较大的 L_a 则降低一些亮度, 他简单地根据场景的中间亮度: $L_{wmid} = \sqrt{L_{wmax} \cdot L_{wmin}}$, 把整个亮度范围分为两部分, 得到

$$p' = p \cdot \left(1 - k + k \cdot \frac{L_a}{L_{wmid}} \right), \quad k \in [0, 1] \quad (16)$$

其中 k 代表了算子局部化的程度. 对于如何求解 L_a , Schlick 提出了低通滤波、图像分割等方法. 低通滤波的方法能产生比较好的结果, 但是与 Chiu^[22] 的方法一样会产生比较明显的失真; 而图像分割等方法得到的结果并不是很好, 甚至差于最初的全局算子.

2.2.3 摄影映射算子 (Photographic TMO)

Reinhard 等^[27] 认为虽然视觉感知模型是解决颜色阶映射问题的一个比较好的途径, 也能得到一些有效的算法, 但是现有模型的精度不高, 会引起两方面的问题: 首先是会引入一些假象, 如前面提到的黑色的环绕带; 其次是视觉体验不仅仅依赖于对比度和亮度等信息, 场景内容、观察条件等因素也很重要, 从而增加了模型的复杂度. “分区曝光显影系统 (Zone system)” 的概念首先在上世纪四十年代被提出, 后来经过 White^[28] 和 Adams^[29-31] 的发展而逐渐在摄影界流行起来. 这种拍照手法的主要目的是希望在拍摄的当时, 就能够预视相片洗出来的效果. 它将场景和显示设备的动态范围分别按照亮度倍增的顺序分成了若干个区域 (Zones), 其中显示设备的

动态范围 (Print zones) 分成了 11 个区域, 摄影师首先获取场景的中间亮度值, 这个亮度将会映射到显示设备的 Zone-5, 然后再获取场景的最高和最低亮度以决定该场景的动态范围. 如果没有超过 11 个 Zones, 那么就能将场景和显示设备的 Zone 一一对应; 如果超过了 11 个 Zones, 就需要使用遮光-增光 (Dodging-and-burning) 的方法来压缩动态范围. 基于这种摄影的手法, Reinhard 等开发了一种新的色阶映射算子. 首先, 利用对数平均值亮度代表场景的中间亮度值

$$\bar{L}_w = \exp\left(\frac{1}{N} \cdot \sum_{i,j} \ln(\delta + L_w(i,j))\right) \quad (17)$$

其中 δ 是很小的正数; 然后所有的像素按照此式线性缩放: $L(i,j) = (\alpha/\bar{L}_w) \cdot L_w(i,j)$, α 是 \bar{L}_w 被映射后的值, 对于一般的图像, $\alpha = 0.18$, 对于比较暗 (或亮) 的图像, α 可以变大到 $[0.36, 0.72]$ (或变小到 $[0.09, 0.045]$); 最后将 $L(i,j)$ 按照 $L_d(i,j) = L(i,j)/(1 + L(i,j))$ 映射到显示设备的 $L_d(i,j)$, 这样就保证了所有的 $L_d(i,j)$ 都在显示设备的范围 $[0, 1]$ 之内, 还可以引入 L_{white} , 将这个映射扩展为

$$L_d(i,j) = \frac{L(i,j) \cdot \left(1 + \frac{L(i,j)}{L_{white}^2}\right)}{1 + L(i,j)} \quad (18)$$

$L_d(i,j)$ 超过 1 的值都取为 1, 这样 L_{white} 为被映射到显示设备最大亮度的最小场景亮度值, 以产生一种 “burn out” 的效果. 可以看出这个色阶映射算子实际上是一种全局算子, 对于动态范围极高的图像还是会丢失一些细节信息, 必须用类似遮光-增光的方法适当压缩动态范围, 基本思想是对每一个像素根据其局部信息选择不同的 α . 具体做法是将图像分成若干亮度相近的区域, 分别求出每个区域的 \bar{L}_w 来代替 $L_d(i,j) = L(i,j)/(1 + L(i,j))$ 分母中的 $L(i,j)$. 为此作者引入了中心-环绕函数 $V(i,j,s)$ ^[32], 其中 s 为高斯滤波器的尺度, 通过 $|V(i,j,s_m)| < \epsilon$ 来确定合适的 s_m , 最后把 $|V_1(i,j,s_m)|$ 作为区域的平均亮度值, 于是得到

$$L_d(i,j) = \frac{L(i,j)}{1 + V_1(i,j,s_m(i,j))} \quad (19)$$

这样就实现了局部的自动遮光-增光. 通过对一系列高动态范围图像的测试分析可以看到, 此算法有很好的鲁棒性, 对大多数图像都能得到令人满意的效果.

2.2.4 自适应对数映射算子

Drago 等^[33] 根据 Stockham^[34] 的研究结果, 提出了一种自适应的对数映射算子. Stockham 建议显示器亮度和场景亮度之间采用如下的映射关系

$$L_d = \frac{\ln(1 + L_w)}{\ln(1 + L_{wmax})} \quad (20)$$

很明显这也是个全局算子. 由于式 (20) 是对 $(1 + L_w)$ 求以 $(1 + L_{wmax})$ 为底的对数, Drago 等根据像素及其邻域信息自适应地调节 $(1 + L_{wmax})$, 并且利用 Perlin 和 Hoffert 的 “偏置” 幂函数^[35] 在 2 和 10 之间平滑地进行基底插值, 提出了自适应对数映射算子. 最终的计算公式如下

$$L_d = \frac{L_{dmax}}{\lg(1 + L_{wmax})} \times \frac{\ln(1 + L_w)}{\ln\left(2 + 8 \left(\frac{L_w}{L_{wmax}}\right)^{\frac{\ln(b)}{\ln(0.5)}}\right)} \quad (21)$$

式中 $L_{wmax} = (e \cdot L_{max})/L_{wa}$, 其中 L_{max} 为场景最大亮度; e 为曝光因子, 用来调节图像的整体亮度; L_{wa} 代表图像局部的适应水平, 是高斯卷积后的图像; 偏置参数 b 最终决定了高亮度像素被压缩的程度以及图像暗区域的细节可见度, 根据经验, b 的取值一般在 0.7 到 0.9 之间, 通过试验可知 $b = 0.85$ 时对大多数图像都能得到令人满意的效果. 这种方法运算量很小, 已经被用来实时压缩高动态范围视频.

2.2.5 Retinex 算法

Retinex 理论最早由 Land 和 McCann^[36] 提出, 它是人类视觉系统感知亮度和色彩的一种计算模型, 能够在进行动态范围压缩的同时保持色彩的恒常性, 因而可以用来压缩图像的动态范围. 早在上世纪 80 年代初, Frankle 和 McCann^[37] 就提出了一种基于 Retinex 算法的 TMO, Funt 等开发了几种 Retinex 算法的 Matlab 实现代码^[38], 并且把这种算法和其他的各种色阶映射算子进行了比较. 在这种算法中, 迭代次数决定了动态范围压缩的程度, 越少的迭代次数将会把动态范围压缩到越小的范围, 然后再通过查表的后续处理将动态范围进行合适的扩展以适应各种显示设备. 后来, 各种关于 Retinex 的变种算法不断被提出来, Jobson 等^[39] 将以前的单尺度 Retinex 推广到了多尺度的情况, 并且测试了超过 100 种的图像, 能较好地校正那些曝光不足或是未进行白平衡的图像, 显著地改善了图像的视觉效果. 受限于当时的图像获取设备, Jobson 未能对动态范围更高的图像作进一步的测试, 但对高动态范围图

像的处理提出了一些指导性的意见,并指出多尺度的 Retinex 算法会放大失真,对于那些进行了有损压缩的 JPEG 之类的图像,有可能使原本觉察不到的失真变得很明显,所以应用多尺度 Retinex 算法一定要在有损压缩之前才能达到较好的效果。

2.2.6 光照度分离压缩

Arend^[40] 认为人类视觉系统在观察图像或场景的时候,并不是把图像或场景当作单一的实体来认知的,而是将其分解为一组“本征图像 (Intrinsic image)”,每个“本征图像”描述了一种场景感知量(如:光照、反射、透射、方位、距离等)。基于这种思想, Tumblin^[41] 提出了一种分层的方法来压缩合成图像的动态范围,该方法将图像分成反射层和照明层,照明层决定了场景的动态范围和亮度信息,而反射层则包含了场景的细节。通过压缩照明层来达到动态范围压缩的目的,而反射层则保留了图像的细节信息。由于真实场景的反射层不像合成图像那样容易获得(事实上,真实场景的反射层也是可以得到的,如基于图像序列的方法^[42-43]),这种方法比较适用于合成图像。对于真实场景图像,也有很多类似的方法,如 Tumblin^[44] 提出的基于 LCIS (Low curvature image simplifier) 的分层压缩方法,该方法借鉴了画家作画的方法:先画出场景中不同区域的分界线,然后再由粗到细地将各个区域的细节加进去。通过 LCIS 可以得到一组对应于不同尺度细节信息的图像,对不同的尺度使用不同的压缩比,然后通过求和得到动态范围压缩后的图像。Durand 和 Dorsey^[45] 则将图像分成了包含图像大尺度和小尺度细节信息的基层和细节层,通过对基层进行大比例的压缩来达到动态范围压缩的目的,该方法利用双边滤波器 (Bilateral filter) 来获得基层图像。随后 Choudhury 等^[46] 也采用了这种分成基层和细节层的方法,所不同的是新方法是利用三边滤波器 (Trilateral filter) 来计算基层图像,而细节层图像则通过原始图像减去基层图像得到。这几种方法的思想都差不多,只是获取各层图像所使用的滤波器不一样,从动态范围压缩的效果来看,LCIS 和三边滤波器要好于双边滤波器。

局部算子能较好地模拟人类视觉系统,因而引起了学者的广泛关注,其他较具代表性的局部算子还有 Fattal 的梯度域压缩的方法^[47] 等。

全局算子和局部算子都只是解决了静态图像的动态范围压缩,但是对于序列图像来说,动态范围压缩后的图像将不可避免地丢失一些只有在高亮度刺激下人眼才会出现的特性(如闪光和光晕等),而且人眼明视觉和暗视觉的适应过程也应考虑在内,因此近年来也有不少时变算子被提出来,有兴趣的读

者可以参考文献 [16, 48-50]。

3 结论

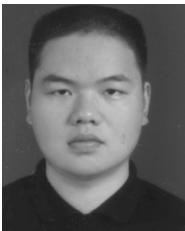
高动态范围图像和色阶映射算子是近年来图像图形学界研究的一个新领域,关于高动态范围图像传感器器件,国际上也只是出现了一些实验性的原理性样机,相信随着半导体技术和传感检测技术的发展,各种新型的传感器将最终使高动态范围图像的获取如同现在拍摄一张数码照片这样简单,同时也将形成一套高效的色域覆盖范围宽广、高保真的编码标准;由于人类视觉系统的生理特性至今为止都没有得到彻底的研究,现有的视觉模型都只是一个或某几个方面的近似,因此目前的 TMO 都存在这样或那样的问题,并不能适用于所有的情形,随着人类认识自身生理特性的进一步深入,必将有更加普遍适用的 TMO 诞生。

References

- 1 Mann S, Picard R W. On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. In: Proceedings of IS & T's 48th Annual Conference. Washington D. C., USA: 1995. 422-428
- 2 Debevec P E, Jitendra M. Recovering high dynamic range radiance maps from photographs. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 1997. 369-378
- 3 Madden B C. Extended Intensity Range Imaging, Technical Report MS-CIS-93-96, GRASP Laboratory, University of Pennsylvania, USA, 1993
- 4 Aggarwal M, Ahuja N. Split aperture imaging for high dynamic range. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 10-17
- 5 Dalsa. Datasheet [Online], available: <http://www.dalsa.com/pi/products/DSC.asp>, August 27, 2008
- 6 Fairchild Imaging. Datasheet [Online], available: <http://www.fairchildimaging.com/products/fpa/ccd/area/>, August 27, 2008
- 7 Hamamatsu Photonics. Datasheet [Online], available: <http://sales.hamamatsu.com/en/products/solid-state-division/image-sensors/ccd.php>, August 27, 2008
- 8 Chen S, Ginosar R. Adaptive SensitivityTM CCD image sensor. In: Proceedings of the 18th Convention of Electrical and Electronics Engineers in Israel. Tel Aviv, Israel: IEEE, 1995. 363-365
- 9 Nayar S K, Mitsunaga T. High dynamic range imaging: spatially varying pixel exposures. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, USA: IEEE, 2000. 472-479
- 10 Ward G. Global illumination and HDRI formats [Online], available: <http://www.debevec.org/IBL2003/GWcourse-TalkIBL2003.pdf>, August 27, 2008

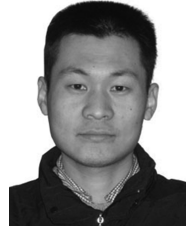
- 11 Stevens S S. To honor fechner and repeal his law: a power function, not a log function, describes the operating characteristic of a sensory system. *Science*, 1961, **133**(3446): 80–86
- 12 Stevens S S, Stevens J C. Brightness function: parametric effects of adaptation. *Journal of the Optical Society of America*, 1963, **53**(3): 375–385
- 13 Miller N J, Ngai P Y, Miller D D. The application of computer graphics in lighting design. *Journal of the IES*, 1984, **14**(1): 6–26
- 14 Upstill S D. The Realistic Presentation of Synthetic Images: Image Processing in Computer Graphics [Ph.D. dissertation], University of California, USA, 1985
- 15 Tumblin J, Rushmeier H. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 1993, **13**(6): 42–48
- 16 Ferwerda J A, Pattanaik S N, Shirley P, Greenberg D P. A model of visual adaptation for realistic image synthesis. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 1996. 249–258
- 17 Blackwell H R. An Analytical Model for Describing the Influence of Lighting Parameters Upon Visual Performance, Technical Foundations CIE 19.21, 1981
- 18 Ward G. A contrast-based scalefactor for luminance display. *Graphics Gems IV*. San Diego: Academic Press Professional, 1994. 415–421
- 19 Larson W G, Rushmeier H, Christine P. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 1997, **3**(4): 291–306
- 20 Scheel A, Stamminger M, Seide H P. Tone reproduction for interactive walkthroughs. *Computer Graphics Forum*, 2001, **19**(3): 301–311
- 21 Cohen J, Tchou C, Hawkins T, Debevec P. Real-time high dynamic range texture mapping. In: Proceedings of the 12th Eurographics Workshop on Rendering. Berlin, Germany: Springer-Verlag, 2001. 313–320
- 22 Chiu K, Herf M, Shirley P, Swamy S, Wang C, Zimmerman K. Spatially nonuniform scaling functions for high contrast images. In: Proceedings of Graphics Interface. Toronto, Canada: KFUPM Eprints, 1993. 245–253
- 23 Land E H. The retinex theory of color vision. *Scientific American*, 1977, **237**(6): 108–128
- 24 Marr D. *Vision*. New York: W. H. Freeman, 1982
- 25 Tumblin J E, Rushmeier H E. Tone Reproduction for Realistic Computer Generated Images, Technical Report GIT-GVU-91-13, Graphics, Visualization, and Usability Center, Georgia Institute of Technology, 1991
- 26 Schlick C. Quantization techniques for visualization of high dynamic range pictures. In: Proceedings of the 5th Eurographics Workshop on Rendering. Berlin, Germany: Springer-Verlag, 1994. 7–20
- 27 Reinhard E, Stark M, Shirley P, Ferwerda J. Photographic tone reproduction for digital images. In: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 2002. 267–276
- 28 White M. *The New Zone System Manual*. New York: Morgan and Morgan, 1990
- 29 Adams A, Baker R. *The Camera (Ansel Adams Photography, Book 1)*. Boston: Bulfinch, 1995
- 30 Adams A, Baker R. *The Negative (Ansel Adams Photography, Book 2)*. Boston: Bulfinch, 1995
- 31 Adams A, Baker R. *The Print (Ansel Adams Photography, Book 3)*. Boston: Bulfinch, 1995
- 32 Blommaert F J J, Martens J B. An object-oriented model for brightness perception. *Spatial Vision*, 1990, **5**(1): 15–41
- 33 Drago F, Myszkowski K, Annen T, Chiba N. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 2003, **22**(3): 419–426
- 34 Stockham T G. Image processing in the context of a visual model. *Proceedings of the IEEE*, 1972, **60**(7): 828–842
- 35 Perlin K, Hoffert E M. Hypertexture. *ACM SIGGRAPH Computer Graphics*, 1989, **23**(3): 253–262
- 36 Land E H, McCann J J. Lightness and retinex theory. *Journal of the Optical Society of America*, 1971, **61**(1): 1–11
- 37 Frankle J, McCann J J. Method and Apparatus for Lightness Imaging, U.S. Patent 4384336, May 1983
- 38 Funt B, Ciurea F, McCann J J. Retinex in Matlab. In: Proceedings of the 8th Color Imaging Conference: Color Science, Systems and Applications. Scottsdale, USA: SFU, 2000. 112–121
- 39 Jobson D J, Rahman Z, Woodell G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 1997, **6**(7): 965–976
- 40 Arend L. Surface colors, illumination, and surface geometry: intrinsic-image models of human color perception. *Lightness, Brightness, and Transparency*. New York: Wiley, 1994. 159–214
- 41 Tumblin J, Hodgins J K, Guenter B K. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 1999, **18**(1): 56–94
- 42 Weiss Y. Deriving intrinsic images from image sequences. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 68–75
- 43 Matsushita Y, Lin S, Kang S B, Shum H Y. Estimating intrinsic images from image sequences with biased illumination. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer-Verlag, 2004. 274–286

- 44 Tumblin J, Turk G. LCIS: a boundary hierarchy for detail-preserving contrast reduction. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 1999. 83–90
- 45 Durand F, Dorsey J. Fast bilateral filtering for the display of high dynamic range images. *ACM Transactions on Graphics*, 2002, **21**(3): 257–266
- 46 Choudhury P, Tumblin J. The trilateral filter for high contrast images and meshes. In: Proceedings of the Eurographics Symposium on Rendering. Los Angeles, USA: ACM, 2005. 186–196
- 47 Fattal R, Lischinski D, Werman M. Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 2002, **21**(3): 249–256
- 48 Pattanaik S N, Tumblin J, Yee H, Greenberg D P. Time-dependent visual adaptation for fast realistic image display. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 2000. 47–54
- 49 Ledda P, Santos L P, Chalmers A. A local model of eye adaptation for high dynamic range images. In: Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa. Stellenbosch, South Africa: ACM, 2004. 151–160
- 50 Durand F, Dorsey J. Interactive tone mapping. In: Proceedings of the Eurographics Workshop on Rendering Techniques. Berlin, Germany: Springer-Verlag, 2000. 219–230



杨克虎 中国科学院自动化研究所博士研究生。2003 年于西北工业大学自动控制系获得学士学位。主要研究方向为图像处理, 计算机视觉和惯性导航。本文通信作者。E-mail: kehu.yang@ia.ac.cn
(**YANG Ke-Hu** Ph. D. candidate at the Institute of Automation, Chinese Academy of Sciences. He received his

bachelor degree from Northwestern Polytechnical University in 2003. His research interest covers image processing, computer vision, and inertial navigation. Corresponding author of this paper.)



姬 靖 中国科学院自动化研究所博士研究生。2001 年于西安交通大学获得学士学位, 2005 年于中国海洋大学获得硕士学位。主要研究方向为图像测量和增强现实。E-mail: jing.ji@ia.ac.cn

(**JI Jing** Ph. D. candidate at the Institute of Automation, Chinese Academy of Sciences. He received his bachelor degree from Xi'an Jiaotong University in 2001 and master degree from Ocean University of China in 2005, respectively. His research interest covers image-based metrology and augmented reality.)



郭建军 1984 年于中南大学获得工学学士学位, 1989 年于清华大学自动化系获得硕士学位, 现在宏易未来(北京)科技有限公司工作。主要研究方向为图像处理和机器视觉。

E-mail: jianjun.guo@soliwell.com

(**GUO Jian-Jun** Received his bachelor degree from Central South University in 1984 and master degree from Tsinghua University in 1989, respectively. Currently, he is working at the HPI Innovation (Beijing) Co., Ltd.. His research interest covers image processing and computer vision.)



郁文生 中国科学院自动化研究所研究员。1998 年于北京大学力学与工程科学系获理学博士学位。主要研究方向为系统鲁棒控制理论, 时滞系统的稳定性分析和数学机械化在控制理论中的应用。

E-mail: wensheng.yu@ia.ac.cn

(**YU Wen-Sheng** Professor at the Institute of Automation, Chinese Academy of Sciences. He received his Ph. D. degree in dynamics and control from Peking University in 1998. His research interest covers robust control theory, stability analysis of time-delay system, and applications of mathematics mechanization in control theory.)