

不同曝光值图像的直接融合方法*

张军^{1,2+}, 戴霞², 孙德全¹, 王邦平¹

¹(四川大学 计算机学院 图形图像研究所, 四川 成都 610064)

²(西华大学 数学与计算机学院, 四川 成都 610039)

Directly Fusion Method for Combining Variable Exposure Value Images

ZHANG Jun^{1,2+}, DAI Xia², SUN De-Quan¹, WANG Bang-Ping¹

¹(Institute of Image and Graphic, College of Computer Science, Sichuan University, Chengdu 610064, China)

²(Institute of Computer and Mathematics, Xihua University, Chengdu 610039, China)

+ Corresponding author: E-mail: mecca_zj@163.com

Zhang J, Dai X, Sun DQ, Wang BP. Directly fusion method for combining variable exposure value images. *Journal of Software*, 2011, 22(4): 813-825. <http://www.jos.org.cn/1000-9825/3795.htm>

Abstract: In this paper, an efficient and robust image fusion method is proposed to extract pixel intensity under the best exposure value directly from a low dynamic range of images with variable exposure values. Using this method, the HDR (high dynamic range) image can be computed and shown on the traditional display device without any prior information on a camera or scene. The proposed image fusion method is based on a mathematical model of a pixel intensity curve, established by a special robust curve fitting arithmetic criteria to evaluate the image quality of each pixel. The sufficient experimentations on various scenes indicate that the proposed method is more efficient than traditional HDRI (high dynamic range imaging) technologies and more robust when imaging noise, object movement, and small camera displacement simultaneously.

Key words: high dynamic range imaging; robust curve fitting; image fusion; least squares fitting

摘要: 提出了一种直接从同一场景多次不同曝光值下成像的 LDR(low dynamic range)图像序列中提取每个像素位置最佳成像信息的图像融合方法,可以在无需任何拍摄相机参数及场景先验信息的情况下,快速合成适合在常规设备上显示的 HDR(high dynamic range)图像.该方法利用特殊设计的鲁棒性曲线拟合算法建立 LDR 图像序列中每个像素位置像素值曲线的数学模型,并由此给出评价单个像素成像时曝光合适程度的标准和融合最佳成像像素信息的方法.对不同场景的大量实验结果显示,该方法的计算结果与传统 HDR 成像技术经过复杂的 HDR 重建和色调映射计算后得到的结果相当,但具有更高的计算效率,并同时图像噪声、相机微小移动和运动目标的影响具有较好的鲁棒性.

关键词: 高动态范围成像;鲁棒性曲线拟合;图像融合;最小二乘拟合

中图法分类号: TP391 文献标识码: A

随着数字相机技术的快速发展和一些应用领域的实际需求,利用数字相机在不同曝光值下对同一场景拍

* 基金项目: 国家高技术研究发展计划(863)(2007AA01Z328, 2009AA01Z332)

收稿时间: 2009-07-23; 定稿时间: 2009-12-02

摄一系列低动态范围图像并合成高动态范围图像的技术得到了快速的发展,被称为高动态范围成像(high dynamic range imaging,简称 HDRI)技术.HDRI 技术目前已被广泛应用于电影、摄影、计算机图形学和摄影测量等领域^[1].从一系列不同曝光值下成像的低动态范围(low dynamic range,简称 LDR)图像序列中合成高动态范围(high dynamic range,简称 HDR)图像,需要一个关键的假设:在整个拍摄过程中,相机位置和场景中的目标都不能有移动.而实际户外场景中的运动目标却是普遍存在的,使合成的 HDR 图像在运动目标区域存在重影^[2,3].在高感光度下成像以缩短拍摄时间是一种克服运动目标影响的有效方法,但会使 CCD 相机成像出现严重的噪声^[4].同时,使用廉价商用相机手工调整曝光值拍摄场景会使相机在拍摄过程中存在微小旋转和平移,需要事先对 LDR 图像序列进行图像配准^[5,6].最后,直接观察 HDR 图像需要特殊的高动态范围显示硬件^[7,8].由于这种显示设备价格昂贵,目前需要对 HDR 图像进行色调映射以便压缩其动态范围,使之可以在常规设备上显示^[9-12].另一方面,为了具有更好的色彩还原能力和动态范围捕捉能力,目前商用相机的内部响应函数(camera response function)较为复杂,造成 HDR 重建^[13]时需要精密地计算相机内部响应函数^[14].

由于受到上述复杂因素的影响,传统 HDRI 技术一般需要“相机内部响应函数标定”、“图像配准”、“噪声抑制”、“重影去除”、“HDR 重建”和“色调映射”等步骤才能将 LDR 图像序列合成为适合在常规设备上观察的 HDR 图像^[2,15].然而,最近的研究结果显示^[16],在常规显示设备上,如此大费周折计算出的 HDR 图像质量并没有明显优于曝光合适的 LDR 图像.实际上,为在常规设备上显示 HDR 图像而重建出场景的真实亮度信息,再通过色调映射压缩场景亮度动态范围是完全没有必要的.

本文通过鲁棒性曲线拟合方法建立 LDR 图像序列每个像素位置像素值曲线的数学模型,并将该曲线的拐点判定为该像素位置的最佳曝光成像值,从而可以由 LDR 图像序列中提取曝光合适的像素信息,直接融合出适合在常规设备上显示的 HDR 图像.与传统 HDRI 技术相比,本文方法具有以下特点:(1) 将图像噪声、运动目标影响和相机微小移动影响都归结为广义的噪声,一次性完成 HDR 图像合成中图像配准、噪声抑制和重影去除这 3 个步骤;(2) 无需繁琐、耗时的 HDR 重建和色调映射计算;(3) 无需相机参数和场景先验信息;(4) 算法较为简洁、高效,适合像素级的并行处理机制.

本文第 1 节介绍与本文紧密相关的前人的工作.第 2 节给出所提算法的构建动机和算法流程.第 3 节给出一种具有鲁棒性的曲线拟合算法,并讨论其拟合误差.第 4 节给出详细的直接融合 LDR 图像序列的方法.第 5 节给出实验结果及分析.第 6 节给出结论和未来的工作.

1 相关工作

HDRI 技术的基本思想是,对同一场景采用不同的曝光值(exposure value)进行采样,通过相机内部相应函数的逆运算得到场景的实际亮度.从图像融合的角度,HDRI 技术可视为融合同一场景多次曝光图像的一种技术:通过提取同一场景在不同曝光值下成像的信息,融合成局部自适应曝光的高动态范围图像^[17].Rovid 等人^[11]提出基于图像梯度、Vavilin 和 Jo^[18]提出基于图像区域方差和熵值作为区域细节水平的评价标准.这两种方法都采用在同一场景的不同曝光值 LDR 图像序列中挑选包含细节水平最多的子区域,直接合成 HDR 图像的思想.这些方法的缺点在于,图像子区域包含细节程度的评价函数计算较为耗时且并不稳定,区域划分大小需要精细调整,不同区域间的过渡需要复杂的后期平滑处理.同时,对同一场景的有限次采样并不能保证场景中每个区域都至少有一次在最佳曝光值下成像,产生融合结果可能存在曝光不足或曝光过度的区域.本文借鉴了直接提取同一场景不同曝光值 LDR 图像序列像素信息的图像融合思想.与前人算法不同,为了提高计算效率和融合效果,本文建立单个像素位置在不同曝光值下成像的像素值曲线数学模型.该模型可以从有限次采样数据中计算出每个像素位置在任意曝光值下成像的像素值,并可判定每个像素位置的最佳曝光值.在此基础上,采用简洁的图像融合方法^[19]可以高效地合成适合在常规设备上显示的高动态范围图像.

2 本文融合算法的构建动机及算法流程

从场景的有限 LDR 图像序列中直接提取每个像素位置最佳成像信息的构想需要解决两个关键问题:

- (1) 建立评价像素点成像时曝光值合适程度的标准;
- (2) 从每个像素位置像素值点列中推测该像素位置在任意曝光值下成像的像素值.

2.1 像素级的曝光适应度评价方法

通过大量的实验和分析,我们总结出一个图像中单个像素位置在不同曝光值下成像的规律.图 1 给出了文献[13]中一个教堂场景在 16 次不同曝光值下成像的 LDR 图像序列.当场景中某一点的 LDR 图像序列从曝光不足逐渐变为曝光过度时,必然存在一个发生质变的点,即为图像的最佳曝光成像点.该点对应于亮度序列一阶差商值最大的点(二阶差商值的零点),数学上称之为“拐点”.故本文提出利用 LDR 图像序列像素值序列的拐点作为判定最佳曝光成像点的依据.该判定标准仅涉及单个像素的像素值序列,相对于前人使用区域像素值梯度、方差或熵等标准要简洁和有效得多.

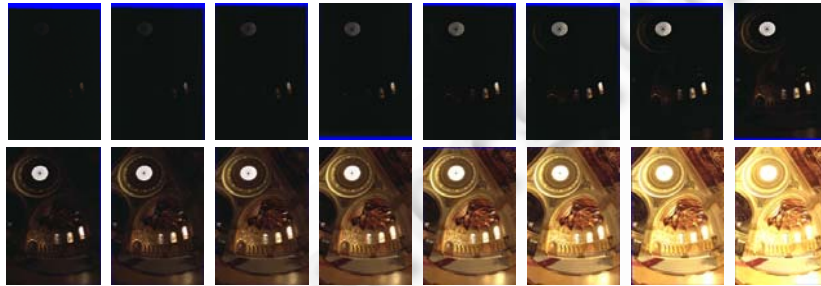


Fig.1 A group of LDR images in Ref.[13]

图 1 文献[13]中的一组 LDR 图像

设 LDR 图像序列为 $\{I_{uv}^i\}_{i=1}^N$, 其中 I_{uv}^i 表示第 i 幅图像在 (u,v) 位置的像素值, 令

$$d_i^1|_{uv} = \frac{I_{uv}^{i+1} - I_{uv}^{i-1}}{(i+1) - (i-1)}, \quad i = 1, 2, \dots, N-1, \quad d_0^1|_{uv} = \frac{I_{uv}^1 - I_{uv}^0}{1-0}, \quad d_N^1|_{uv} = \frac{I_{uv}^N - I_{uv}^{N-1}}{(N) - (N-1)} \quad (1)$$

图 2 给出了教堂场景中 5 个不同区域对应的像素值点列 $\{I_{uv}^i\}_{i=1}^N$ 及一阶差商点列 $\{d_i^1|_{uv}\}_{i=1}^N$. 仔细观察即可发现, LDR 图像序列中每个像素位置的最佳曝光成像发生点都与其一阶差商值的最大点相对应(如图 2 中方形点所示). 即若有 $d_{index}^1|_{uv} = \arg \max_{1 \leq i \leq N} d_i^1|_{uv}$, 则 I_{uv}^{index} 为最佳曝光值下成像的像素值.

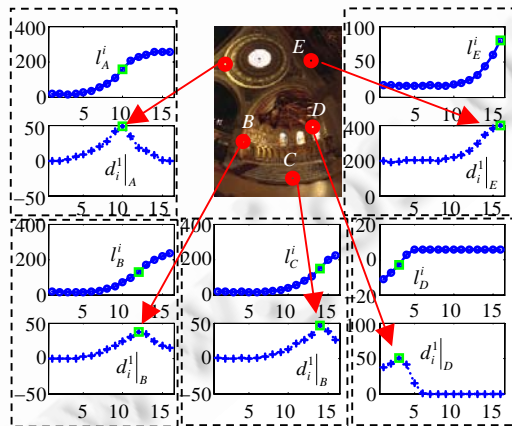


Fig.2 Pixel intensity curves and their first-order difference quotient at different points of church scene's LDR images

图 2 教堂场景中不同区域的 LDR 图像序列像素值曲线及其一阶差商曲线

2.2 LDR图像序列像素值曲线模型

LDR 图像序列仅为场景在不同曝光值下的有限次采样,场景中某些区域的最佳曝光成像点可能发生在两次采样之间.同时,在融合 LDR 图像序列时也需要平滑不同区域间的亮度差异.这些都需要得到场景中每个像素点在任意曝光值下的成像结果,即一条平滑的像素值曲线.

相机成像的过程非常复杂,但仍可抽象为如下方程^[13]:

$$I_{uv}^i = f(X) = f(\eta E_{uv} \Delta t_i) \quad (2)$$

其中, I_{uv}^i 表示 LDR 图像序列中第 i 幅图像在 (u,v) 像素位置的像素值, X 为曝光值, f 为相机内部响应函数, η 为由相机感光度 ISO 和光圈值大小决定的尺度参数, E_{uv} 为场景中 (u,v) 像素位置的真实亮度值, Δt_i 为第 i 幅图像的曝光时间.理论分析和实际测试都显示,相机响应函数 f 在有限区间内是光滑函数^[13].

然而,实际场景存在以下情况,从而由 LDR 图像序列中建立像素值曲线模型是一个具有挑战性的问题:

- (1) 场景中不存在动态目标,其 LDR 图像序列中的 $\{I_{uv}^i\}_{i=1}^N$ 可能并非对应于同一个亮度 E_{uv} ;
- (2) 手工操作相机造成的微小相机位置移动会使 $\{I_{uv}^i\}_{i=1}^N$ 并非对应于同一个亮度 E_{uv} ;
- (3) 成像过程存在不可避免的噪声 $I_{uv}^i = f[\eta E_{uv} \Delta t_i + \omega_1(\Delta t_i, \eta)]$, 其中 $\omega_1(\Delta t_i, \eta)$ 是与曝光时间和 CCD 感光度相关的随机噪声^[4];
- (4) 图像存储时,由于压缩或其他原因同样会引入噪声 $I_{uv}^i = f(\eta E_{uv} \Delta t_i) + \omega_2$, 其中, ω_2 是随机噪声.

最终的相机成像方程比公式(2)要复杂得多:

$$I_{uv}^i = f[\eta E_{uv}(i) \Delta t_i + \omega_1(\Delta t_i, \eta)] + \omega_2 \quad (3)$$

传统 HDR 技术的预处理步骤分别从单幅图像中消除 ω_2 , 从序列图像中消除 $\omega_1(\Delta t_i, \eta)$, 进行图像配准和运动目标消除以便使 $E_{uv}(i) = E_{uv}$. 但从方程(3)中可以看到,影响 LDR 图像序列的各种因素是纠缠在一起的,加上 f 的非线性特性,传统 HDR 技术的 3 个预处理步骤是相互制约且低效的.

本文仅以方程(2)中 $f(X)$ 光滑性作为基础,将方程(2)与方程(3)中不同的项都归结为广义噪声,通过选择基图像和鲁棒性曲线拟合的方式直接对 $\{I_{uv}^i\}_{i=1}^N$ 建立数学模型(本文第 3 节给出算法细节),能够有效克服由于运动目标和相机移动造成异常数据(outliers)的影响.

2.3 融合算法流程

在建立像素级曝光适应度评价标准和 LDR 图像序列像素值曲线数学模型之后,融合 LDR 图像序列的过程就可以表述为:逐像素建立 LDR 图像序列像素值曲线数学模型、寻找曲线拐点并按照拐点所在位置的索引提取 LDR 图像序列中该图像的像素信息.

本文提出的 LDR 图像序列融合算法的流程如图 3 所示.

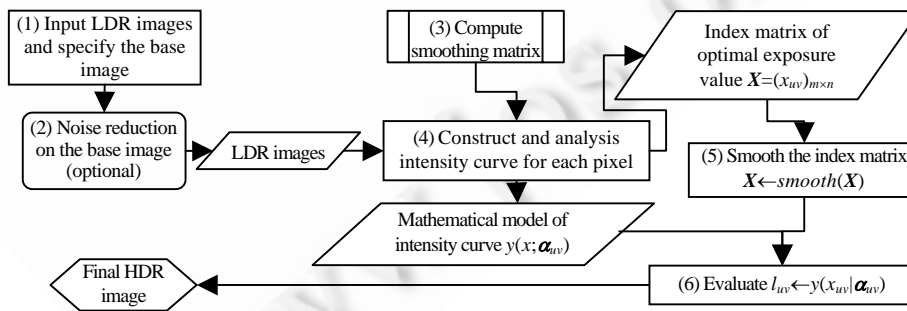


Fig.3 Flowchart of our arithmetic

图 3 本文算法流程图

与传统 HDRI 技术流程^[2](如图 4 所示)相比,本文算法具有更加简洁的流程,并排除了传统方法中耗时的迭

代反馈预处理(如图4所示中的虚线)、HDR重建和Tone Mapping等步骤.本文算法流程的第(2)步是可选的,利用较为成熟的针对单幅图像的噪声滤波方法^[20]对基图像进行处理.这一步骤将针对序列图像的噪声滤波方法与传统针对单幅图像的噪声滤波方法相结合,可进一步提高噪声抑制效果,是本文方法灵活性的一种体现.注意到该步骤仅需对单幅图像进行噪声滤波,所以并不影响本文方法的计算速度.

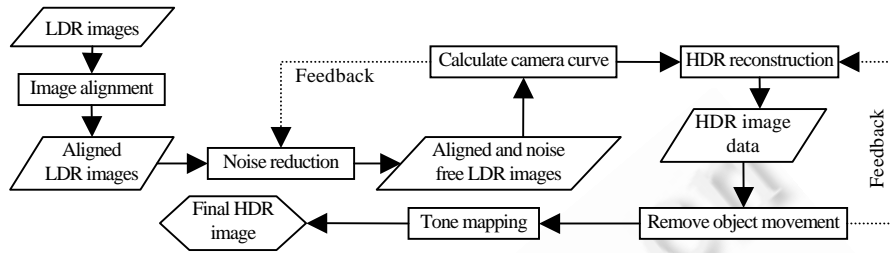


Fig.4 Flowchart of traditional HDRI technologies

图4 传统 HDRI 技术算法流程图

3 鲁棒性曲线拟合算法

根据本文第2节的叙述,需要建立LDR图像序列像素值曲线的数学模型,并对异常数据具有一定的鲁棒性.传统多项式曲线拟合和其他一些参数曲线^[21]拟合能够有效平滑原始数据,但较易受到异常数据的干扰.非线性鲁棒统计(nonlinear robust statistics)方法^[22]具有较好的抵抗异常数据的能力,但其实现是一个迭代过程,对图像序列处理这种涉及千万个点列处理的情况非常耗时.本节构造一种基于拟合点列二阶差商方法建立LDR图像序列像素值曲线的数学模型,可兼顾计算速度和鲁棒性.

设曲线型值点列为 $\{(x_i, y_i)\}_{i=0}^N$,可用其二阶差商 $\{(x_i, d_i^2)\}_{i=0}^N$ 来近似表示其二阶导数:

$$d_i^2 = \frac{d_{i+1}^1 - d_{i-1}^1}{x_{i+1} - x_{i-1}}, i = 1, 2, \dots, N-1, d_0^2 = \frac{d_1^1 - d_0^1}{x_1 - x_0}, d_N^2 = \frac{d_N^1 - d_{N-1}^1}{x_N - x_{N-1}} \quad (4)$$

$$d_i^1 = \frac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}}, i = 1, 2, \dots, N-1, d_0^1 = \frac{y_1 - y_0}{x_1 - x_0}, d_N^1 = \frac{y_N - y_{N-1}}{x_N - x_{N-1}}$$

构造多项式曲线

$$y(x; \alpha) = l(x) + n(x; \alpha) \quad (5)$$

其中: $l(x) = y_0 + \frac{x-x_0}{x_N-x_0}(y_N-y_0)$ 为线性部分; $n(x; \alpha) = (x-x_0)(x-x_N) \left[\sum_{k=0}^3 a_k x^k \right]$ 为非线性部分,曲线的二阶导函数

数主要决定于该部分; $\alpha = (a_0, a_1, a_2, a_3)^T$ 为曲线参数.

简单地分析即可得出:

- (1) 曲线方程(5)插值点列的两端点 (x_0, y_0) 和 (x_N, y_N) .
- (2) 其二阶导数函数为

$$\frac{d^2 y}{dx^2} = A(x)a_0 + B(x)a_1 + C(x)a_2 + D(x)a_3 \quad (6)$$

其中, $A(x) = [14x^3 - 6x^2(x_0+x_N) + 6x(x-x_0)(x-x_N)]$, $B(x) = [10x^2 - 4x(x_0+x_N) + 2(x-x_0)(x-x_N)]$, $C(x) = [6x - 2(x_0+x_N)]$, $D(x) = 2$.

通过大量的实验分析发现,利用拟合点列的二阶差商和点列中人为指定的“基点(basis point)” (x_τ, y_τ) 可以有效地平滑点列并较少受到异常数据的影响.

令 $\left. \frac{d^2 y}{dx^2} \right|_{x=x_i} = d_i^2$ 且 $y(x_\tau) = y_\tau, i=0, 1, 2, \dots, N, 1 \leq \tau \leq N-1$,则有,

$$\begin{bmatrix} A(x_0) & B(x_0) & C(x_0) & D(x_0) \\ \vdots & \vdots & \vdots & \vdots \\ A(x_N) & B(x_N) & C(x_N) & D(x_N) \\ \gamma(x_\tau - x_0)(x_\tau - x_N)x_\tau^3 & \gamma(x_\tau - x_0)(x_\tau - x_N)x_\tau^2 & \gamma(x_\tau - x_0)(x_\tau - x_N)x_\tau & \gamma(x_\tau - x_0)(x_\tau - x_N) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} d_0^2 \\ \vdots \\ d_N^2 \\ \gamma[y_\tau - l(x_\tau)] \end{bmatrix} \quad (7)$$

其中, γ 为加权系数.

将公式(7)写成矩阵形式为

$$P\alpha = \beta \quad (8)$$

其中, $\alpha = (a_0, a_1, a_2, a_3)^T$, $\beta = (d_0^2, d_1^2, \dots, d_N^2, \gamma[y_\tau - l(x_\tau)])^T$.

在最小二乘意义下, 待定参数 α 的解为

$$\alpha = (P^T P)^{-1} P^T \beta = S \beta \quad (9)$$

其中,

$$S = (P^T P)^{-1} P^T \quad (10)$$

注意到, 公式(9)中矩阵 S 只与点列横坐标 $\{x_i\}_{i=0}^N$ 相关, 而与其纵坐标 $\{y_i\}_{i=0}^N$ 无关, 故称其为平滑矩阵 (smoothing matrix). 在对 LDR 图像序列进行平滑处理时, 图像中所有像素位置对应的平滑矩阵是同一个矩阵, 可以预先进行计算. 这意味着, 在对 LDR 图像序列进行分析时, 每个像素位置只涉及两次 N 维向量差商运算、一次低阶“矩阵-向量”乘积运算和若干次低阶多项式求值运算. 本文算法的速度优势主要来自于此.

在曲线拟合过程中, γ 的取值会影响拟合曲线对噪声的抑制能力和克服异常数据的能力. 较小的 γ 值可以使拟合曲线对噪声具有良好的抑制效果, 但更容易受到异常数据的影响; 较大的 γ 值可以使拟合曲线对异常数据不敏感, 但有可能将“基点”的噪声带入最终的计算结果. 故在实际融合 LDR 序列图像的实际应用中, 需要调整合适的 γ 值, 以获得“运动目标消除”和“噪声抑制”间的平衡.

方程(5)对原曲线的逼近精度可以通过公式(6)对曲线二阶导数的逼近精度间接决定.

定理 1. 设 $f(x) \in C^2[a, b]$, $f''(x) = y''(x; \alpha) + e(x)$, $f(x) = y(x; \alpha) + E(x)$. 若有 $y(a; \alpha) + f(a)$ 和 $y(b; \alpha) + f(b)$, 则有

$$|E(x)| \leq \frac{1}{8} \max_{a \leq x \leq b} |e(x)| (b-a)^2 \quad (11)$$

证明: 见附录 A.

4 多曝光图像序列直接融合算法实现

在本文第 3 节构造的 LDR 序列图像单个像素位置像素值曲线数学模型的基础上, 可以设计出简洁、高效的 LDR 图像序列融合方法. 设分辨率为 $m \times n$ 的 LDR 图像序列为 $\{I_i = (r_{uv}^i, g_{uv}^i, b_{uv}^i)_{m \times n}\}_{i=1}^N$, 可以分别对图像的 3 个通道中每个像素位置对应的像素值点列进行平滑. 为便于叙述, 用 $\{I_{uv}^i\}_{i=1}^N$ 表示 RGB 的 3 个通道中任意一个通道在像素位置 (u, v) 点的像素值点列, 并对 LDR 序列图像每个位置的像素值点列补充无曝光 (像素值为 0) 的左端点和完全过度曝光 (像素值为 255) 的右端点. 这样, 最少仅需包含 3 幅图像 (整体曝光不足、曝光适度和曝光过度) 的 LDR 图像序列即可求解方程组(8).

在得到 LDR 序列图像像素值曲线后, 该曲线的拐点即为最佳曝光成像点. 考虑到计算精度和稳定性的影响, 实际求解拐点时, 采用计算 RGB 的 3 个通道中曲线一阶导数平均最大值的方式, 即

$$x = \arg \max_{0 \leq x \leq 1} \{[r'(x; \alpha_r) + g'(x; \alpha_g) + b'(x; \alpha_b)]/3\} \quad (12)$$

所有像素位置的最佳曝光点将记录在索引矩阵 $X = (x_{uv})_{m \times n}$ 中. 受到文献[19]的启示, 可以通过平滑索引矩阵 X 来消除图像噪声对最佳曝光点计算精度的影响. 同时, 合理的平滑索引矩阵 X 也能使合成图像的不同亮度区域间的过渡更加平缓, 融合结果更贴近人眼视觉习惯.

本文提出的 LDR 序列图像融合方法对某一通道的处理流程可表述为:

1. 读入图像序列在 l 通道的像素值 $\{I_{uv}^l\}_{i=1}^N$, 并指定基图像 (basis image) I_τ 和权系数 γ ;

2. 对基图像进行处理 $I_r \leftarrow \Phi(I_r)$ (可选);
3. 令 $x_i = \frac{i}{N+1}$, $i = 0, 1, \dots, N, N+1$, 并按照公式(7)和公式(10)预先计算平滑矩阵 S ;
4. 遍历所有的像素位置 (u, v) , 执行如下操作:
 - 4.1. 令 $y_i = \begin{cases} 255, & i = N+1 \\ I_{uv}^i, & i = 1, 2, \dots, N \\ 0, & i = 0 \end{cases}$, 并按照公式(4)计算其二阶差商 $\{d_i^2\}_{i=0}^{N+1}$.
 - 4.2. 按照公式(9)计算拟合多项式参数 $\alpha_{uv} = (a_0, a_1, a_2, a_3)^T|_{uv}$.
 - 4.3. 按照公式(12)计算最佳曝光成像发生点 x_{uv} .
 - 4.4. 调整索引矩阵值 $x_{uv} \leftarrow ((p-q)x_{uv} + q)^c$, 其中 $p, q \in [0, 1], p \geq q$;
5. 平滑最佳曝光成像位置索引矩阵 $X \leftarrow \text{smooth}(X)$;
6. 遍历所有的像素位置 (u, v) , 计算融合结果: $I_{uv} \leftarrow y(x_{uv} | \alpha_{uv})$.

其中,对平滑矩阵的平滑操作 $\text{smooth}()$ 可以使用简单、快速的高斯函数卷积方法^[20].由于本文第3节给出的数学模型可以推测出图像在任意曝光值下成像的像素值,故平滑过的索引矩阵会自动在场景不同亮度区域间寻找合适的曝光值达到平滑过渡的效果,产生自然的局部自适应曝光的高动态范围图像.算法中第4个步骤仅涉及图像序列中单个像素位置的信息,完全符合数据并行计算的要求,可以在现代可编程图形、图像硬件上(GPU)高速地并行执行.由于让所有像素都在最佳曝光值下成像有可能会使图像内容出现亮度逆转(原场景中的高亮区域可能会成为融合图像中的低亮区域),并不一定符合每个人的视觉习惯,故算法通过 $x_{uv} \leftarrow ((p-q)x_{uv} + q)^c$ 的步骤对索引矩阵进行预先调整,可形成不同风格的融合图像.减小 $p-q$ 值,会使融合图像从更少的 LDR 图像中提取信息.当 $p-q=0$ 时,融合图像与索引位置为 q 的 LDR 图像一致.当 $c \rightarrow 0$ 时,融合图像为 LDR 图像序列中的最后一幅图像;当 $c \rightarrow \infty$ 时,融合图像为 LDR 图像序列中的第 1 幅图像.

5 实验结果及分析

本文的实验平台是一台 CPU 主频为 2.2GHz、2G 内存的 PC, Microsoft Visual Studio 2008 编译环境,使用 QtPfsGui 开源代码(<http://qtpfsgui.sourceforge.net/> Version 1.9.3 Apr 26 2009)实现前人 HDR 重建和 Tone Mapping 技术,下载互联网上常见的 LDR 图像序列(<http://www.hdrsoft.com/examples.html>)和 HDR 图像数据(<http://www.anywhere.com/gward/hdrenc/pages/originals.html>)以及我们手工操作相机(Olympus E510 型号数字相机)拍摄的 LDR 图像序列作为实验数据源.所有实验结果中,传统 HDR 技术采用文献[13]的方法进行 HDR 重建.

5.1 对静态场景的实验结果及分析

对如图 1 所示的静态场景,采用本文算法融合结果与采用文献[10]方法色调映射后的合成图像对比结果如图 5 所示.文献[10]中基于梯度域色调映射方法具有较强的局部自适应能力,形成的图像更加自然;其缺点是会改变原 LDR 图像序列局部区域的色度和饱和度,人为添加原图像没有的边缘和亮度波动,并在高亮区域(教堂场景中最右边的窗户区域)丢失部分细节.相对来说,本文方法直接提取原 LDR 图像序列像素值信息,更加忠实于原图像序列的成像风格,并保留更多的细节内容;但使每个像素都在最佳曝光值下成像会导致图像对比度下降而显得有些自然,需要通过平滑和调整索引矩阵以牺牲部分像素的最佳曝光性而获得更加自然的融合图像(如图 6(b)所示).图 6(c)中最亮点表示 1,最暗点表示 0.

图 6 显示了本文方法与文献[23]方法在计算结果上的对比分析.可以看出,对于亮度范围变化剧烈的场景,本文方法表现出更好的空间自适应提取最佳成像信息的能力.图 7 给出了本文方法与文献[10,23]方法的进一步对比分析结果.对于亮度差异较大的静态场景,文献[10]中的方法对场景中间部分(低亮区域与高亮区域存在复杂交叉情形)具有更好的自适应性,而本文方法的计算结果在场景非中心区域的曝光度更加合适.从整体细节保持程度以及饱和度保持程度上来说,本文算法的计算结果要优于文献[23]而略逊于文献[10].第 5.5 节对各种方法计算效率的对比分析将会显示,本文方法在综合考虑性能和效果时更具优势.

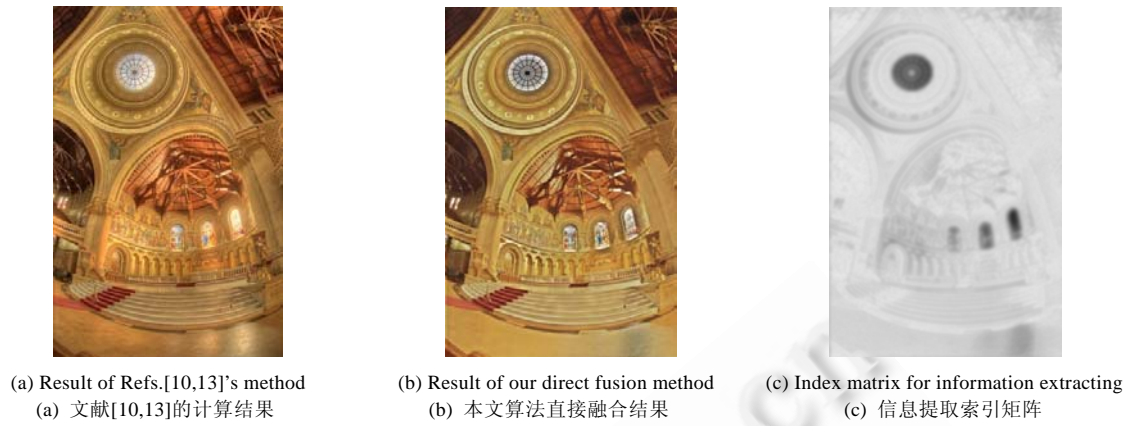


Fig.5 Comparison of Refs.[10,13]'s method and ours

图5 文献[10,13]中的方法与本文方法的比较

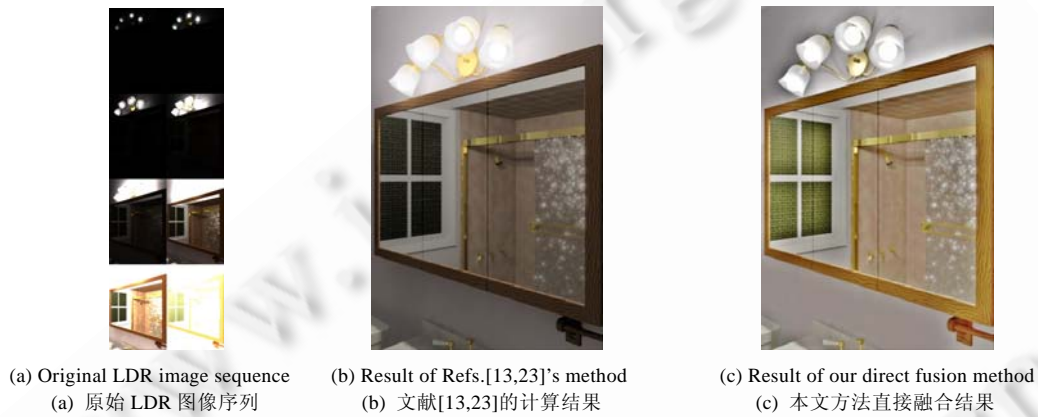


Fig.6 Comparison of Refs.[13,23]'s method and ours

图6 文献[13,23]中的方法与本文方法的比较



Fig.7 Further comparison results to the static scene

图7 静态场景进一步的对比分析

5.2 对具有运动目标场景和相机移动情况的实验结果及分析

当场景中出現运动目标时,传统 HDR 方法需要事先进行重影消除(ghost remove)才能进行 HDR 重建计算.如果拍摄过程存在相机移动,还需要事先进行图像配准预处理步骤.本文算法通过指定合适的基图像以及公式(7)中 γ 取较大值的方式,无需进行任何预处理步骤即可同时克服运动目标和相机移动的影响,具有较高的鲁棒性.图 8 给出了对文献[3]中一个含有运动目标场景的实验分析结果.对比分析显示,文献[3]中的方法采用复杂的迭代算法对重影消除效果与本文的简单曲线拟合算法的效果类似,有效克服了运动目标的影响.所不同的是,由于本文方法指定的基图像(图 8(a)中虚线包围的图像)包含有运动目标,造成最终合成的 HDR 图像保留了基图像的运动目标,这是本文方法目前不可避免的一个缺陷.

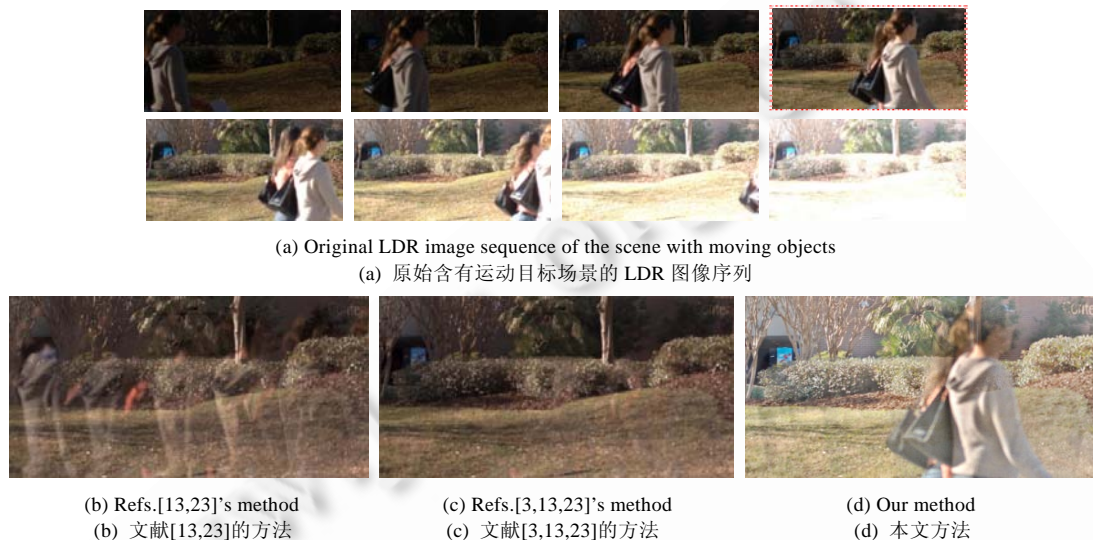


Fig.8 Comparison results to the scene with moving objects

图 8 含运动目标场景的实验结果的对比分析

通过 γ 取较大值的方式,本文方法对相机微小移动(相机在三脚架上由于手工调整曝光值而产生的移动)具有一定的鲁棒性.图 9 显示的是从具有相机微小移动 LDR 图像序列中融合的结果.计算图 9(a)时 $\gamma=0$,算法对相机微小移动克服能力较差,形成重影;而计算图 9(b)时 $\gamma=100$,算法对相机微小移动具有较强的鲁棒性,融合出清晰的结果.

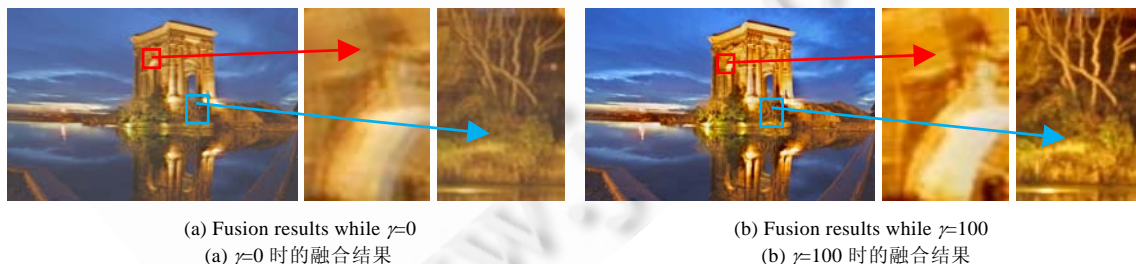


Fig.9 γ 's influence to our method for overcoming the small camera displacement

图 9 γ 取值对本文方法克服相机微小移动能力的影响

5.3 对含有噪声的LDR图像序列的实验结果及分析

图 10 中 LDR 图像序列是由高感光度相机(ISO1600)拍摄的,包含严重的 CCD 成像和 JPEG 格式压缩造成

的噪声.本文算法利用 $\gamma=0$ 时对该 LDR 图像序列作处理,可有效减小原图像的噪声(如图 10(b)所示).与目前经典 LDR 图像序列噪声抑制算法^[4]的处理结果(如图 10(c)所示)相比,本文算法去噪能力略显不足.但考虑到文献[4]中的算法需要事先得到精确的相机内部响应函数,并在图像序列中进行耗时的迭代计算,本文算法在效率和性能的平衡上仍具有一定优势.

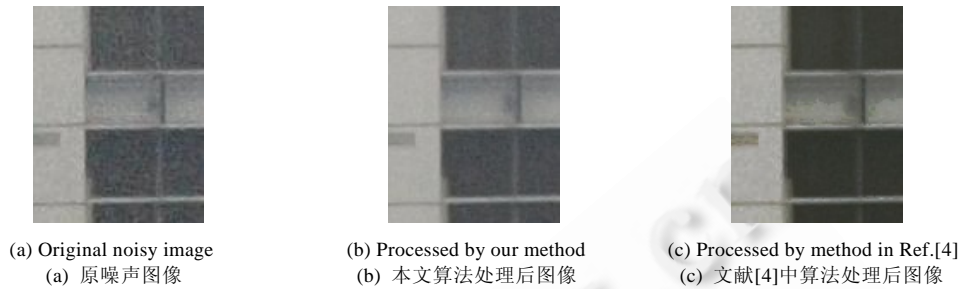


Fig.10 Comparison results to the noisy LDR image sequence

图 10 对含噪声 LDR 图像序列处理结果的对比分析

5.4 真实户外动态场景的LDR图像序列融合结果

利用本文算法在对实际户外动态场景的 LDR 图像序列进行分析时,需要同时考虑图像噪声、运动目标消除和相机微小移动对准等因素,故需要若干次尝试性地调整 γ 的取值并选择合适的基图像.图 11 中的真实户外动态场景包含运动汽车、行人、树枝随风摆动和高感光度成像(ISO1600)形成的噪声.经过若干次尝试,本文算法在 $\gamma=40$ 、选取第 4 幅图像为基图像(虚线框)时取得了较好的图像融合结果,如图 12(b)所示.



Fig.11 LDR image sequence of real outdoor scene with moving objects

图 11 真实户外动态场景的 LDR 图像序列

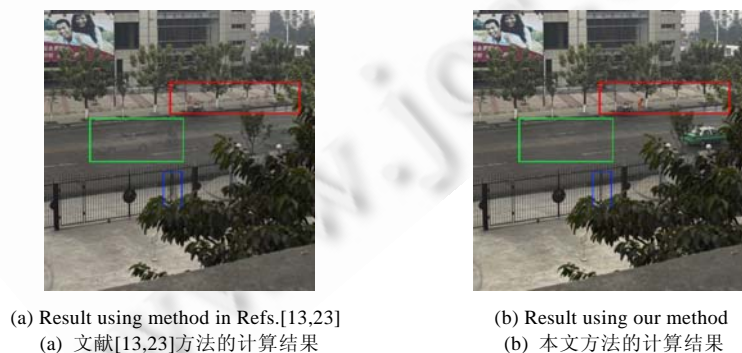


Fig.12 Fusion results of real outdoor scene with moving objects

图 12 真实户外动态场景的图像融合结果

5.5 计算效率对比分析结果

由于本文方法基于高效的曲线拟合算法,直接从 LDR 图像序列中逐像素地提取有效信息参与融合,其计算效率是非常占有优势的.表 1 中的计算时间统计结果表明,本文算法的计算速度要比传统 HDRI 技术快一个数量级.在统计 HDR 重建步骤的计算时间时,包含了图像配准^[5]、噪声抑制^[4]和重影消除^[3]这 3 个预处理步骤的计算时间,程序由 Qtpfsgui 源代码编写.

需要说明的是,由于不同计算平台的硬盘 I/O 时间差异很大,故表 1 中的数据不包含硬盘 I/O 时间.

Table 1 Comparative analysis of computational efficiency (s)
表 1 本文算法计算效率的对比分析 (秒)

LDR image size	HDR reconstruction	Tone mapping ^[10]	Time-Consuming of traditional method	Time-Consuming of our method
456×342×6	1.82	2.22	4.04	0.51
912×684×6	6.67	8.44	15.11	1.52
1824×1368×6	24.03	34.89	58.92	5.54
3648×2736×6	100.17	136.01	236.18	21.43

6 结 论

通过建立同一场景不同曝光值 LDR 图像序列中每个像素位置像素值曲线的数学模型,本文给出了一种简洁、高效的计算 LDR 图像序列中每个像素位置最佳曝光值的算法,并由此给出直接融合 LDR 图像序列的图像融合方法.该方法先对 LDR 图像序列中每个像素位置像素值曲线的二阶差商进行曲线拟合,再通过计算该曲线拐点的方式提取每个像素位置最佳曝光值下成像的像素信息,直接融合 LDR 图像序列.对实验结果的详细对比和分析表明,本文算法的计算效率约为传统 HDRI 技术的 10 倍,并在常规显示设备上能够达到与传统 HDRI 技术相当的高动态范围图像显示效果.同时,本文算法对原 LDR 图像序列中存在的图像噪声、运动目标、图像间微小平移和旋转等影响具有较好的鲁棒性.特别地,本文算法对图像序列中每个像素位置的核心处理仅涉及简单的局部运算,并不需要其周边像素的信息,非常适合利用现代可编程图形处理硬件(graphics processing unit)进行并行计算.

本文算法中存在一个全局参数 γ 用来平衡算法对噪声的抑制能力和对运动目标的鲁棒性,但场景中不同区域的噪声和所包含的运动目标情况有所不同,局部自适应决定 γ 的最优取值是我们下一步的研究方向.另外,本文算法的融合结果很大程度上依赖于基图像的选取,如何自适应地选取基图像也是我们将来的研究重点.

References:

- [1] Reinhard E, Ward G, Pattanaik S, Debevec P. High Dynamic Range Imaging. San Francisco: Elsevier Inc., 2006. 1–17
- [2] Jacobs K, Loscos C, Ward G. Automatic high-dynamic range image generation for dynamic scenes. IEEE Computer Graphics and Applications, 2008,28(2):84–93. [doi: 10.1109/MCG.2008.23]
- [3] Khan EA, Akyuz AO, Reinhard E. Ghost removal in high dynamic range images. In: Hayes M, Lanterman A, eds. Proc. of the IEEE Int'l Conf. on Image Processing. Washington: IEEE Computer Society, 2006. 2005–2008. [doi: 10.1109/ICIP.2006.312892]
- [4] Akyuz AO, Reinhard E. Noise reduction in high dynamic range imaging. Journal of Visual Communication & Image Representation, 2007,18(5):366–376. [doi: 10.1016/j.jvcir.2007.04.001]
- [5] Ward G. Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures. Journal of Graphics Tools, 2004,8(2):17–30.
- [6] Hua SG, Wang LD, Ou ZY. Alignment of images with varying exposure level from a same scene and HDR image construction. Journal of Computer-Aided Design and Computer Graphics, 2007,19(4):528–534 (in Chinese with English abstract).
- [7] Rempel AG, Trentacoste M, Seetzen H, Young HD, Heidrich W, Whitehead L, Ward G. Ldr2Hdr: On-the-Fly reverse tone mapping of legacy video and photographs. ACM Trans. on Graphics, 2007,26(3):391–396.
- [8] Seetzen H, Heidrich W, Stuerzlinger W, Ward G, Whitehead L, Trentacoste M, Ghosh A, Vorozcovs A. High dynamic range display systems. ACM Trans. on Graphics, 2004,23(3):760–768.

- [9] Durand F, Dorsey J. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. on Graphics*, 2002,21(3): 257–266.
- [10] Fattal R, Lischinski D, Werman M. Gradient domain high dynamic range compression. *ACM Trans. on Graphics*, 2002,21(3): 249–256.
- [11] Rovid A, Hashimoto T, Varlaki P. Improved high dynamic range image reproduction method. In: Fodor J, Prostean O, eds. *Proc. of the 4th Int'l Symp. on Applied Computational Intelligence and Informatics*. Washington: IEEE Computer Society, 2007. 203–207. [doi: 10.1109/SACI.2007.375511]
- [12] Song LM, Wang HQ, Chen C, Ye SQ, Gu WK. Tone mapping for high dynamic range image using a probabilistic model. *Journal of Software*, 2009,20(3):734–743 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/20/734.htm> [doi: 10.3724/SP.J.1001.2009.03371]
- [13] Debevec PE, Malik J. Recovering high dynamic range radiance maps from photographs. In: Owen SG, Whitted T, Mones-Hattal B, eds. *Proc. of the ACM SIGGRAPH*. New York: ACM Press, 1997. 369–378. [doi: 10.1145/258734.258884]
- [14] Grossberg MD, Nayar SK. What is the space of camera response functions. In: Flynn P, Dyer C, Perona P, eds. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. Washington: IEEE Computer Society, 2003. 18–20.
- [15] Battiato S, Castorina A, Mancuso M. High dynamic range imaging for digital still camera: An overview. *Journal of Electronic Imaging*, 2003,12(3):459–469. [doi: 10.1117/1.1580829]
- [16] Akyuz AO, Fleming R, Riecke BE, Reinhard E, Bulthoff HH. Do HDR displays support LDR Content? A psychophysical evaluation. *ACM Trans. on Graphics*, 2007,26(3):381–387. [doi: 10.1145/1276377.1276425]
- [17] Grossberg MD, Nayar SK. High dynamic range from multiple images: Which exposures to combine? In: Ma S, Shum H, Lee SW, eds. *Proc. of the Int'l Conf. on Computer Vision Workshop on Color and Photometric Methods in Computer Vision*. Washington: IEEE Computer Society, 2003. http://www1.cs.columbia.edu/CAVE/publications/pdfs/Grossberg_CPMCV03.pdf
- [18] Vavilin A, Jo K. Recursive HDR image generation from differently exposed images based on local image properties. In: Park Y, Choi HR, Yun WS, eds. *Proc. of the Int'l Conf. on Control, Automation and Systems*. Washington: IEEE Computer Society, 2008. 2791–2796. [doi: 10.1109/ICCAS.2008.4694234]
- [19] Li H, Manjunath BS, Mitra SK. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 1995,57(3):234–245. [doi: 10.1109/ICIP.1994.413273]
- [20] Russ JC. *The Image Processing Handbook*. Boca Raton: CRC Press, 2007. 373–377.
- [21] Farin G. *Curves and Surfaces for Computer Aided Geometric Design*. 4th ed., New York: Academic Press, 1997. 363–370.
- [22] Maronna R, Martin D, Yohai V. *Robust Statistics: Theory and Methods*. Chichester: John Wiley & Sons, Ltd., 2006. 87–114.
- [23] Reinhard E, Stark M, Shirley P, Ferwerda J. Photographic tone reproduction for digital images. *ACM Trans. on Graphics*, 2002, 21(3):267–276. [doi: 10.1145/566654.566575]

附中文参考文献:

- [6] 华顺刚,王丽丹,欧宗璞.同一场景不同曝光图像的配准及 HDR 图像合成. *计算机辅助设计与图形学学报*,2007,19(4):528–534.
- [12] 宋明黎,王慧琼,陈纯,叶秀清,顾伟康.基于概率模型的高动态范围图像色调映射. *软件学报*,2009,20(3):734–743. <http://www.jos.org.cn/1000-9825/20/734.htm> [doi: 10.3724/SP.J.1001.2009.03371]

附录 A. 定理 1 的证明

设 $g(x) \in C^2[a, b]$, 且有 $g(a)=g(b)=0$, 对 $g(x)$ 构造一次插值多项式 $P_1(x)$, 使 $P_1(a)=g(a), P_1(b)=g(b)$, 则 $P_1(x)=0$.

根据插值多项式的误差公式有

$$\begin{aligned}
 g(x) - P_1(x) &= \frac{1}{2} g''(\zeta)(x-a)(x-b), \quad \zeta \in [a, b] \\
 \Rightarrow |g(x)| &= \frac{1}{2} |g''(\zeta)| |(x-a)(x-b)| \leq \frac{1}{2} \max_{a \leq x \leq b} |g''(x)| |(x-a)(x-b)| \leq \frac{1}{8} \max_{a \leq x \leq b} |g''(x)| (b-a)^2 \quad (13) \\
 \Rightarrow \max_{a \leq x \leq b} |g(x)| &\leq \frac{1}{8} \max_{a \leq x \leq b} |g''(x)| (b-a)^2
 \end{aligned}$$

显然有,公式(11)中 $E(x)=f(x)-y(x;\boldsymbol{\alpha})\in C^2[a, b]$ 且 $\frac{d^2 E(x)}{dx^2} = \frac{d^2 [f(x)-y(x;\boldsymbol{\alpha})]}{dx^2} = e(x)$, 以及 $E(a)=f(a)-y(a;\boldsymbol{\alpha})=0$, $E(b)=f(b)-y(b;\boldsymbol{\alpha})=0$.

于是,可令 $g(x)=E(x)$,根据公式(13)可立即得到 $|E(x)|\leq \frac{1}{8} \max_{a\leq x\leq b} |e(x)|(b-a)^2$. 证毕. \square

附录 B. 更多直接融合的高动态范围图像



Fig.13 More results of our method

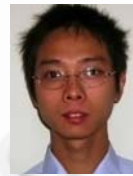
图 13 更多本文方法的计算结果



张军(1978—),男,甘肃临潭人,博士生,讲师,主要研究领域为计算机视觉,图形学.



戴霞(1976—),女,讲师,主要研究领域为计算机辅助几何设计,图形学.



孙德全(1983—),男,硕士,主要研究领域为计算机图形学,图像处理.



王邦平(1973—),男,博士,主要研究领域为计算机视觉,图形图像信息处理,模式识别.