

用于图像照明控制的本征图像传输算法

Zebediah

2024 年 5 月 5 日

- Matlab code: https://github.com/QingXin96/Intrinsic_image_transfer
- Article link: [Intrinsic Image Transfer for Illumination Manipulation](#)

1 引言

首先, 我们需要知道一个著名的本征图像模型

$$\mathcal{I} = \mathcal{L} \odot \mathcal{R}, \quad (1.1)$$

其中 \mathcal{L} 和 \mathcal{R} 分别表示光照和反射率, \odot 表示 Hadamard 积. 如果没有对光照 \mathcal{L} 或反射率 \mathcal{R} 施加进一步的假设, (1.1) 是一个约束条件很少的问题. 有时我们可以考虑这样一个线性模型:

$$\hat{\mathcal{I}} = a + b\mathcal{L} + c\mathcal{R}. \quad (1.2)$$

然而这个模型仍然具有局限性.

将内在图像分解、层次重映射和图像重构集成到一个广义的基于优化的框架中是一种很好的方法, 它使我们能够以隐式的方式控制图像光照, 而无需进行明确的本征图像分解.

在此, 我们给出以下假设, 这些假设可以减少未知数的数量, 使得问题变得适定:

- 光照 L 的空间平滑性;
- 反射率 R 是光照不变的;
- 图像可见度主要取决于光照 L .

2 主要目标

令 $s = \{s_i\}_{i=1}^N$ 和 $o = \{o_i\}_{i=1}^N$ 分别表示输入图像和输出图像的像素强度, 我们的目标是最小化损失函数

$$\min_{\mathbf{o}} E(\mathbf{o}) = \alpha E^l(\mathbf{o}) + \beta E^r(\mathbf{o}) + \gamma E^c(\mathbf{o}), \quad (2.1)$$

其中 E^l, E^r 和 E^c 分别定义在光照、反射率和内容上, α, β, γ 控制这三个项的平衡.

3 模型构建

3.1 照明损失

由于不断变化的照明条件, 图像可能遭受严重的光照退化. 在这种情况下, 我们希望将退化的光照恢复到一个良好的平衡状态. 假设有一个额外的图像 $\mathbf{c} = \{c_i\}_{i=1}^N$ 具有理想的潜在光照 \mathbf{c}^l , 那么使输出光照 \mathbf{o}^l 接近潜在光照 \mathbf{c}^l 是合适的, 同时导致光照损失 $E^l(\mathbf{o})$, 从而

$$E^l(\mathbf{o}) = \sum_i (o_i^l - c_i^l)^2. \quad (3.1)$$

基于假设 (a), 我们考虑一个简单的类高斯平滑核 \mathcal{K} , 从而有

$$o_i^l = \sum_{j \in \mathcal{N}_i} \mathcal{K}_{i,j} o_j, \quad \sum_{j \in \mathcal{N}_i} \mathcal{K}_{i,j} = 1, \quad \mathcal{K}_{i,j} \propto \exp\left(-\frac{1}{\delta_f^2} \|\mathbf{f}_i - \mathbf{f}_j\|_2^2\right), \quad (3.2)$$

其中 \mathcal{N}_i 表示像素 i 的邻域集合, $\mathcal{K}_{i,j}$ 是像素 i 和 j 之间的高斯核权重, \mathbf{f} 是一个具有标准差 δ_f 的特征向量. 由 (3.1) 和 (3.2) 可得

$$E^l(\mathbf{o}) = \sum_i \sum_{j \in \mathcal{N}_i} \left(\mathcal{K}_{i,j}^o o_j - \mathcal{K}_{i,j}^c c_j \right)^2, \quad (3.3)$$

其中 $\mathcal{K}_{i,j}^o$ 和 $\mathcal{K}_{i,j}^c$ 是相应的滤波核.

3.2 反射率损失

基于假设 (b), 我们期望输出的反射率 \mathbf{o}^r 与原始图像的反射率 \mathbf{s}^r 相同, 于是定义反射率损失为

$$E^r(\mathbf{o}) = \sum_i (o_i^r - s_i^r)^2, \quad (3.4)$$

其中 o_i^r 和 s_i^r 分别是相应反射层的第 i 个像素值. 然而 (3.4) 需要一个难以处理的本征图像分解, 又注意到反射层 R 被假设为具有丰富的特征, 并具有相同的光照不变性质, 这启发我们通过正则化这些特征来控制 R . 因此我们使用局部线性模型来编码这些特征:

$$s_i^r = \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} s_j^r, \quad (3.5)$$

其中 Ω_i 是像素 i 的邻域, 权重 $\omega_{i,j}^{s^r}$ 满足 $\sum_{j \in \Omega_i} \omega_{i,j}^{s^r} = 1$. 如果反射率 s^r 保持不变, 那么权重 $\omega_{i,j}^{s^r}$ 也是不变的. 对于 \mathbf{o}^r 也有类似的结果, 于是我们可以得到

$$\begin{cases} o_i^r = \sum_{j \in \Omega_i} \omega_{i,j}^{o^r} o_j^r, & \sum_{j \in \Omega_i} \omega_{i,j}^{o^r} = 1, \\ s_i^r = \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} s_j^r, & \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} = 1, \\ \omega_{i,j}^{o^r} = \omega_{i,j}^{s^r}. \end{cases} \quad (3.6)$$

在 (3.6) 中, 我们强制令 $\omega_{i,j}^{o^r} = \omega_{i,j}^{s^r}$, 以给出反射层的结构一致性约束. 注意到空间中光照的变化相对较小, 于是有 $s_k^l \approx \bar{s}_k^l$, 其中 \bar{s}_k^l 是局部块 k 中的均值. 此外, 如果两个块 i 和 j 彼此靠近, 还成立 $\bar{s}_i^l \approx \bar{s}_j^l$. 因此

$$\begin{aligned} s_i &= (s_i^r + s_i^l) \approx (s_i^r + \bar{s}_i^l) = \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} s_j^r + \bar{s}_i^l = \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} (s_j^r + \bar{s}_i^l) \\ &\approx \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} (s_j^r + \bar{s}_j^l) \approx \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} (s_j^r + s_j^l) = \sum_{j \in \Omega_i} \omega_{i,j}^{s^r} s_j. \end{aligned} \quad (3.7)$$

对比 (3.5) 和 (3.7), 当添加一个恒定光照 \bar{s}_i^l 时, $\omega_{i,j}^{s^r}$ 仍保持不变. 由于局部线性模型的平移不变性, 反射层可以直接进行正则化, 而无需将反射从图像中解耦, 因此综合 (3.4)-(3.7) 得

$$\begin{aligned} E^r(\mathbf{o}) &= \sum_i \left(o_i - \sum_{j \in \Omega_i} \omega_{i,j}^o o_j \right)^2, \\ \text{s.t. } \quad \omega_{i,j}^o &= \omega_{i,j}^s, \quad s_i = \sum_{j \in \Omega_i} \omega_{i,j}^s s_j, \end{aligned} \quad (3.8)$$

其中 $\omega_{i,j}^s = \omega_{i,j}^{s^r}$ 和 $\omega_{i,j}^o = \omega_{i,j}^{o^r}$ 表示局部“编码”权重.

3.3 内容损失

为了避免过拟合, 在此引入内容损失, 定义如下:

$$E^c(\mathbf{o}) = \sum_i (o_i - s_i)^2. \quad (3.9)$$

4 IIT 算法

将三个损失函数组合, 并以矩阵形式重写 (2.1) 为

$$\begin{aligned} E(\mathbf{o}) &= \alpha \| \mathbf{K}^o \mathbf{o} - \mathbf{K}^c \mathbf{c} \|_2^2 + \beta \| \mathbf{M} \mathbf{o} \|_2^2 + \gamma \| \mathbf{o} - \mathbf{s} \|_2^2, \\ \text{s.t. } \quad \omega_{i,j}^o &= \omega_{i,j}^s, \quad s_i = \sum_{j \in \Omega_i} \omega_{i,j}^s s_j, \end{aligned} \quad (4.1)$$

其中 $\mathbf{K}^o(\mathbf{K}^c)$ 是一个核矩阵, 其 i, j 项为 $\mathcal{K}_{i,j}^{o(c)}$; $\mathbf{M} = [\mathbf{I} - \mathbf{W}]$ 是一个稀疏矩阵, 其中 \mathbf{I} 是单位矩阵, \mathbf{W} 的 i, j 项为 $\omega_{i,j}^o$.

首先计算滤波核 \mathcal{K} . 最简单的情形是二维高斯滤波器, 权重 \mathbf{f} 依赖于像素位置 $\mathbf{p} = [p_x, p_y]$. 为了获得更好的结果, 可以使用双边滤波器, 它考虑了像素的位置 $\mathbf{p} = [p_x, p_y]$ 和颜色强度 (R_i, G_i, B_i) . 这里设高斯滤波器的标准差为 $\delta_{f_i} = \delta_s$, 双边滤波器的标准差为 $\delta_{f_i} = (\delta_s, \delta_r)$. 为简单起见, 由于预先不可用图像 \mathbf{o} , 我们在实验中将直接令 $\mathbf{K} = \mathbf{K}^o = \mathbf{K}^c$.

对于 LLE 权重, 当每个点的近邻点数量大于空间维度时, 采用直接求解可能不稳定, 可以通过解决逐点正则化问题来计算 \mathbf{W} :

$$\min \sum_i \left(s_i - \sum_{j \in \Omega_i} \omega_{i,j}^s s_j \right)^2 + \epsilon \|\boldsymbol{\omega}^s\|_2^2, \quad \text{s.t. } \sum_{j \in \Omega_i} \omega_{i,j}^s = 1. \quad (4.2)$$

在得到 \mathbf{K} 和 \mathbf{W} 后, 可以通过 $dE/d\mathbf{o} = 0$ 来求解 (4.1), 因此

$$(\alpha \mathbf{K}^T \mathbf{K} + \beta \mathbf{M}^T \mathbf{M} + \gamma \mathbf{I}) \mathbf{o} = \alpha \mathbf{K}^T \mathbf{K} \mathbf{c} + \gamma \mathbf{s}, \quad (4.3)$$

其中 $\mathbf{L} = \alpha \mathbf{K}^T \mathbf{K} + \beta \mathbf{M}^T \mathbf{M} + \gamma \mathbf{I}$ 是一个大规模且稀疏的拉普拉斯矩阵. 由于 \mathbf{L} 对称且半正定, (4.3) 可以通过 Gauss-Seidel 算法和预条件共轭梯度法 (PCG) 等方法来求解.

总之, 本文 IIT 算法包括: 确定用于光照拟合的滤波核, 计算用于编码图像反射率的 LLE 权重, 并嵌入反射率层以进行图像重建. IIT 算法过程如下所示:

算法 1 IIT

输入: $\{s_i\}_{i=1,\dots,N}, \{c_i\}_{i=1,\dots,N}$ 和 α, β, γ ;

输出: $\{o_i\}_{i=1,\dots,N}$;

过程:

1: 识别滤波器

- 1) 设置参数: $\mathcal{N}_i, \delta_s, \delta_r$ 或者 $(\mathcal{N}_i, \delta_s)$;
- 2) 从 (3.2) 得到 $\mathcal{K}_{i,j}$, 以及从 (4.1) 得到 $\mathbf{K}^o, \mathbf{K}^c$;

2: 计算 LLE 权重: \mathbf{W}

- 1) 设置参数: Ω_i, ϵ ;
- 2) 为每个像素 i 寻找近邻 Ω_i ;
- 3) 从 (4.2) 得到 $\omega_{i,j}^s$;
- 4) 令 $\omega_{i,j} = \omega_{i,j}^s$, 计算 $\mathbf{M} = \mathbf{I} - \mathbf{W}$;

3: 图像重建

- 1) 从 (4.3) 得到拉普拉斯矩阵 \mathbf{L} ;
 - 2) 使用 PCG 算法求解 (4.3).
-

5 实验结果

如图 1 所示, 图中的示例是由 CLAHE 算法生成的, 具有精细平衡的照明, 但局部细节略有扭曲, 而 IIT 算法在抑制局部噪声和失真方面取得了显著改进, 特别是天鹅“颈部”和“翅膀”区域周围. 此外, 使用高斯和双边滤波器之间存在微小的视觉差异.

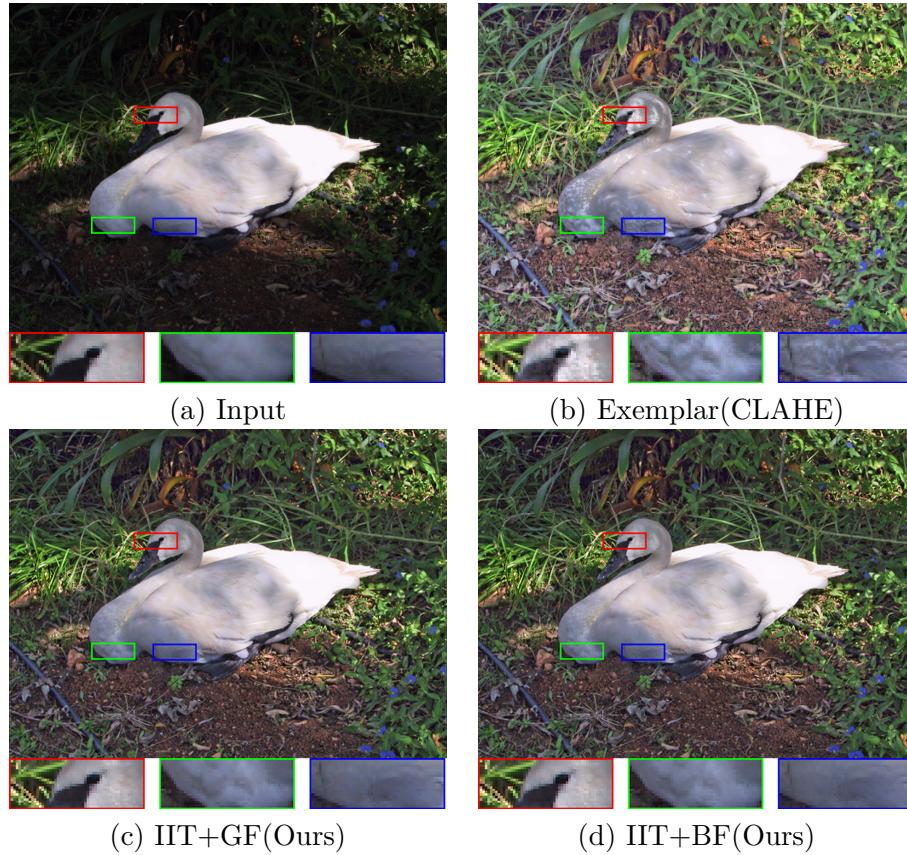


图 1: 示例 1

如图 2 所示, (b) 和 (d) 中的示例具有不同的基础照明, 而在 (c) 和 (e) 中, 示例中不一致的细节在两种情况下都得到了显著抑制. 因此平衡的图像照明是 IIT 算法的关键因素, 而不是高质量的局部细节.

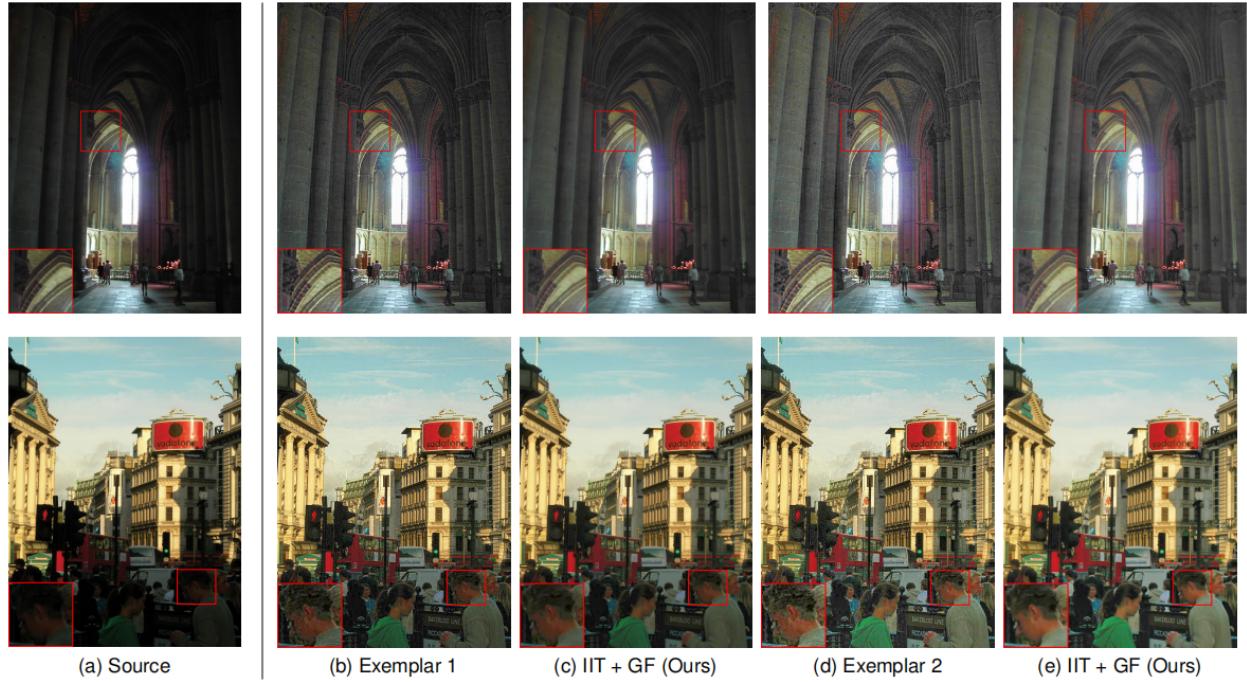


图 2: 示例 2

此外, IIT 算法对具有严重局部退化细节的示例的有很强的鲁棒性, 还有其他许多优点, 缺点之一是速度一般, 在此不一一例举.