arXiv:2303.13031v1 [cs.MM] 23 Mar 2023

# 利用新数据集和退化模型实现实用的 SDR 到 HDRTV 升格转换

Cheng Guo[1,2], Leidong Fan[3,2],ZiyuXue[4,1] 和 Xiuhua Jiang[2,11]中国传媒大学媒体融合与传播国家重点实验室[2]鹏城实验室 [3]北京大学[4]国家广播电视总局广播电视科学研究院

{guocheng,jiangxiuhua}@cuc.edu.cn, fanleidong@stu.pku.edu.cn, xueziyu@abs.ac.cn

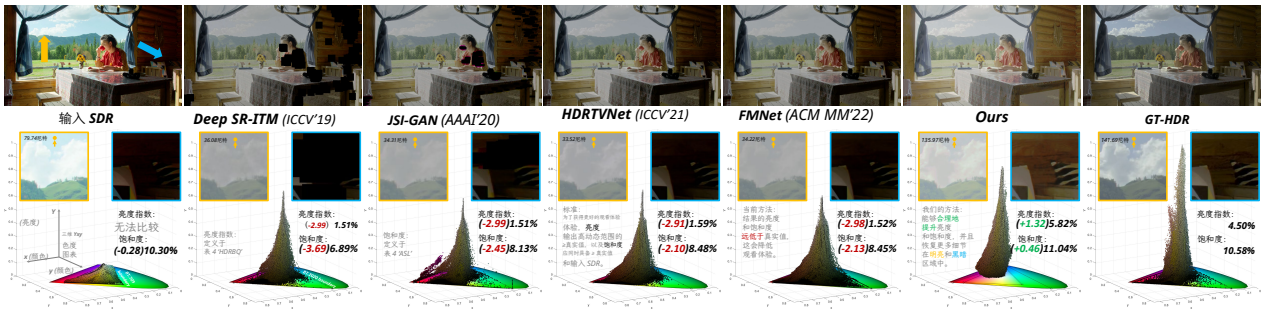



图 1. **将 SDR 内容升格转换，以供 HDR-WCG 电视显示**。为确保获得更好的观看体验 [1], , 我们使用 (1) 三维 *Yxy* 色度图(纵轴 *Y* 表示 HDR/亮度, *xy* 平面表示 WCG/颜色)，以观察 HDRTV 在 HDR与广色域体积方面的优势如何得到恢复；并使用 (2) 详细可视化 (黄色和蓝色方框)评估该方法的恢复能力。请注意，印刷版本中的 HDR 显得较暗，因为其大范围的亮度和颜色容器是由容量较小的 SDR 呈现的；正确可视化时则会呈现正常效果（图 7 上方）。尽管如此，从 (3)“亮度指数”和“饱和度”可以看出，**当前方法**的结果比真实值更**暗淡且饱和度更低**。

## 摘要

在媒体行业，当用户拥有 *HDR-WCG*（高动态范围-宽色域）电视，而市面上大多数现成视频素材仍为 *SDR*（标准动态范围）时，便产生了将 *SDR* 转换为 *HDRTV* 的需求。研究界已开始通过基于学习的方法应对这一低层视觉任务。然而，将当前方法应用于真实 *SDR* 内容时，往往会产生偏暗、色彩不饱和的结果，几乎无法改善观看体验。与其他以网络为中心的方法不同，我们认为这种不足源于训练集（*HDR-SDR* 配对数据）。因此，我们提出了新的 *HDRTV* 数据集（称为 *HDRTV4K*）和新的 *HDR* 到*SDR*退化模型。随后，我们使用这些数据训练亮度分段网络（*LSN*），该网络由一个全局映射主干和两个分别处理高、低亮度范围的 *Transformer*分支组成。我们还通过定制指标和主观实验更新了评估标准。最后，我们进行了消融研究，以验证方法的有效性。我们的工作可在

地址：*https://github.com/AndreGuo/HDRTVDM*。

## 1. 引言

图像的动态范围定义为记录的最大亮度与最小亮度之比。更大的亮度容纳范围使高动态范围（HDR）能够更好地呈现场景。在媒体和电影行业，先进的电光转换函数（EOTF）例如，感知量化/HLG [2], 以及广色域（WCG）RGB原色例如，BT。2020 [3]。

虽然广色域高动态范围显示器在消费市场日益普及，但由于制作流程成本高得令人望而却步，大多数商业影像素材仍采用标准动态范围（SDR），广色域高动态范围版本仍十分稀缺。因此，市场对将大量现有SDR内容转换为HDR电视服务所需格式的需求日益增长。这类SDR内容可能包含无法再现的场景，但更有可能存在由老旧成像系统和传输带来的瑕疵。这表明，SDR到HDRTV上转换是一项病态的低层视觉——

觉任务，因此研究界已开始采用基于学习的方法（ [4–9]等）。

然而，尽管他们使用了多功能网络（§2.1），我们发现，将真实的标准动态范围图像输入当前方法后，结果会变暗且饱和度不足（图1），这与 SDR 到 HDRTV 升格转换的感知动机相悖。正如 CVPR22 第 1 届视觉数据集理解研讨会所报告的那样， [10], 大多数方法以网络为中心，低估了训练集的影响。对于复原类低层视觉任务，训练集包含两个要素：标签真值本身的质量，以及真值到低质量图像退化模型（DM），即网络学习复原的对象。其他低层视觉任务正逐渐弥补这方面的忽视， [11–16], 但在基于学习的 SDR 到 HDRTV 升格转换中，这种情况仍然普遍存在。

并非偶然，我们发现，当前方法表现不佳的原因在于数据集。我们采用若干针对HDRTV定制的指标（表4）来评估当前训练集：（1）通过测量标注HDR的*HDR/WCG*范围等（表5），我们注意到其质量和多样性不足以激励网络生成令人满意的结果；（2）通过退化后标准动态范围图像的统计数据，我们发现当前HDR到SDR退化模型倾向于过度改变饱和度和亮度（见表6），因此网络会学到从SDR到HDR的劣化过程。为此，我们提出HDRTV4K数据集（§3.2），其中包含高质量且多样化的HDRTV帧（图4）作为标签。随后，我们采用3种新的HDRTV到SDR退化模型（§3.3），以避免上述不足，同时具备适当的退化能力（表6），使网络能够学习合理的复原能力。

随后，我们将该任务形式化为对完整亮度范围进行全局映射，并恢复低／高亮度范围。相应地，我们提出亮度分段网络（LSN，§3.1），其中全局主干网络和两个 Transformer 风格的 U 形网络 [17] 分支分别负责执行不同分段亮度范围（区域）所需的不同操作。

最后，正如 [18, 19], 所发现的，传统的基于距离的指标在纯重建任务（例如去噪）中表现良好，但在感知驱动的 HDR 重建中却不适用，因此我们采用更细粒度的指标（§4.2）和主观实验（§4.3）等评估标准。

我们的贡献主要有三点：**(1)** 强调并验证数据集对 SDR 到 HDRTV 任务的影响，这一影响长期以来被低估。**(2)**利用新型 HDR 电视数据集和 HDR到SDR退化模型来训练网络。**(3)**引入新的问题形式化，并据此提出新型亮度分段网络。

## 2. 相关工作

### 2.1. 术语与研究范围

已有许多与高动态范围相关的学习方法 [20]已被提出，但它们的预期应用有显著差异。

用。参见表1，“线性高动态范围”指场景参考高动态范围图像，旨在记录线性辐射度，供图形应用使用，例如基于图像的照明。[21, 22], 而“高动态范围电视”指我们所说的显示参考宽色域高动态范围 [2]格式。显然， ① 旨在合成可在标准动态范围显示器上观看的高动态范围画面，② 着重处理多个标准动态范围图像之间的错位，并针对新的成像流程设计；而 ③和 ④ 则都面向现有标准动态范围，但应用各不相同。由于许多研究是在学界尚未明确上述术语之前提出的，我们将某种方法归类为 SDR到HDRTV上转换（④，即我们的任务），前提是其目标高动态范围采用感知量化/HLG 电光转换函数和 BT.2020 广色域。

| | 任务名称 | From | 到单个 | 方法 |
|---|---|---|---|---|
| ① | 高动态范围风格 增强 | 单一 SDR | 增强后的 SDR | 例如。 [23–26] |
| ② | 多重曝光 HDR 成像 | 多种 SDR | 线性 HDR | 例如. [27–32] |
| ③ | SI-HDR* | 单一 SDR | | 例如. [33–37] |
| ④ （我们的方法） | 逆色调映射* 或 升转换 | | 高动态范围电视 帧 | [4–9] [38–53] |

*：SI-HDR：单图像高动态范围重建，iTM：逆色调映射。

表 1. 各类基于学习的高动态范围相关任务。绿色/青色/ 洋红色术语分别来自 [18]/我们的/ [7]术语表。

在我们的研究范围内（④），各方法基于不同理念设计网络：除例如需要语义信息的高光恢复外，我们任务的主要流程属于类似于图像润饰/增强的全局映射。基于此， [7] 将任务划分为 3 个步骤，并在全局部分使用 $1 \times 1$卷积， [43] 学习小型排序图像块之间的映射，而 [51] 则利用学习到的潜向量执行明确定义的逐像素映射。

特征调制也常用于引入全局先验信息。与 [7], 相比，它们在以下方面有所不同：先验类型 [47, 49], 、调制张量的形状 [48–50], 、调制机制 [9] 以及调制位置 [50]。

为协调连续的视频帧， [42]采用带有额外时间维度的3D卷积， [8] 以3帧为输入，并在U形网络瓶颈处部署多帧交互模块。此外，为了实现对齐， [44,49]采用可变形卷积，其偏移图由不同模块预测。

由于标准动态范围电视与高动态范围电视之间存在分辨率差异，一些方法 [4–6, 46, 48, 52] 联合进行超分辨率和升格转换。此外，一些方法借助非学习式预处理 [53] 或旁路[43, 45], ，而 [41] 则针对曝光不良的标准动态范围图像进行训练。

### 2.2. HDR电视数据集

它们使用的网络各不相同，目前公开的高动态范围电视训练集仅有 3 个（表。2）。所有数据集（包括我们的方法）都包含采用（D65白点）BT。2020 [3]RGB原色（色域）、感知量化 [54]电光转换函数和 $1000nit$ 峰值亮度的高动态范围电视帧，以及采用 BT。709 [55] 色域、伽马电光转换函数和 $100nit$ 峰值亮度的标准动态范围对应帧。

| 数据集（用途） | #pair | 分辨率 | HDR 格式 |
|---|---|---|---|
| KAIST [4] ( [5,48,52]) | 39840 | 160×160 | 无符号 16 位 MATLAB .mat YUV |
| 曾20 [6] | 23229 | UHD | H.265 Main 10 YUV |
| HDRTV1K [7] ( [9,47,50]) | 1235 | | 16 位 PNG RGB |
| HDRTV4K（我们的方法 新） | 3878 | 高清和超高清 | 16 位无损 .tif RGB |

表 2. 不同高动态范围电视数据集的状况：训练集部分。我们的高动态范围帧分辨率包括高清（1920×1080）和超高清（3840×2160），这些帧是从 >220 不同视频片段中人工挑选的，并封装为无损（*LZW* 或 *Deflate*）TIFF。其他数据则分别仅从 7 [4]/18 [7]电视演示片和一部调色影片中提取；其质量和多样性将在后文的表5 和图4中量化。

## 2.3. HDRTV 到 SDR 降级模型

大多数方法都遵循常见做法，将高动态范围（真实值）降级为输入 SDR（低质量）。降级模型（DM）至关重要，因为它决定了网络能够学到什么；即使在 SI-HDR 中，人们也相对更加关注这一问题 [16, 33, 35]。然而，在 SDR 到 HDRTV 领域对此鲜有讨论（表3）：

| DM | *YouTube* | *Reinhard* | 其他退化模型 |
|---|---|---|---|
| 用途 | [4, 5, 7–9, 40, 44, 46–50, 52] | [6, 38, 41–43] | ***2446a*** [45] |
| 数据集 | KAIST 和 HDRTV1K | Zeng20 | 等等。 [51, 53, 56] |

表3. 当前高动态范围电视到标准动态范围退化模型（显示映射）。「数据集」表示其中的标准动态范围内容是使用该显示映射从高动态范围内容退化而来。

***YouTube***表示 YouTube 对 BT.2020/感知量化1000 高动态范围内容应用的默认转换，以生成适用于标准动态范围的版本；***Reinhard/2446a*** 表示使用 *Reinhard* 色调映射算子 [57]/BT.2446[58] 方法*A*将高动态范围映射到标准动态范围。 [51]/ [53]/ [56]分别通过调色/*Habel*色调映射算子/另一种学习型网络将高动态范围退化为标准动态范围。

在其他复原任务中 [11–16], 退化模型的设计应使退化具有适当的范围和多样性，从而让网络学会恰当的复原能力并具备良好的泛化能力。因此，我们认为当前的退化模型不利于训练。具体而言，***YouTube*** 旨在为只有标准动态范围显示器的用户合成高动态范围画面，因此往往会增强/提高亮度和饱和度，网络反过来就会学着降低/削弱它们。此外，色调映射例如，***Reinhard*** 和***2446a*** 致力于尽可能多地保留高动态范围中的信息，其映射曲线单调递增（图5），且不会产生高光剪切，因此训练出的网络不太可能恢复过曝区域中的大量信息。后文将在表6、图1&6等中验证上述观察。

# 3. 提出的方法

我们的方法概览如图2所示。

## 3.1. 网络结构

**问题形式化** 有助于研究人员明确网络应恢复哪种退化。给定$\mathbf{x}$ 这个标准动态范围，$\mathbf{y}$ 则为高动态范围。以往方法 [7, 47, 49] 假设 $\mathbf{x}$ 和$\mathbf{y}$ 来自同一台相机拍摄的单个 RAW文件。

这适用于 SDR-HDR同步制作，在这种情况下，$\mathbf{x}$ 的质量也很高，因此升格转换（$\mathbf{y} = f(\mathbf{x})$)只需遵循特定风格并恢复较少的缺失信息。然而，由于我们的方法面向现有 SDR，我们假设 SDR（$\mathbf{x}$）及其虚构的 HDR 对应版本（$\mathbf{y}$）由不同的摄像机和后续处理流程同时拍摄，类似于 [18]。SDR处理流程的不完善使现有 SDR 容易受到宽容度有限、色域狭窄及其他退化（$\mathbf{x} = d(\mathbf{y})$)的影响。

一些研究 [7, 35] 假设 $d(\cdot)$ 按顺序引入，并为 $f(\cdot)$ 分配分层子网络。由于级联网络在实际应用中体量庞大，我们转而将特定退化建模为在不同亮度范围内更为显著。具体而言，SDR 中的过度曝光发生在高亮度范围 （$\mathbf{x}_h$），而低亮度范围（$\mathbf{x}_l$）更容易受到噪声等影响；与此同时，中亮度范围（$\mathbf{x}_m$）的质量相对较高。一些传统升格转换算子（ [59] 等）会分别处理分段的亮度范围，我们在此将这一思路引入深度神经网络（DNN）——LSN：

**亮度分段网络（LSN）**由一个作用于全亮度范围（$\mathbf{x}$）的主干和两个分别作用于$\mathbf{x}_l$ 与$\mathbf{x}_h$ 的分支组成，它们通过以下方式进行分段：

$$\mathbf{x}_l = max(0, \frac{t - \mathbf{x}}{t}),\ \mathbf{x}_h = max(0, \frac{\mathbf{x} - 1}{t} + 1) \quad (1)$$

其中$\mathbf{x} \in [0, 1]$。也就是说，亮度范围低于/高于阈值 $t$（经验设定为0。05）的部分会被线性映射为更显著的值 $[0, 1]$, , 如图2右上角所示。

分段后，$\mathbf{x}$、$\mathbf{x}_l$ 和$\mathbf{x}_h$ 需要采用不同的深度神经网络操作。首先，正如 [7] 等人（§2.1）所发现的，$f(\cdot)$ 大多属于全局（与像素无关）操作，类似于图像增强/图像修饰。因此，我们在全亮度范围（$\mathbf{x}$）上设置 5 个级联的无感受野 $1 \times 1$ 卷积层，构成**全局映射主干**。与 [7, 9, 47–50], 类似，我们添加一个调制分支（图3 左），将$\mathbf{x}$的全局先验纳入深度神经网络的决策中。

另一方面，$f(\cdot)$ 仍涉及非全局复原，需要聚合局部语义。具体而言，DNN 在$\mathbf{x}_l$ 中的任务类似于去噪和低照度增强，而在$\mathbf{x}_h$ 中的任务则类似于图像修复（由 [60, 61] 提出），在这一任务中，丢失的内容需要通过幻觉补全。这些低层视觉任务需要更大的感受野，因此我们选用 [17] 擅长建模长距离依赖的 Transformer 块，并将其排列为带有跳跃连接的编码器-解码器结构（U形网络）。**Transformer风格U形网络的分支**在$\mathbf{x}_l$ 和$\mathbf{x}_h$ 上采用相同结构，但参数不同。LSN 的细节见图3。

## 3.2. HDRTV4K数据集

设计完 LSN 后，我们需要更好的训练集来释放其能力。首先从 labelHDR 的质量入手：

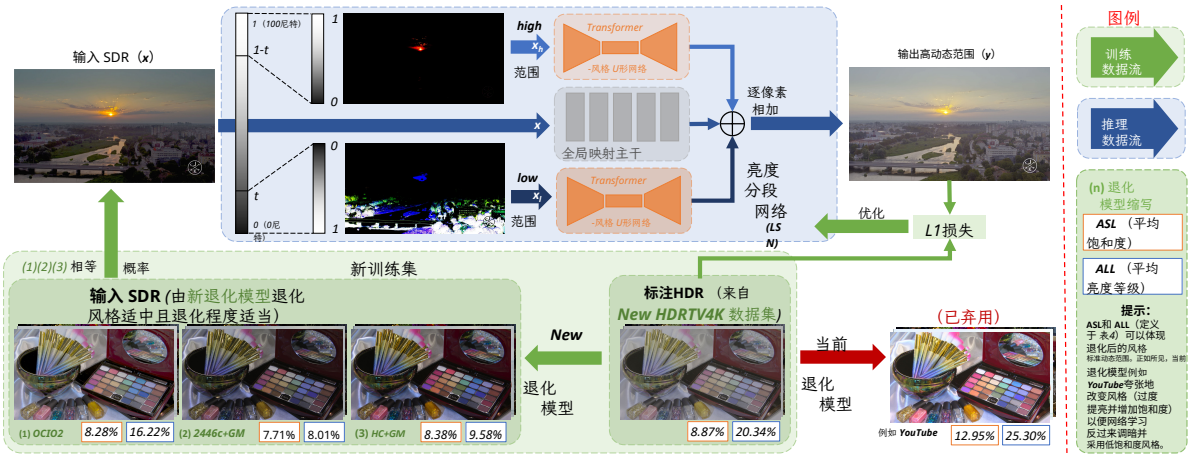



图 2. 本文的整体概览。我们的亮度分段网络（LSN，§3.1）基于新颖的问题表述进行设计，随后使用所提出的 HDRTV4K数据集（§3.2）中的标注HDR进行监督训练，并以新颖退化模型（DMs，§3.3）退化后的输入 SDR作为输入。我们的 LSN、HDRTV4K数据集和退化模型的**主要关注点**分别是：恢复暗部和亮部区域、提升标签真值 GT-HDR 的质量（表5）和多样性（图4），以及确保低质量标准动态范围具有恰当的风格和退化（表6 和图5）。

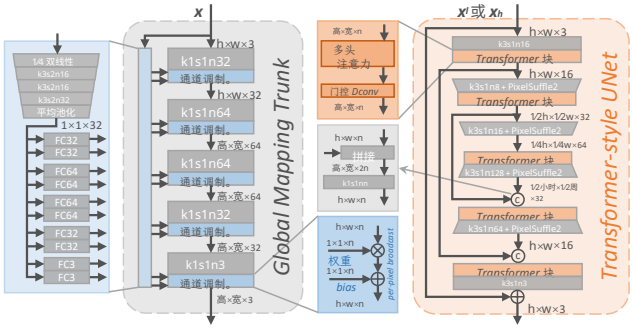



图 3. LSN 模块的详细结构。“k3s1n16”表示卷积层，其 3×3 卷积核、步幅=1、输出通道数=16分别为相应数值；“FC32”表示输出通道数为输出通道数=32的全连接层。为了使 LSN 更轻量（参数量 [17]），我们在 U 形网络的每个编码 / 解码层级仅使用 1 个 Transformer 块，以便将其应用于 4K 分辨率。=325

**动机**。在 §1中，我们论证了当前数据集中标注HDR的质量问题。此处依据以下考量，通过表4中的 10 项指标，从 3 个方面对其进行评估：

首先，HDR/WCG 范围越大，网络就越有可能学习超出 SDR 能力范围的先进 WCG-HDR 色域中的像素。同时，帧内多样性越高，网络就能在较小批量中学习得越好，即更大的 **SI/CF/stdL**分别意味着网络可以学习到更多样的高频模式、更丰富的色彩体积和更大的亮度对比度。

此外，风格蒸馏 [51] 已成为类似任务中影响方法性能的重要因素，例如图像增强/图像修饰，因此我们还增加了 3 项指标，用于量化整体风格以及标注 HDR。请注意，网络学到的风格在很大程度上会受到退化模型（显示映射）的影响，后文将对此进行讨论。

| | |
|---|---|
| **HDR/WCG 范围的指标 ↓** | |
| **FHLP** | **高光像素**占比：高光像素的空间比例<br>即其归一化亮度 $Y = 0.2627R + 0.6780G + 0.0593B > 0.1(100nit$，SDR 的峰值亮度。) |
| **EHL** | **高光**范围：平均像素 ($i$) 距离<br>高动态范围的亮度及其剪裁至-100$nit$版本[1]:<br>$\frac{1}{n}\sum_{i=1}^{n}\sqrt{[Y_i - clip(Y_i)]^2}$，$clip(x) = clamp(x,\ 0,\ 0。1)$ |
| **FWGP** | **广色域像素**占比：广色域像素的空间比例，即。<br>其 $[x, y]$ 坐标位于 BT.2020 内，但位于<br>SDR 的 BT.709 色域之外，在 $Yxy$ 色度图表中。 |
| **EWG** | **广色域覆盖范围** [62]：平均像素距离（介于<br>宽色域高动态范围及其色域硬裁剪 [63]版本[2]:<br>$\frac{1}{n}\sum_{i=1}^{n}\|\mathbf{S}_i - HC(\mathbf{S}_i)\|_{\ 2}$，$\mathbf{S} = [X,^{Y,\ Z]^{\mathrm{T}}}$ |
| **帧内多样性指标 ↓（均基于方差）** | |
| **SI** | **空间信息**：像素的标准差<br>附录 6 中定义的经 Sobel 滤波的帧 [64]。 |
| **CF** | **颜色丰富度**：定义见 [65]。 |
| **stdL** | 帧中所有像素的**亮度**标准差。 |
| **整体风格指标 ↓** | |
| **ASL** | **平均饱和度水平**：归一化像素平均长度<br>高动态范围色度分量 $\mathbf{C} = [C_t, C_p]^{\mathrm{T}}$ [66]:<br>$\frac{\sqrt{2}}{n}\sum_{i=1}^{n}\|\mathbf{C}_i\|_2$，$\mathbf{C} \in [-0.5, 0.5]$ |
| **ALL** | **平均亮度级别**：$Y$in**FHLP**的像素平均值。 |
| **HDRBQ** | **高动态范围亮度量化** [67],涉及视觉显著性。 |

1: **EHL** 用于补偿某些情况例如，全是101$nit$ HDR帧且<FHLP>占比为100%，但高光程度较低。
2: 色域硬裁剪的公式可见于（$HC(\cdot)$）在公式5。

表4. 标注HDR的质量和多样性通过上述10项指标从3个方面进行衡量（均呈正相关），结果见表。5. 完整说明请参见补充材料。

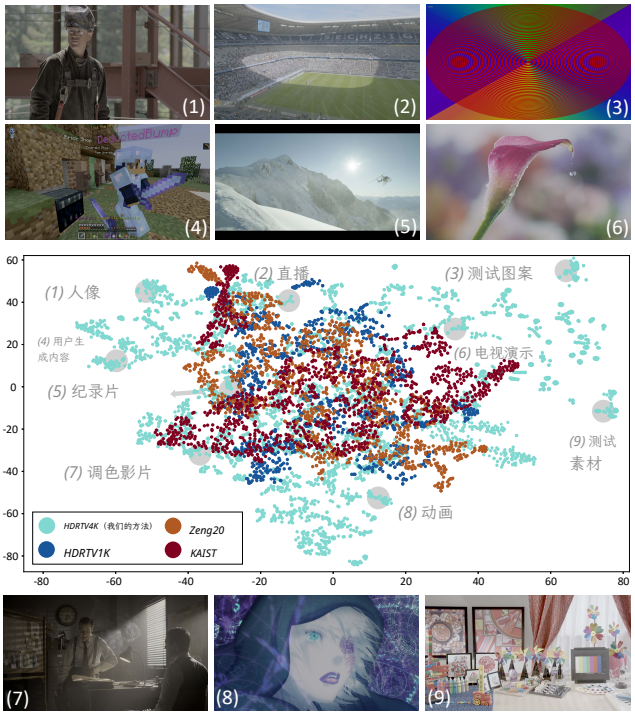
**我们的工作**。表格中的统计数据5 证实了当前数据集的不足，即较低的帧内多样性，而最重要的是，*HDR/WCG* 范围较窄，这使得其网络无法生成真实的 HDR-WCG 体积。为此，我们提出了一个新的数据集 HDRTV4K，其中包含 3878 个 BT.2020/感知量化1000（表格2）HDR 帧，且质量更高。

具体而言，这些帧是从各种开放内容（ [68–72]等）中手动提取并对齐的，具有更大的*HDR/WCG*范围、更高的帧内多样性

| 指标 / 数据集 | *HDR*范围 | | 广色域范围 | | 帧内多样性 | | | 整体风格 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **FHLP** | **EHL** | **FWGP** | **EWG** | **SI** | **CF** | **stdL** | **ASL** | **ALL** | **HDRBQ** |
| KAIST [4] | 1.5250 | 0.2025 | 5.4771 | 0.1104 | 1.9372 | 5.9485 | 0.9597 | 8.9087 | 17.2854 | 1.8597 |
| Zeng20 [6] | 0.0197 | 0.0012 | 0.4792 | 0.0034 | 0.1231 | 4.2048 | 0.3146 | 3.8061 | 6.0805 | 0.3781 |
| HDRTV1K [7] | 1.2843 | 0.1971 | 2.9089 | 0.1633 | 2.2378 | 11.0722 | 1.8006 | 10.9414 | 15.1626 | 2.7970 |
| HDRTV4K（我们的方法） | 5.3083 | 0.9595 | 2.6369 | 0.5123 | 3.5508 | 10.5882 | 3.4837 | 9.8274 | 21.1996 | 5.1593 |

表 5. 标注HDR帧的质量，以表中 10 项指标的帧均值表示。4。更大的高动态范围/广色域范围有助于网络在标准动态范围色域之外生成更多像素，而帧内多样性则有助于提升网络的泛化能力。所有数值均为百分比（%），我们将最有利于训练的项目标出。图4进一步展示了它们的多样性。

以及合理的风格（表5）。请注意，我们使用 DaVinci Resolve 对部分原始素材重新调色，以在多样性上增加一些扰动。我们的数据集缩略图及其帧分布对比见图4。




多样性比较：我们的 HDRTV4K数据集与其他数据集。这里，每个二维坐标都是单帧 10 维向量（包含表中 10 项指标）的投影，采用 t-SNE [73] (`exg=3`、`prep=50`)。我们还展示了数据集中 HDR 帧的缩略图（1）至（9），并用灰色圆圈标出它们对应的二维坐标。如图所示，我们的数据集涵盖了分布更广的帧，即，为网络学习提供了更多样化的场景。

## 3.3. 新退化模型

获得质量更高的标注HDR后，我们将重点放在退化模型（DM）上，它决定了复原网络需要学习的内容。正如§2.3中所述，当前的退化模型在不合理的饱和度和亮度变化以及微弱的过度曝光退化方面表现不佳，**ALL**验证了这一点，**ASL**、**FOEP**，见表6。也就是说，***YouTube/2446a*** 往往会同时增加或降低**ASL** & **ALL**，因此网络会学会同时降低或过度提高饱和度，并不合理地调暗或提亮。此外，***2446a*** 和***Rienhard*** 提供的**FOEP** 较少，不利于网络形成对过度曝光的幻觉能力。

| 经退化处理的标准动态范围 | | FOEP[1] | ALL | ASL[2] | SI | CF |
|---|---|---|---|---|---|---|
| Current DM | ***2446a*** | 0.181 | 5.880 | 3.624 | 6.041 | 6.573 |
| | ***Reinhard*** | 1.264 | 21.887 | 7.442 | 14.147 | 12.568 |
| | ***YouTube*** | 5.439 | 28.219 | 14.641 | 18.545 | 25.225 |
| Ours DM | ***2446c+GM*** | 1.739 | 11.669 | 10.183 | 12.503 | 19.391 |
| | ***HC+GM*** | 4.252 | 14.062 | 10.377 | 15.090 | 20.146 |
| | ***OCIO2*** | 1.580 | 18.887 | 9.977 | 13.578 | 18.052 |
| 标准↑: 越小越好 | | 保持[3] | 更小 | kept | - | - |
| 源HDR（表5）[4] | | 5.308 | 21.200 | 9.827 | 3.551 | 10.588 |

[1]: **FOEP**: **过度曝光像素**比例：**过度曝光**像素的归一化亮度空间比例 $Y=1$。[2]: 对于标准动态范围，**ASL** 也按表4计算，但使用 $\mathbf{C} = [C_b, C_r]^{\mathrm{T}}$ [55] 而非 $[C_t, C_p]^{\mathrm{T}}$。[3]: 假设高动态范围的高光（$>100nit$）部分在标准动态范围中都应被裁剪至 1（过度曝光）。[4]: 灰色 表示容器差异，即即使对于退化良好的标准动态范围图像，高动态范围/标准动态范围的指标也应有所不同。

表 6. 给定 HDRTV4K 数据集的标注HDR后，标准动态范围图像经不同退化模型退化后，其统计特性会发生变化。结果以百分比表示；根据§不利 / 中性 / 有利于 [illegible] 己，见2.3 & 3.3。如图所示，我们的退化模型提供了适当的过度曝光（**FOEP**），且不会造成不当的亮度（**ALL**）和饱和度（**ASL**）变化，便于网络学习。不同退化方式生成的标准动态范围图像示例见补充材料。

这促使我们采用/利用新的退化模型，确保退化程度适当且风格方面没有缺陷：

(1) ***OCIO2***：*OpenColorIOv2* [74], 从 BT.2020/感知量化1000 高动态范围容器转换到 BT.709/伽马标准动态范围的商业方法，此处通过三维查找表 (LUT) 实现。

(2) ***2446c+***色域映射：在 Rec.2446 [58]，的 Method C 色调映射基础上进行修改，随后进行色域映射。这里，每个高动态范围编码值 $\mathbf{E}' = [R', G', B']^{\mathrm{T}}$（对于非线性值，我们使用上标′）通过 PQ 电光转换函数线性化为 $\mathbf{E}$，然后转换到 $Yxy$。接着，$Y$ 进行全局映射至 $Y_{SDR}$，而 $xy$ 则保持不变，随后与 $Y_{SDR}$ 重新合成，形成$\mathbf{E}_{SDR}$，之后再形成$\mathbf{E}'_{SDR_{709}}$。我们的修改体现在：

$$Y_{SDR} = clamp(TM_{2446c}(Y), 0, 100nit) \qquad (2)$$

其中 $TM_{2446c}(\cdot) \in [0, 118nit]$ 是 [58]（图5 中的原始全局映射曲线（虚线）。也就是说，我们将原始输出范围裁剪至 $100nit$，以生成更多 **FOEP**，而非采用线性缩放，因为它与 DM***2446a***存在相同的缺点。色域映射将在公式4和5中介绍。

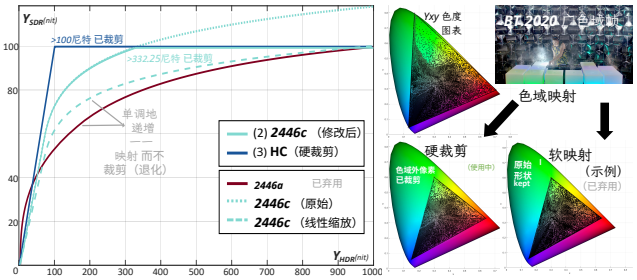



图 5. 所提出的退化模型（DMs，§3.3）的关键思想之一：应适当裁剪高动态范围的亮度（左）和颜色（右）色域，以便网络能够学习相应的复原。另一个思想是在退化过程中保持合理的亮度和颜色/饱和度风格，请参阅表6。

(3)***HC+GM***：硬裁剪与色域映射，这是一种容器转换，标准动态范围部分保持不变，而高动态范围/广色域色域中的所有亮度/颜色均进行硬裁剪。对于亮度部分，我们将所有 >100$nit$ 像素裁剪至 100$nit$：

$$\mathbf{E}_{SDR} = 10 \times clamp(\mathbf{E}, 0, 0.1),\ \mathbf{E} \in [0,1] \qquad (3)$$

其中，0.1 对应于归一化线性高动态范围 $\mathbf{E}$ 中的 100$nit$，而 $\times 10$ 用于适配容器差异（标称峰值 1 表示 1000$nit \rightarrow$ 100$nit$）。

到目前为止，$\mathbf{E}_{SDR}$ 仍处于 BT.2020 色域内。因此，我们加入色域映射（**+GM**）：首先，将颜色空间从 BT.2020 RGB原色转换到 BT.709：

$$\mathbf{E}_{SDR_{OOG}} = \mathbf{M}\mathbf{E}_{SDR},\ \mathbf{M} = \begin{bmatrix} 1.6605 & 0.5876 & -0.0728 \\ -0.1246 & 1.1329 & 0.0083 \\ -0.0182 & -0.1006 & 1.1187 \end{bmatrix} \qquad (4)$$

其中，$\exists\, \mathbf{E}_{SDR_{OOG}} \notin [0,1]$ 即经过颜色空间转换到较窄色域后，色域外像素（OOG）将落在有效范围之外。处理色域外像素是色域映射的关键。我们不采用软映射 [63, 75]（[76]的显示映射（DM）所使用的方法），而是尽可能保留广色域信息并将其映射到较窄色域，采用硬裁剪，将所有色域外像素裁剪到 BT.709 边界（图5）：

$$\mathbf{E}_{SDR_{709}} = clamp(\mathbf{E}_{SDR_{OOG}}, 0, 1) \qquad (5)$$

然后，$\mathbf{E}^{SDR_{709}}$ 将式（2）和式（3）中的结果转换为 SDR 编码值$\mathbf{E}'_{SDR_{709}}$，使用 BT.1886 [77] 光电转换函数（近似采用伽马 *2.22*）和质量因子为=80的 8 位 JPEG 压缩。

我们的理念是，大多数定义明确的操作都可以简化为 $clamp(\cdot)$ 函数：只有在亮度和色彩体积被裁剪时，显示映射中才会发生多对一映射，训练后的网络才能学习相应的恢复能力。此外，**ALL** & **ASL**在表6中表明，这些退化模型生成的标准动态范围内容具有合理的风格，确保网络不会学习到风格劣化。

# 4. 实验

**训练细节**。对于每个 HDR 帧，随机调整其大小至 [0.25x,1x], 6，然后从随机位置裁剪尺寸为 $600 \times 600$ 的图像块。随后，每个 6×3878 HDR 图像块都会以**相等的概率**通过 3 种提出的退化模型之一退化为标准动态范围图像，并再次以质量因子 QF =75 存储为 JPEG。

所有 LSN 参数均采用 Kaiming 初始化，并使用 $l_1$ loss 和 AdaM 进行优化，学习率从 $2 \times 10^{-4}$ 开始，每经过 $10^5$ 次迭代减半，直至 $4 \times 10^5$。

## 4.1. 标准与配置

**标准**。由于 SDR 到 HDRTV 上转换的目标是满足人类感知，为了使高动态范围电视相比标准动态范围电视实现“跨越式”提升，评估方法时应考察以下几点：[1]***C1***：结果在视觉上是否令人愉悦，包括亮度（***C1A***）和色彩外观（***C1B***）等；***C2***：HDRTV 在 HDR/WCG 体积方面的优势是否得到恰当恢复和利用；***C3***：明亮区域和暗区是否得到恢复或至少有所增强；以及***C4***：是否避免了伪影。然而，正如 [18,19], 当前的基于距离的指标（如峰值信噪比）所表明的那样，此类指标并不适用，因为图像是否具有吸引力，主要取决于所学习数值分布的准确性，而非***C1-4***。因此，我们采用表7中的新评估标准。

| | 视觉效果见图1&6 | 指标 见表8 | 主观实验 (§4.3) |
|---|---|---|---|
| ***C1A*** | - | ALL，HDRBQ | 总体评分 & 选择归因 |
| ***C1B*** | | ASL | |
| ***C3*** | 黄色和蓝色方框 | - | |
| ***C4*** | 详细可视化 | | |
| ***C2*** | 三维 *Yxy* 图 | FHL（WG）P，EHL（WG） | - |

表 7。各项标准***C1-4*** 的评估方式：**详细可视化、精细定制指标**和**主观实验**。

**对比方法**。如表8所示，我们的方法与 6 种基于学习的方法进行比较 [4–9],，以及两款商业软件*DaVinci* 和*Nuke*。来自联合超分辨率方法的结果 [4–6] 在比较前经过双线性下采样。请注意，基于学习的方法未使用我们的数据集重新训练，因为我们将数据集视为影响性能的关键归因因素。

**测试集**。我们选取了 12 段 4K标准动态范围视频，时长为 10$s$，并且 50$fps$。其中 3 段是通过***YouTube*** 显示映射（§2.3）从对应的 HDR-GT 降级得到，其余 9 段则由 HDR 序列编号***2***、13、15、20、21、22、42、44、55 的视频进行*2*、*13*、*15*、*20*、*21*、*22*、*42*、*44*、*55*，来自 [78]。实验公平性得以保证，因为测试集中的两种退化模型（标准动态范围-高动态范围关系）均未被我们的网络见过（这甚至让使用***YouTube*** 显示映射 [4,5,7–9]训练的方法占了优势）。

## 4.2. 结果

所有竞争者的结果如表8 和图6所示。

**指标**。从表8 可知，采用***YouTube*** 显示映射训练的方法往往会产生较低的饱和度（**ASL**）和亮度（**ALL**& **HDRBQ**），且其**ASL** 甚至低于输入 SDR。此外，SR-ITM-GAN [6] 恢复的 HDR与广色域体积最小，因为其标注HDR（来自 Zeng20数据集）的*HDR/WCG*范围最小（表5）。请注意，当前方法能生成足够的**FWGP**，但从图三维*Yxy*色度图图中6 可知，其广色域

| 方法（网络） | 网络的训练方式 数据集（真实值） | DM | （恢复率%，真实值为100%）*HDR/WCG* 色域的恢复方式 FHLP | EHL | FWGP | EWG | （偏移率%，真实值为0%）整体风格 ASL | ALL | HDRBQ | 传统指标 PSNR | SSIM | ΔE | VDP3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 输入 SDR | - | | 0 | 0 | 0 | 0 | 6.570 | 10.76 | - | 23.92分贝 | 0.8861 | 44.97 | 6.571 |
| Deep SR-ITM [4] | KAIST | *YouTube* | (14.03)0.2323 | (9.70)0.372 | (175.2)1.0964 | (71.96)0.172 | (-20.33)5.485 | (-15.06)9.580 | (-72.80)1.428 | 26.59分贝 | 0.8115 | 32.54 | 6.917 |
| JSI-GAN [5] | | | (12.33)0.2041 | (3.55)0.136 | (213.1)1.3334 | (88.23)0.212 | (-16.62)5.741 | (-14.36)9.659 | (-79.20)1.092 | 27.87分贝 | 0.8420 | 30.23 | 7.452 |
| SR-ITM-GAN [6] | Zeng20 | *Reinhard* | (8.04)0.1332 | (14.33)0.550 | (00.00)0.0000 | (00.00)0.000 | (-7.04)6.400 | (-30.37)7.853 | (-57.12)2.251 | 28.04分贝 | 0.8831 | 24.78 | 6.707 |
| HDRTVNet [7] | HDRTV1K | *YouTube* | (18.59)0.3078 | (16.29)0.625 | (388.8)2.4334 | (28.82)0.069 | (-15.51)5.817 | (-13.47)9.759 | (-69.65)1.593 | 30.82分贝 | 0.8812 | 27.58 | 8.120 |
| KPN-MFI [8] | own | | (1.17)0.0193 | (0.02)0.001 | (392.9)2.4592 | (15.82)0.038 | (-21.74)5.388 | (-16.10)9.462 | (-82.47)0.920 | 29.37分贝 | 0.8746 | 27.47 | 7.785 |
| FMNet [9] | HDRTV1K | | (13.67)0.2264 | (12.35)0.474 | (396.5)2.4813 | (91.29)0.220 | (-16.91)5.770 | (-13.48)9.758 | (-71.24)1.510 | 30.91分贝 | 0.8855 | 27.16 | 8.069 |
| LSN（我们的方法） | HDRTV4K | 我们的方法 | (256.6)4.2509 | (71.96)2.599 | (109.8)0.6873 | (150.0)0.361 | (+9.25)7.522 | (+81.12)20.42 | (-27.04)3.829 | 24.47分贝 | 0.8310 | 37.84 | 8.130 |
| DaVinci | - | ×3 | (54.96)0.9103 | (43.16)1.655 | (310.0)1.9399 | (85.22)0.205 | (-21.36)5.414 | (-21.42)8.863 | (-39.23)3.190 | 26.39分贝 | 0.8918 | 35.47 | 8.528 |
| Nuke | | | (112.1)1.8565 | (24.70)0.384 | (00.00)0.0000 | (00.00)0.000 | (-27.53)4.990 | (+17.93)13.30 | (-51.19)2.562 | 20.87分贝 | 0.7273 | 64.29 | 7.479 |
| HDR-GT（参考文献。） | | | (100.0)3.835 | (100.0)0.6258 | (100.0)0.240 | | (0.00)6.885 | | (0.00)5.248 | - | - | - | - |
| 消融研究 ↓ | | | | | | | | | | | | | |
| LSN（我们的方法） | HDRTV4K | *YouTube* | (13.09)0.2168 | (6.63)0.254 | (401.3)2.5118 | (59.81)0.144 | (-14.68)5.874 | (-13.46)9.760 | (-76.89)1.213 | 30.15分贝 | 0.8858 | 28.04 | 7.902 |
| | HDRTV4K | *Reinhard* | (9.06)0.1501 | (19.14)0.734 | (00.00)0.0000 | (00.00)0.000 | (-5.74)6.490 | (-25.65)8.387 | (-52.84)2.475 | 27.70分贝 | 0.8436 | 26.03 | 6.832 |
| | Zeng20 | 我们的方法×3 | (109.6)1.8147 | (9.00)0.345 | (0.95)0.0593 | (0.76)0.002 | (+3.36)7.117 | (+66.64)18.79 | (-59.52)2.124 | 25.17分贝 | 0.8179 | 28.58 | 7.874 |
| | HDRTV1K | 我们的方法 ×3 | (177.5)2.9405 | (52.49)2.013 | (279.5)1.7494 | (70.43)0.169 | (+6.56)7.337 | (+52.30)17.18 | (-34.83)3.420 | 24.30分贝 | 0.8351 | 38.03 | 8.002 |

表 8. 指标。我们使用**精细定制指标**(第 *3-9* 列，定义见表4) 来评估标准***C1*&*2*** (§4.1)：第 3-6 列 中的指标越显著，特定方法恢复 HDR 与广色域体积的效果就越好（***C2***）。此外，第 8-10 列 越接近真实值，表示亮度和色彩外观越相似。请注意，我们允许第 *8-10* 列略高于真实值，这意味着结果 HDR 的视觉体验合理地比真实值更令人愉悦（***C1***）。据此，我们突出显示了不利于观看体验的结果。

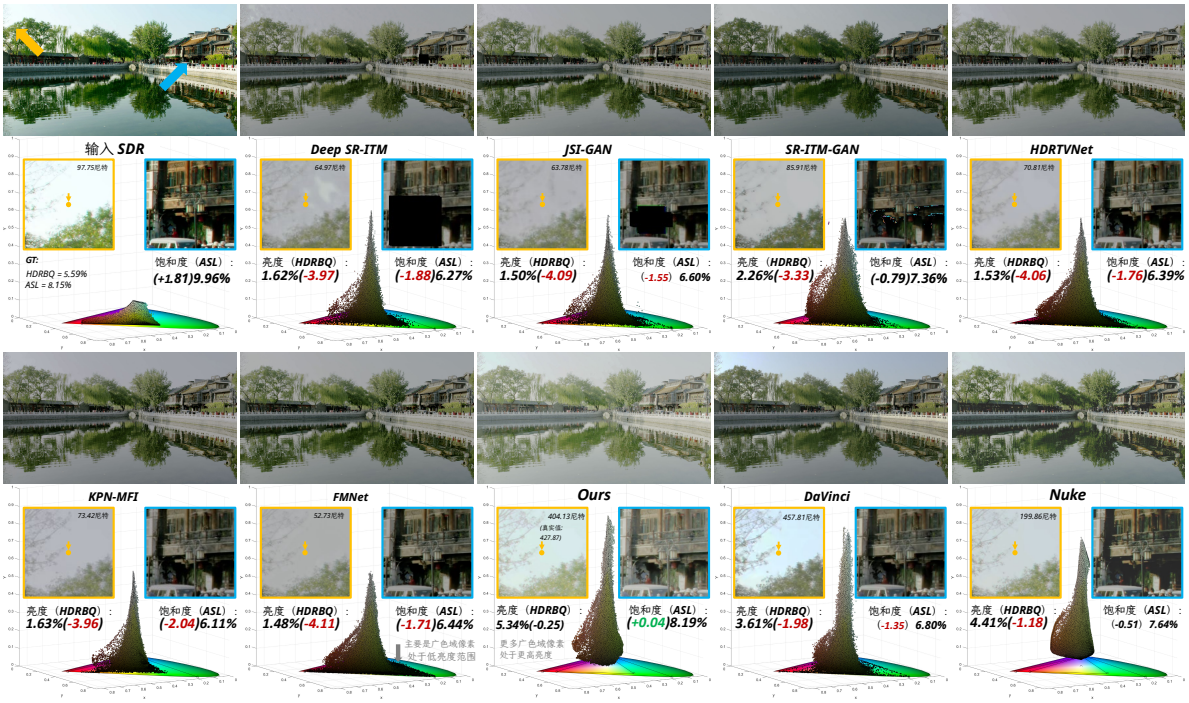



图 6. 视觉效果。我们通过三维*Yxy*色度图比较来评估 HDR/WCG 体积的恢复情况（***C2***），并展示明亮区域（黄色箭头，标示同一像素的亮度）和暗部区域（蓝色箭头）的详细可视化，以评估方法的恢复能力（***C3***）。请注意，不能得出“标准动态范围比高动态范围更鲜艳”这一结论，因为在印刷版本中，高动态范围看起来会比标准动态范围更暗（见图1中的解释）。因此，我们转而查看**ASL** & **HDRBQ** (第 *8-10* 列) 和后续的主观实验，以确认高动态范围是否确实更暗（***C1***）。更多场景的比较将在补充材料中提供。

像素主要分布在低亮度范围内，这意味着它们对观看体验帮助不大。我们的方法能够恢复足够的 HDR/WCG 体积，同时合理地增强亮度和饱和度。[1]

**视觉效果**。图中展示了各方法的恢复能力。6：当前方法在亮区（黄色）和暗区（蓝色）的表现都不佳。具体而言，使用***YouTube***显示映射训练的方法会产生调暗的（**HDRBQ**）和饱和度不足的（**ASL**）结果，甚至在相同位置的亮度也低于输入 SDR（黄色方框中的箭头）。这印证了 §2.3中的发现：如果显示映射倾向于提亮并过度增加饱和度，网络也会学会调暗并降低饱和度。此外，我们的方法能在亮区和暗区恢复更多细节，且风格更佳。

[1]我们还提供传统的峰值信噪比（PSNR）、结构相似性（SSIM）、 ΔE（颜色差异[79]）和 VDP3（HDR-VDP-3 [80]），但这些指标主要反映输出值与真实值的接近程度（例如，结果比真实值更暗（例如，Deep SR- ITM [4]）或更鲜艳（我们的方法），得分都可能同样偏低），因此对我们的评估毫无帮助。首次发现这一现象的是 [18, 19], , 并将在补充材料中进一步解释。

## 4.3. 主观实验

目前，很少有主观研究 [19, 43, 59, 81–83] 针对 SDR到HDR转换流程（更多是针对不同高动态范围内容）而设计。与 [19], 类似，我们评判输出高动态范围是否优于原始标准动态范围。具体而言，如图7所示，我们使用两台并排放置的支持标准动态范围和高动态范围电视的 *SONY KD85X9000H* 显示器，一台校准为 100$nit$/BT.709 标准动态范围，另一台校准为感知量化 1000$nit$/BT.2020 高动态范围。每组（输入）标准动态范围—（输出）高动态范围视频对显示两次，间隔 3$s$ 灰度级，并按照以下 10$s$ 方式评分：要求每位参与者在 -5（标准动态范围好得多）到 0（相同）再到 5（高动态范围好得多）的范围内连续评分，同时至少选择一项评分依据（图7下方）。该过程重复 9（参与者人数）×9（比较对象数量）×12（片段数量）次。

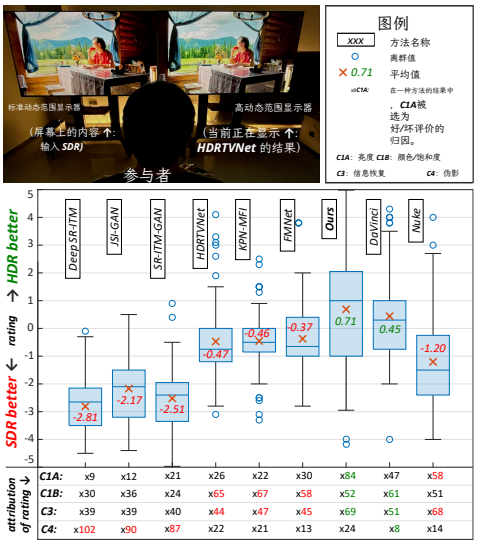



图7. 主观实验的环境和结果。对两台显示器进行了不同的校准，以确保标准动态范围和高动态范围均能**正确显示**。结果以四分位图呈现。

结果表明，只有两种方法（我们的方法和*DaVinci*）的评价优于输入 SDR。对于方法 [4–6], , 其主要“归因”（图7下方）是伪影（***C4***，见图6中的蓝框）。对于无伪影的“***YouTube***-显示映射”方法 [7–9],, 其评分略低，主要是因为饱和度较低（***C1B***），且无法恢复信息；这与表8和图6一致。从“归因”中我们还注意到，我们的方法在观看体验（***C1***）和信息恢复（***C3***）方面得分更高。此外，我们的“不良案例”在于恢复出的大片暗区带有固有噪声等，且被放大而未消除。这就是为何与商业方法***C4***(×24)相比，我们在*DaVinci* 和*Nuke*中获得了更多勾选项。

## 4.4. 消融研究

**关于显示映射**。 当显示映射改为 ***YouTube***时，表8显示 **FHLP**、**EHL**、**ASL**、**ALL**的指标显著下降和 **HDRBQ**，而图8确认结果与那些‘***YouTube***-显示映射’方法类似，即。观看体验更差，恢复能力更弱。此外，使用***Reinhard*** 显示映射（不包含裁剪）时，结果中的高光区域保持未学习状态。

**在数据集上**。 这里，我们使用原始退化模型，但使用其他数据集中的标注HDR。在图8 黄色方框中，我们的退化模型促使网络输出更高的亮度，但由于 Zeng20 的 HDR范围最小，即这些高光在标注HDR中并不存在，因此我们的 LSN 无法“识别”它们，进而产生伪影。由于该数据集的广色域范围也最小，**FWGP&EWG** 在表8 中明显下降。使用稍逊的 HDRTV1K 作为标注时，差异相对较小
不太显著。不过，在这两种情况下，**ASL&ALL**都相似，因为显示映射即网络的风格 倾向没有改变。

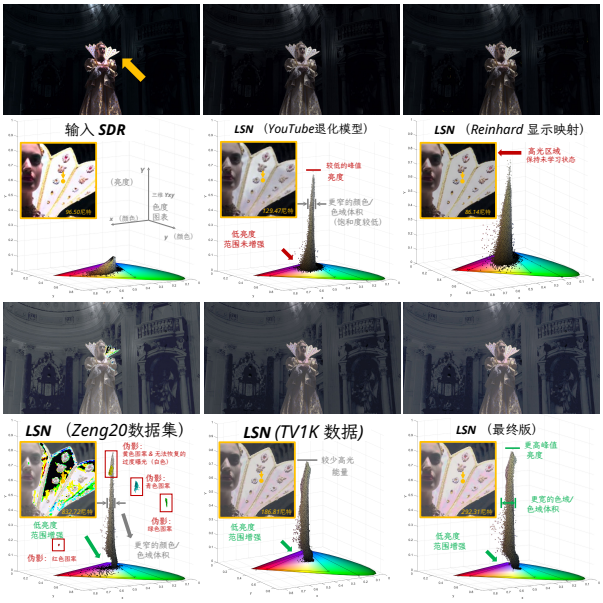



图8. 消融研究的结果证明了高质量标注HDR和合理退化模型的重要性。具体而言，缺少标注HDR的HDR/WCG体积（Zeng20）以及显示映射的退化（***Reinhard***）都会削弱LSN的恢复能力；与此同时，***YouTube***显示映射的风格 会使我们的LSN变得和其他方法一样平庸。

## 5. 结论

低层视觉有两种类型：“单纯复原”，其目标仅仅是获得干净图像或真实值例如去噪；以及“感知驱动”，旨在带来更好的观看体验例如图像增强/图像修饰。SDR 到 HDRTV 的升转换同时属于这两种类型。然而，当前方法只实现了前者，而忽略了后者，因此只关注网络机制。

为此，我们从两个方面作出回应：**(1)** 关注训练集的影响，并通过提出新数据集和退化模型来提升其质量；**(2)** 引入基于“感知”原则的新评估标准。

结果证明了我们观点的有效性，并表明我们的方法已可用于实际应用。

## 参考文献


[1] 国际电信联盟，瑞士日内瓦，报告 ITU-R BT.2381-0：高动态范围电视（HDR-TV）系统的要求，第 0 版，2015 年 7 月。1, 6[2]国际电信联盟，瑞士日内瓦，ITU-R BT.2100-2建议书：用于制作和国际节目交换的高动态范围电视图像参数值，第 2 版，2018 年 7 月。1, 2[3]国际电信联盟，瑞士日内瓦，ITU-R BT.2020-2建议书：用于制作和国际节目交换的超高清电视系统参数值，第 2 版，2015 年 10 月。1, 2[4] S. Y. Kim、J. Oh 和 M. Kim，“Deep SR-ITM：面向 4K超高清HDR应用的超分辨率与逆色调映射联合学习”，载于 ICCV 论文集，第 3116–3125 页，2019年。2, 3, 5, 6, 7, 8[5] S. Y. Kim、J. Oh 和 M. Kim，“JSI-GAN：基于生成对抗网络的联合超分辨率与逆色调映射，采用逐像素任务专用滤波器处理超高清HDR视频”，载于 AAAI 论文集，第 34 卷，第 11287–11295 页，2020年。2, 3, 6, 7, 8[6] H. Zeng、X. Zhang、Z. Yu 和 Y. Wang，“SR-ITM-GAN：利用生成对抗网络学习 4K超高清HDR”，IEEE Access，第 8 卷，第 182815–182827 页，2020年。2, 3, 5, 6, 7, 8[7] X. Chen、Z. Zhang、J. S. Ren、L. Tian、Y. Qiao 和 C. Dong，“从标准动态范围电视迈向高动态范围电视的新旅程”，载于 ICCV 论文集，第 4500–4509 页，2021年。2, 3, 5, 6, 7, 8[8] G. Cao、F. Zhou、H. Yan、A. Wang 和 L. Fan，“KPN-MFI：一种用于视频逆色调映射、具有多帧交互的核预测网络”，载于 IJCAI 论文集，第 806–812 页，2022年。2, 3, 6, 7, 8 – .[9] G. Xu、Q. Hou、L. Zhang 和 M.-M. Cheng，“FMNet：用于标准动态范围到高动态范围转换的频率感知调制网络”，载于 ACMMM 论文集，第 6425–6435 页，2022年。2, 3, 6, 7, 8[10] Z. Liang，“CVPR 2022 第一届视觉数据集理解研讨会。” https://sites.google.com/view/vdu-cvpr22，2022年。2[11] K. Zhang、J. Liang、L. Van Gool 和 R. Timofte，“为深度盲图像超分辨率设计实用的退化模型”，载于 ICCV 论文集，第 4791–4800 页，2021年。2, 3[12] V. Dewil、A. Barral、G. Facciolo 和 P. Arias，“自监督与合成数据集：在视频去噪领域，哪种弊端较小？”，载于 CVPR 论文集，第 4900–4910 页，2022年。2, 3[13] J. Jiang、K. Zhang 和 R. Timofte，“迈向灵活的盲JPEG伪影去除”，载于 ICCV 论文集，第 4997–5006 页，2021年。2, 3[14] K. Jiang、Z. Wang、Z. Wang、C. Chen、P. Yi、T. Lu 和 C.-W. Lin，“退化即升级：学习低光照图像增强的退化”，载于 AAAI 论文集，第 36 卷，第 1078–1086 页，2022年。2, 3

[15] Y. Zhou、Y. Song 和 X. Du，《面向屏下摄像头的模块化退化模拟与复原》，载于 ACCV 论文集，第 265–282 页，2022年。2, 3[16] C. Guo 和 X. Jiang，《Lhdr：利用轻量级深度神经网络为旧内容重建高动态范围图像》，载于 ACCV 论文集，第 3155–3171 页，2022年。2, 3[17] S. W. Zamir、A. Arora、S. Khan、M. Hayat、F. S. Khan 和 M.-H. Yang，《Restormer：用于高分辨率图像复原的高效 Transformer》，载于 CVPR 论文集，第 5728–5739 页，2022年。2, 3, 4[18] G. Eilertsen、S. Hajisharif、P. Hanji、A. Tsirikoglou、R. K. Mantiuk 和 J. Unger，《如何在单幅图像 HDR 重建中利用指标“作弊”》，载于 ICCV 论文集，第 3998–4007 页，2021 年。2, 3, 6, 7[19] P. Hanji 、 R. Mantiuk、G. Eilertsen、S. Hajisharif 和 J. Unger，《单幅图像 HDR 重建方法比较——质量评估中的注意事项》，载于 SIGGRAPH 论文集，第 1–8 页，2022年。2, 6, 7, 8[20] L. Wang 和 K.-J. Yoon，《HDR 成像的深度学习：最新进展与未来趋势》，IEEE PAMI 汇刊，2021年。2[21] P. Debevec，《基于图像的照明》，载于 ACM SIGGRAPH 2006 课程，第 4–es 页，2006年。2[22] E. Reinhard、W. Heidrich、P. Debevec 等，《高动态范围成像：采集、显示与基于图像的照明》。Morgan Kaufmann，2010年。2[23] M. Gharbi、J. Chen、J. T. Barron、S. W. Hasinoff 和 F. Durand，《用于实时图像增强的深度双边学习》，《ACM图形学汇刊》，第 36 卷，第 4 期，第 1–12 页，2017年。2[24] X. Yang、K. Xu、Y. Song、Q. Zhang、X. Wei 和 R. W. Lau，《通过深度往复式 HDR 变换进行图像校正》，载于 CVPR 论文集，第 1798–1807 页，2018年。2[25] Z. Zheng、W. Ren、X. Cao、T. Wang 和 X. Jia，《通过协同双边学习进行超高清图像 HDR 重建》，载于 ICCV 论文集，第 4449–4458 页，2021年。2[26] B. Mildenhall、P. Hedman、R. Martin-Brualla、P. P. Srinivasan 和 J. T. Barron，《暗光中的 NeRF：从含噪 RAW 图像进行高动态范围视图合成》，载于 CVPR 论文集，第 16190–16199 页，2022年。2[27] N. K. Kalantari、R. Ramamoorthi 等，《动态场景的深度高动态范围成像》，《ACM图形学汇刊》，第 36 卷，第 4 期，第 144–1 页，2017年。2[28] S. Wu、J. Xu、Y.-W. Tai 和 C.-K. Tang，《具有大幅前景运动的深度高动态范围成像》，载于 ECCV 论文集，第 117–132 页，2018年。2[29] Q. Yan、D. Gong、Q. Shi、A. v. d. Hengel、C. Shen、I. Reid 和 Y. Zhang，《用于无重影高动态范围成像的注意力引导网络》，载于 CVPR 论文集，第 1751–1760 页，2019年。2[30] G. Chen、C. Chen、S. Guo、Z. Liang、K.-Y. K. Wong 和 L. Zhang，《HDR 视频重建：由粗到细的网络与真实世界基准数据集》，载于 ICCV 论文集，第 2502–2511 页，2021年。2

[31] Y. Niu、J. Wu、W. Liu、W. Guo 和 R. W. Lau，《Hdr-gan：利用多曝光 LDR 图像和大幅运动进行 HDR 图像重建》，《IEEE图像处理汇刊》，第 30 卷，第 3885–3896 页，2021年。2[32] E. Pérez-Pellitero 等，《NTIRE 2022高动态范围成像挑战赛：方法与结果》，载于 CVPR 论文集，第 1009–1023 页，2022年。2[33] G. Eilertsen、J. Kronander 等，《使用深度卷积神经网络从单次曝光图像重建 HDR 图像》，《ACM图形学汇刊》，第 36 卷，第 6 期，第 1–15 页，2017年。2, 3[34] D. Marnerides、T. Bashford-Rogers 等，《ExpandNet：用于从低动态范围内容扩展高动态范围的深度卷积神经网络》，《计算机图形论坛》，第 37 卷，第 2 期，第 37–49 页，2018年。2[35] Y.-L. Liu、W.-S. Lai、Y.-S. Chen、Y.-L. Kao、M.-H. Yang、Y.-Y. Chuang 和 J.-B. Huang，《通过学习逆转相机处理流程进行单幅图像 HDR 重建》，载于 CVPR 论文集，第 1651–1660 页，2020年。2, 3[36] M. S. Santos、T. I. Ren 和 N. K. Kalantari，《使用带掩码特征和感知损失的卷积神经网络进行单幅图像 HDR 重建》，《ACM图形学汇刊》，第 39 卷，第 1 期，第 1–15 页，2020年。2, 4[37] X. Chen、Y. Liu 等，《HDRUNet：通过去噪和反量化进行单幅图像 HDR 重建》，载于 CVPR 论文集，第 354–363 页，2021年。2[38] S. Ning、H. Xu、L. Song、R. Xie 和 W. Zhang，《利用生成对抗正则项学习逆色调映射网络》，载于 ICASSP 论文集，第 1383–1387 页，2018年。2, 3[39] K. Hirao、Z. Cheng、M. Takeuchi 和 J. Katto，《使用 LuCore 的卷积神经网络高动态范围显示逆色调映射》，载于 2019 IEEE 国际消费电子会议（ICCE），第 1–2 页，2019年。2[40] S. Y. Kim、D.-E. Kim 和 M. Kim，《ITM-CNN：使用卷积神经网络学习将低动态范围视频逆色调映射至高动态范围显示器》，载于 ACCV 论文集，第 395–409 页，2018年。2, 3[41] Y. Xu、S. Ning、R. Xie 和 L. Song，《基于生成对抗网络的多曝光逆色调映射》，载于 ICIP 论文集，第 4365–4369 页，2019年。2, 3[42] Y. Xu、L. Song、R. Xie 和 W. Zhang，《深度视频逆色调映射》，载于 2019 IEEE第五届多媒体大数据国际会议（BigMM），第 142–147 页，2019年。2, 3[43] D.-E. Kim 和 M. Kim，《基于学习的低复杂度逆色调映射与线性映射》，《IEEE视频技术电路与系统汇刊》，第 30 卷，第 2 期，第 400–414 页，2019年。2, 3, 8[44] J. Zou、K. Mei 和 S. Sun，《利用可变形对齐进行多尺度视频逆色调映射》，载于 2020 IEEE 国际视觉通信与图像处理会议（VCIP），第 9–12 页，2020年。2, 3[45] T. Chen 和 P. Shi，《基于多尺度双分支网络的逆色调映射算法》，载于 2021 国际文化科技会议（ICCST），第 187–191 页，2021年。2, 3

[46] G. Xu、Y. Yang、J. Xu、L. Wang、X.-T. Zhen 和 M.-M. Cheng，《联合超分辨率与逆色调映射：一种特征分解聚合网络与新基准》，2022年。2，3

[47] G. He, K. Xu, L. Xu, C. Wu, M. Sun, X. Wen 和 Y.-W. Tai，“通过分层动态上下文特征映射进行 SDRTV 到 HDRTV 转换”，载于 *ACMMM* 论文集，第 2890–2898 页，2022年。2，3

[48] G. He, S. Long, L. Xu, C. Wu, J. Zhou, M. Sun, X. Wen, 和 Y. Dai，“用于联合超分辨率与逆色调映射的全局先验引导调制网络”，2022年。2, 3[49] K. Xu, L. Xu, G. He, C. Wu, Z. Ma, M. Sun, 和 Y.-W. Tai，“通过时空特征融合实现 SDRTV 到 HDRTV 的转换”，2022年。2, 3[50] T. Shao, D. Zhai, J. Jiang, 和 X. Liu，“混合条件深度逆色调映射”，载于 ACMMM 论文集，第 1016–1024 页，2022年。2, 3[51] A. Mustafa, P. Hanji, 和 R. K. Mantiuk，“从图像对中提取风格，用于全局正向与逆色调映射”，载于第十九届 ACM SIGGRAPH 欧洲视觉媒体制作会议（CVMP），第 1–10 页，2022年。2, 3, 4[52] M. Yao, D. He, X. Li, Z. Pan, 和 Z. Xiong，“UHD-HDR 与 HD-SDR 视频之间的双向转换”，IEEE 多媒体汇刊，2023年。2, 3[53] R. Tang, F. Meng, 和 L. Bai，“从 SDR 视频到 HDR 视频的分区映射网络”，载于 2022 年 IEEE 第五届先进信息管理、通信、电子与自动化控制会议（IMCEC），第 5 卷，1472，2022 1466 2, 3，第 – 页。[54] SMPTE，美国纽约，ST 2084:2014 - SMPTE 标准 - 母版参考显示器的高动态范围电光传递函数，2014年。2[55] 国际电信联盟，瑞士日内瓦，ITU-R BT.709-6建议书：高清电视标准制作与国际节目交换的参数值，第 6 版，2015年 6 月。2, 5[56] Z. Cheng, T. Wang, Y. Li, F. Song, C. Chen, 和 Z. Xiong，“迈向真实场景 HDRTV 重建：一种基于数据合成的方法”，载于 ECCV 论文集，第 199–216 页，2022年。3[57] E. Reinhard, M. Stark, P. Shirley, 和 J. Ferwerda，“数字图像的摄影色调再现”，载于第 29 届 SIGGRAPH 论文集，第 267–276 页，2002年。3[58] 国际电信联盟，瑞士日内瓦，ITU-R BT.2446-1 号报告：高动态范围内容与标准动态范围内容之间的转换方法，第 1 版，2021年 3 月。3, 5[59] P. Mohammadi, M. T. Pourazad, 和 P. Nasiopoulos，“一种面向高动态范围视频应用的基于感知的逆色调映射算子”，IEEE 电路与系统视频技术汇刊，第 31 卷，第 5 期，第 1711–1723 页，2020年。3, 8[60]D. Marnerides, T. Bashford-Rogers, 和 K. Debattista，“用于逆色调映射的深度 HDR 幻觉”，《传感器》，第 21 卷，第 12 期，第 4032 页，2021年。3

[61] Y. Zhang 和 T. Aydın，“利用生成式细节重建进行深度 HDR 估计”，载于 Computer Graphics Forum，第 40 卷，第 179–190 页，2021年。3[62] L. Bai, Y. Yang, 和 G. Fu，“超高清电视高动态范围与宽色域分析”，载于 2021 年 I EEE 第五届先进信息技术、电子与自动化控制会议（I AEAC），第 5 卷，第 1750–1755 页，2021年。4[63] 国际电信联盟，瑞士日内瓦，ITU-R BT.2407-0号报告：从 ITU-R BT.2020建议书到 ITU-R BT.709建议书的色域转换，第 0 版，2017年 10 月。4, 6[64] 国际电信联盟，瑞士日内瓦，ITU-R BT.500-14建议书：电视图像质量主观评估方法，第 14 版，2019年 10 月。4[65] D. Hasler 和 S. E. S uesstrunk，“自然图像中色彩丰富度的测量”，载于《 Human Vision and Electronic Imaging VIII》，第 5007 卷，第 87–95 页，2003年。4[66] 杜比，“什么是 ICtCp？——简介。” https : / / professional.dolby.com/siteassets/pdfs/ ictcp_ dolbywhitepaper_v071.pdf。4[67] S. Ploumis, R. Boitard, 和 P. Nasiopoulos，“HDR 亮度量化”，载于 2020 年第 28 届欧洲信号处理会议（EUSIPCO），第 640–644 页，2021年。4[68]Arri，“摄像机样片与参考图像。” https : / / www . arri . com / en / learn - help / learn-help-camera-system/camera-sample-footage-reference-image，2022年。4[69] 奈飞，“奈飞开放内容。” https : / / opencontent.netflix.com/，2020年。4[70]J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, 和 H. Brendel，“制作用于评估色调映射算子和 HDR 显示器的电影级宽色域 HDR 视频”，载于《Digital Photography X》，第 9023 卷，第 279–288 页，2014年。4[71] Y. Wang, S. Inguva, 和 B. Adsumilli，“用于视频压缩研究的 YouTube 用户生成内容数据集”，载于 2019 年 IEEE 第 21 届多媒体信号处理国际研讨会（ MMSP），第 1–5 页，2019年。4[72] Z. Shang, J. P. Ebenezer, A. C. Bovik, Y. Wu, H. Wei, 和 S. Sethura man，“不同环境条件下高动态范围视频的主观评估”，载于 ICIP 论文集，第 786–790 页，2022年。4[73] L. Van der Maaten 和 G. Hinton，“使用 t-SNE 可视化数据”，J ournal of Machine Learning Research，第 9 卷，2008 5，第 11 期。[74] D. Walker, C. Payne, P. Hodoul, 和 M. Dolan，“使用 OpenColorIO v2 进行色彩管理”，载于 ACM SIGGRAPH 2021 Courses，第 1–226 页，2021年。5[75] J. Morovič，《色域映射》。John Wiley & Sons，2008年。6[76] H. Le, T. Jeong, A. Abdelhamed, H. J. Shin, 和 M. S. Brown，“GamutNet：为相机拍摄图像还原宽色域色彩”，载于《色彩与成像会议》，第 2021 卷，第 7–12 页，2021年。6

[77]国际电信联盟，日内瓦，瑞士，*ITU-R BT.1886-0*号建议书：用于高清电视演播室制作的平板显示器参考电光传递函数，0 3 2011年6版，。[78]国际电信联盟，日内瓦，瑞士，*ITU-R BT.2245-10*号报告：用于评估图像质量的高清电视和超高清电视（包括*HDR-TV*）测试材料，第10版，9 2022年。6[79]国际电信联盟，日内瓦，瑞士，*ITU-R BT.2124-0*号建议书：用于评估电视中色彩差异潜在可见性的客观度量指标，第0版，1 2019年。7[80] K. Wolski、D. Giunchi，等，“用于预测局部可见差异的数据集和度量指标”，《*ACM*图形学汇刊》，第37卷，第5期，第1–14页，2018年。7[81] C. Bist、R. Cozot、G. Madec 和 X. Ducloux，“运用照明风格美学进行色调扩展”，《计算机与图形学》，第62卷，第77–86页，2017年。8[82] G. Luzardo、J. Aelterman、H. Luong、S. Rousseaux、D. Ochoa 和 W. Philips，“基于动态中阶色调映射的全自动逆色调映射算法”，《*APSIPA*信号与信息处理汇刊》，第9卷，文章编号e7，2020年。8[83] A. Stojkovic、J. Aelterman、H. Luong、H. Van Parys 和 W. Philips，“高光分析系统（hans）：将电影低动态范围内容从低动态范围转换为高动态范围”，*IEEE Access*，第9卷，第43938–43969页，2021年。8