

指导教师： 杨涛

提交时间： 2016/03/16

# CVPR2015 Paper Translation

No: 01

姓名： 李海军

学号： 2013302599

班号： 10011306



## Learning lightness from human judgement on relative reflectance

## 从人对相对反射率的判断中学习亮度

## 概要

我们开发了一个新的方法，可以仅从一个单一的形象来推断亮度和表面的反射率。以往的方法常从固有图像分解的角度来对待此问题，他们将图像被分离成反射和阴影两部分。但我们发现直接预测像素之间的亮度存在差异的原因却不是因为反射和阴影。大规模的培训，对人类相对反射率的数据的判断，以及深度网络的补丁表示的使用为我们的模型提供了基础。基于内在图像在野外数据分析[4]，我们的局部光照模型达到与国家最先进的全球亮度模型的匹配程度。其中包括在多个光照/反射前科条件下，像素点在一个密集的环境中随机分配的情况。

## 第一部分

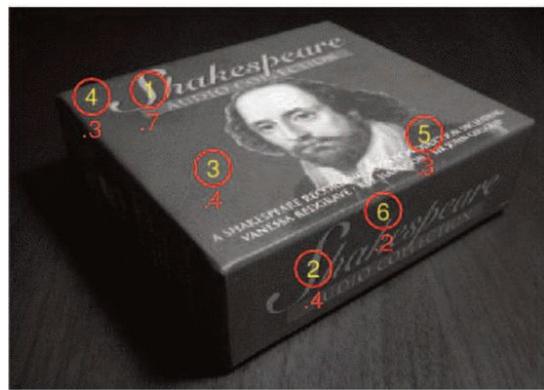
## 介绍

人类的视觉在解码物理反射率反面表现非常突出，哪怕存在着光照的变化。此主观亮度的持久性，这是指所感知的反射率，使得我们计算图像中记录的基于客观发光体的光强变得很困难。

图1示出对像素对强度和亮度之间的复杂的排名的关系[25]:

- 不同强度相同明度: 字母 S 在位置 1 和 2 圈是一样的白漆无论是在盒子的阴影面。同样的, 在亮度不均匀的位置 3 和 4 的油漆颜色是不会判断出错的。

- 强度不同的亮度差异: 黑色的衣服在位置 5 总是正确地看到, 尽管实际上比在位置 6 背景灰色更多, 光线更暗。
- 相同强度的不同明度: 位置 2 和 3, 4 和 5 亮度是相同的水平但灰度不同, 易被分为不同颜色。



intensity:  $I_1 > I_2 = I_3 > I_4 = I_5 > I_6$   
lightness:  $L_1 = L_2 > L_3 = L_4 = I_6 > L_5$

图1.强度与亮度。亮度的主观经验与光在目标图像中的强度有着不平凡的关系。此图像具有3D的盒子, 它的顶部和前表面具有相同的文字和背景颜色的油漆。六个像素被标记, 并标有0和1之间及其亮度即感知的反射率的强度值, 从他们的测量强度排序非常不同。源: [25]。

常规亮度的建模的目的是通过从所述图像强度分开强度和反射率这两个混淆因素, 在所谓的固有图像分解[3]框架以回收物理反射率。

基本思路是, 虽然在该像素本身的强度不足以作为亮度指标, 但其相邻像素通常具有足够的变化来揭示底层照明和反射率, 其中每一个具有在该组的自然图像的鲜明特征[1]。亮度

从识别强度和扣除照明的两方面组成。

左右的反射率和阴影的规律性知识来解决：

具体而言，本征图像模型假定图像强度是反射率图像和阴影图像的产物，而且亮度是简单地通过给出其

$$I = R \cdot S, \tag{1}$$

$$L = R^*, \text{ s.t. } \text{priors}(R, S) \tag{2}$$

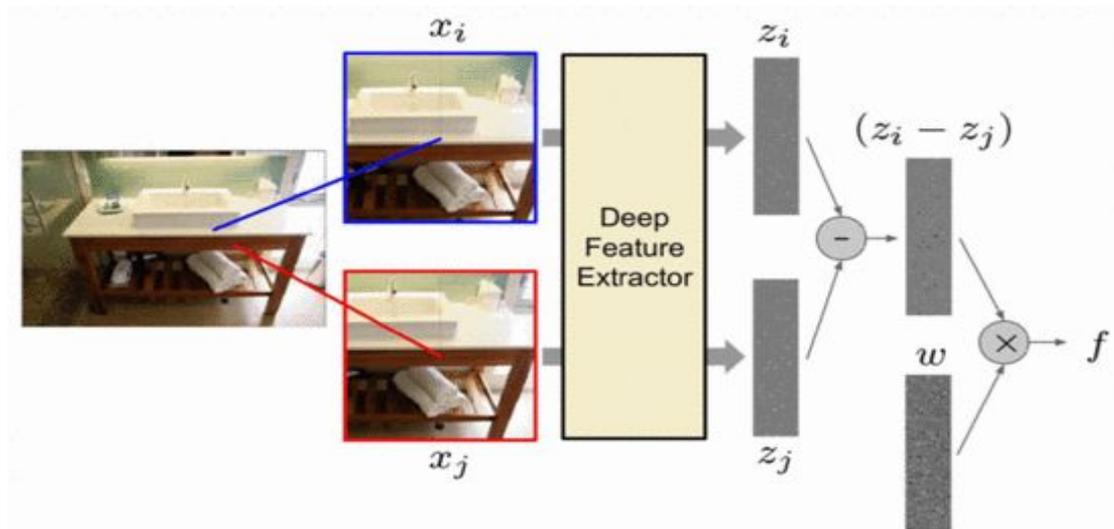


图 2.成对亮度的关系直接学习。给定两个图像补丁，我们提取其深厚的特征并进行训练分类，把它们的亮度值成对按亮度进行排序。

这样的分解是不明确而且没有先例，而一般的策略是利用强大的先验来满足每像素分解所需的搜索空间约束。

我们的方法包括一个完整的变化，涉及直觉和战略。我们学习亮度模型，直接利用数据，通过人类受试者进行一组训练集中许多相对反射率的比较。这就是我们所做的。我们专注于学习的相对顺序，即直接从存在于两个局部图像中找出  $x_i$  和  $x_j$ ，并不需要一个绝对的像素值被分解成可能的和。

许多形式的先验知识，对在更广泛的范围内的空间和统计规律进行了探讨。例如，反射率是假定分段常数 [ 15 ]，[ 16 ]，[ 2 ]，或从稀疏集 [ 18 ]，[ 9 ]，[ 22 ]，而阴影是是临近像素 [ 8 ]，或更可能采取某些值比如 [ 2 ]，[ 16 ]。好的先验知识，也可以了解到生产的使用深层信念网 [ 23 ]。该解决方案已经应用在全球寻求通过局部纹理约束 [ 26 ]或密集的连接条件随机域的所有像素对之间 (CRF) [ 4 ]。

一般的，我们会找出相对位置  $i$  和  $j$  的亮度级别差作为特征值，从本地图像块中提取  $x_i, x_j$ ，然后计算：

$$L_i - L_j = f(z_i, z_j) \tag{3}$$

我们的直觉是，在越来越大的两个像素的邻域提取可捕获的照明和反射背景等的特征。我们需要知道的不是他们究竟是什么，而是他们是如何不同的，以便对这 2 个像素的亮度排序作出判断。

我们的模型是建立在最近的工作的 2 个关键单元：

- 在野外的内在图像数据集 (IIW) [4] 在相对反射率的人工判断的形式提供大量的地面真理：5230 室内图像共有 872161 对比较，对每个图像的比较。与麻省理工学院

的内在影像数据集 [24]。在那里，作为绝对的地面真理，这些成对的比较采取三个值：相同，更轻，或更暗，他们对这些具有发言权。

- 通过分层稀疏编码 (HSC) [6], [17] 或深卷积神经网络 (CNN) 计算的丰富的上下文特征 [14], [12] 提供的信息从细到粗，从小到大在输入图像中每个补丁的上下文特征矢量表示，使一个简单的和直接的分类方法不用依赖任何手工设计全局先验的或昂贵的推理算法。

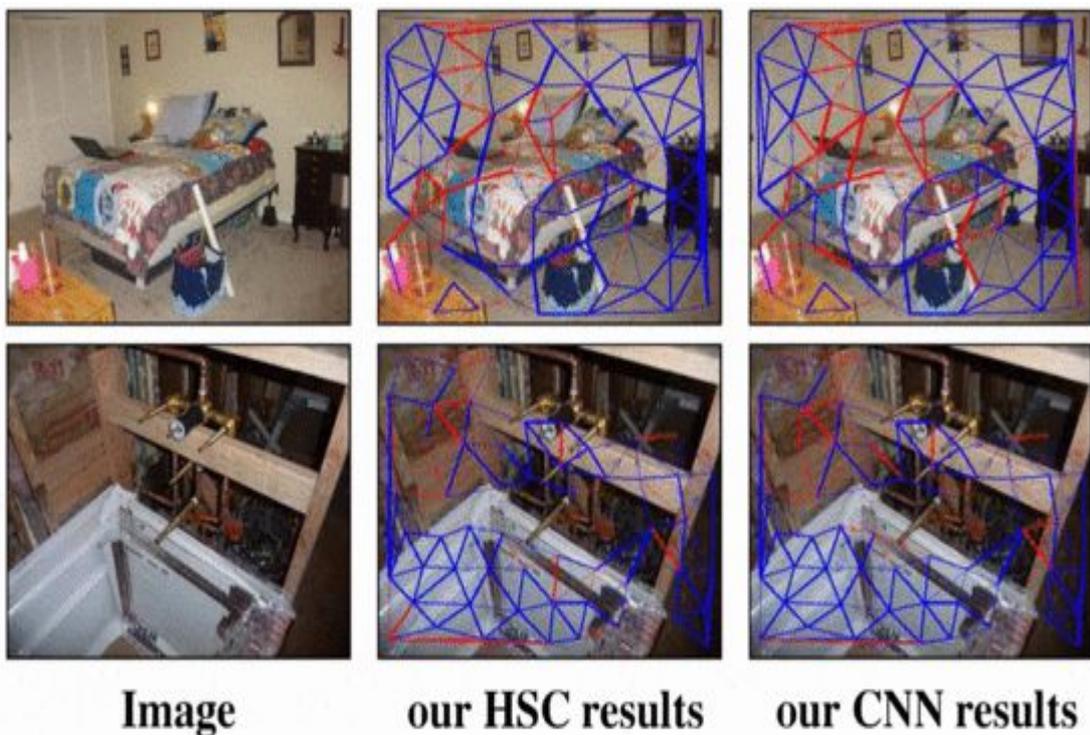


图 3。例如图中两两比较结果。我们展示出 HSC 和美国有线电视新闻网分类结果对一些 IIW[4] 测试图像。图上的箭头从亮到暗的区域，根据地面的情况，用无向边表示等亮度。边缘厚度表示人的信心，而边缘颜色表示分类正确性（红色的错误）。我们的错误往往发生在具有挑战性的地方，如人的信心较低（细红线）和独立的轻背景（厚红线）区域。

几乎所有过去的方法集中于区分反射率和阴影这两个方面，我们直接的方法（图 2）是从这里出发，通过学习，或通过强加或发现更多先验来区分这两个方面[24]之间的特性和边缘可用于约束各个方面[15]，[16]，[2]，[18]，[9]，[22]，[8]，[26]，[4]。任何一方面的设计完全没有先验算

法；只有那些自己从数据中了解到隐含的先验在我们的系统中存在作用。

令人意外的是，有大量成对的原油相对反射率，我们能够利用简单的局部线性分类器与亮度比较精度等同于或优于国家的最先进的模型依赖于多个先验和致密的 CRF 全球推理[ 4 ]。参见图 3 中的示例结果。

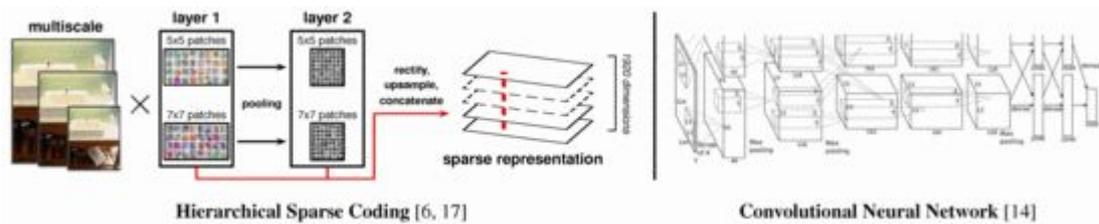


图 4. 深层特征提取。我们考虑两种不同的程序，如图 2 左侧的深特征提取器的实现：我们适应[17]的稀疏编码策略，它建立在[6]，和编码对多个补丁字典多个图像尺度的基础上。第二编码层操作于第一次汇集输出。在相应的空间位置串联跨图层稀疏编码产生每个像素的描述符。仅词典的一个子集在这里是可视化的。右：我我们利用[ 14 ]一个图像块作为输入的卷积神经网络结构和治疗在最后一层的激活作为一种局部描述子的补丁。

我们的工作是在[25]中提出的全球集成方法，其中，在多个尺度简单成对强度差被用作特征和全局亮度，从所有这些存在成对线索的整个图像中对排序结果做补充。虽然这种模式表现在具有挑战性的合成图像，目前还不清楚它是如何被应用到自然的图像与更复杂的视觉外观中。

有图像分解，还可能被用于由本征图像模型用作更多的信息提高例如基于 CRF 的[ 4 ]和[ 25 ]全球化以嵌入方法。

我们的工作第一个使用复杂的深功能，通过一个简单的本地分类规则预测自然图像的亮度。它绕过了固

第 2 节中详细介绍了两种丰富的特征表示，HSC 和 CNN，我们认为用作补丁信息。第 3 节涵盖了我们所学到的模型为成对的亮度关系，以及提取的过程中，从这两个模型，利用线性分类器直接构建的亮度地图。第 4 条规定的实验结果和基准，而第 5 节总结。

## 第二部门

### 补丁描述

在其他视觉应用取得成功的情况下，我们认为这两个分层稀疏编码 [17], [6] 和卷积神经网络 [14] 用作我们的系统架构特征提取工具。图 4 说明这些方法，我们现在简单总结一下每个。

#### 2.1. 分级稀疏编码

我们借用 [17] 的补丁表示策略，反过来又建立在 [6] 的工作之上。对于 [17]，分层稀疏编码是一种有效的算法，用于获得丰富、高维、但稀疏的像素特征描述符。这种稀疏性转化为快速应用线性分类器在图像密集。

显著特点相比，CNN 特点包括 HSC 生成性和深刻网络的多层片提取。因此，他们捕捉多个不同的抽象层次。在这里，我们回顾了 HSC，读者参考 [17] 了解更多详情。

在一个标准的稀疏编码设置，补丁是表示为一个稀疏线性组合。大部分原子来自一个超完备补丁字典。用组合系数向量，编码此问题：

$$\operatorname{argmin}_z \|x - Dz\|^2 \quad \text{s.t.} \quad \|z\|_0 \leq K$$

在实践中虽然没有一个很效率的计算程序用来寻找确切的最优，但是利用贪婪逼近算法的工作可以获得特

征，如 [19], [21]。我们可以简单地把它作为一个特征向量的分类或回归的设置。

在针对相同的字典编码时，每个补丁都会对应于同一二维网格的每一个像素，或者等价地，对应于一个（稀疏）图像通道。在分层稀疏编码设定时候，我我们把这个通道的输出从一个初始层稀疏编码作为输入，汇集后，就得到另一层稀疏编码的高维原子的新字典。

我们的稀疏编码层的层次结构也是在这个意义上，在每一层中，我们针对几个不同的字典进行编码，并且我们在三个不同尺度编码图像。图 4 表示出该属性。一个输入块的特征向量是从通过网络的所有路径稀疏系数矢量的串联。拼接后，我们纠正特征向量：

$$z \leftarrow [\max(z^T, 0), \max(-z^T, 0)]^T \quad (5)$$

对每个阶段的字典在编码生成的方式学习，从训练图像采样的补丁的集合（或从上一层的输出训练图像表示），并应用 MI-K SVD 算法 [6] 近似解来解决：

$$\operatorname{argmin}_{D, Z} \left[ \|X - DZ\|_F^2 + \lambda \sum_{i=0}^{N-1} \sum_{j=0, j \neq i}^{N-1} |d_i^T d_j| \right] \quad (6)$$

s.t.  $\forall i, \|d_i\|_2 = 1$  and  $\forall n, \|z_n\|_0 \leq K$

其中  $\|\cdot\|_F$  表示 Frobenius 范数，是指定的稀疏水平。

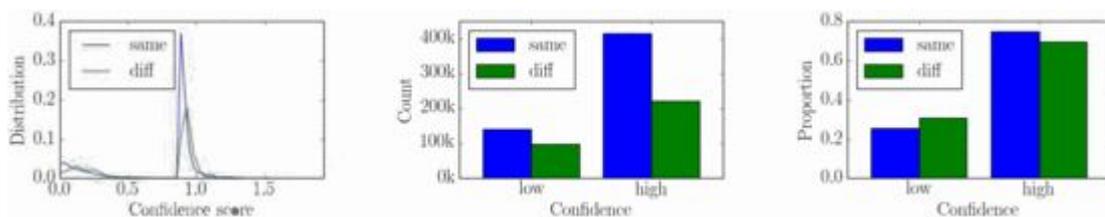


图 5. 人类反射光谱的统计。左：在标准数据集[ 4 ]，人类通常在他们的相对反射率高的地方标注同反射率和轻/暗对。中：每一个反射型的高和低置信地真对。权利：每种类型的地面真理对具有高和低的信心的比例。

在我们的实验中，我们学习词典  $N = 32$  和 64 个原子的每个  $5 \times 5$ ,  $7 \times 7$  和  $9 \times 9$ ，以及第一层的补丁。第二层采用 64 原子和 32 维的第一层输出补丁词典。级联和整顿在图像中的每个像素产生一个 1920 维的特征向量。

### 2.1. 卷积神经网络

我们使用的是[ 12 ]实施的 7 层卷积神经网络[ 14 ]，以最终完全连接层 4096 维激活作为输入到网络提供了一个补丁的特征描述符。每个补丁是从一个窗口周围的位置（顶点）围绕的标记对（边缘）的位置（顶点）的窗口取标准数据集。注意，整个网络可以通过反向传播的判别方式进行传播，而相比之下，HSC 更注重学习词典权重生成。

## 第三部分

### 亮度划分

人类能够在两个不同地点的同一场景[ 4 ]反射率进行比较。我们充分

利用这一点对建筑物进行自动分类，从而复制人对相对亮度的判断。

以下等式（3）和选择一个线性形式为与用作特征前述部分的任一的 HSC 或 CNN 表示，我们有

$$f(z_i, z_j) = w^T(z_i - z_j) \quad (7)$$

我们学会山脊上反射人体地面实况数据排名回归 HSC 分类权重：

$$\min \varepsilon(w) = \sum_{i,j} \log(1 + \exp(-J_{ij} w^T(z_i - z_j))) + \gamma w^T w \quad (8)$$

在这里：

$$J_{ij} = \begin{cases} 1, & R_i^h > R_j^h \\ -1, & R_i^h < R_j^h \end{cases} \quad (9)$$

$$C_{ij} = \text{confidence in } J_{ij} \quad (10)$$

这里是指人对相对反射率（亮度）标准数据集的评级。对规范进行校验。对于在人类判断反射每个例子中，我们创建了两个虚拟的例子，以迫使预测向零。尽管目标是使得  $(R_i^h = R_j^h)$ ，我们需要使得  $R_{ij} = 1$  和  $R_{ij} = -1$ 。而且要培养我们的 CNN

模型，我们使用标准的交叉熵分类避免损失。

图 5 提供了一些反映人类对相对反射率评级的数据。虽然标准数据集包含更多具有不同反射率（由人来判断）的一对多的例子。我们认为在每种类型的训练实例的总体数量的轻微不平衡因而不存在问题。

我们可以等价地从特征角度描述相对亮度的差异，并进行线性预测，或利用作用于点的特征表示一个亮度势函数。定义亮度势：

$$g(z) = w^T z \quad (11)$$

然后我们的相对亮度分类就会产生如下差异：

$$f(z_i, z_j) = w^T (z_i - z_j) = g(z_i) - g(z_j) \quad (12)$$

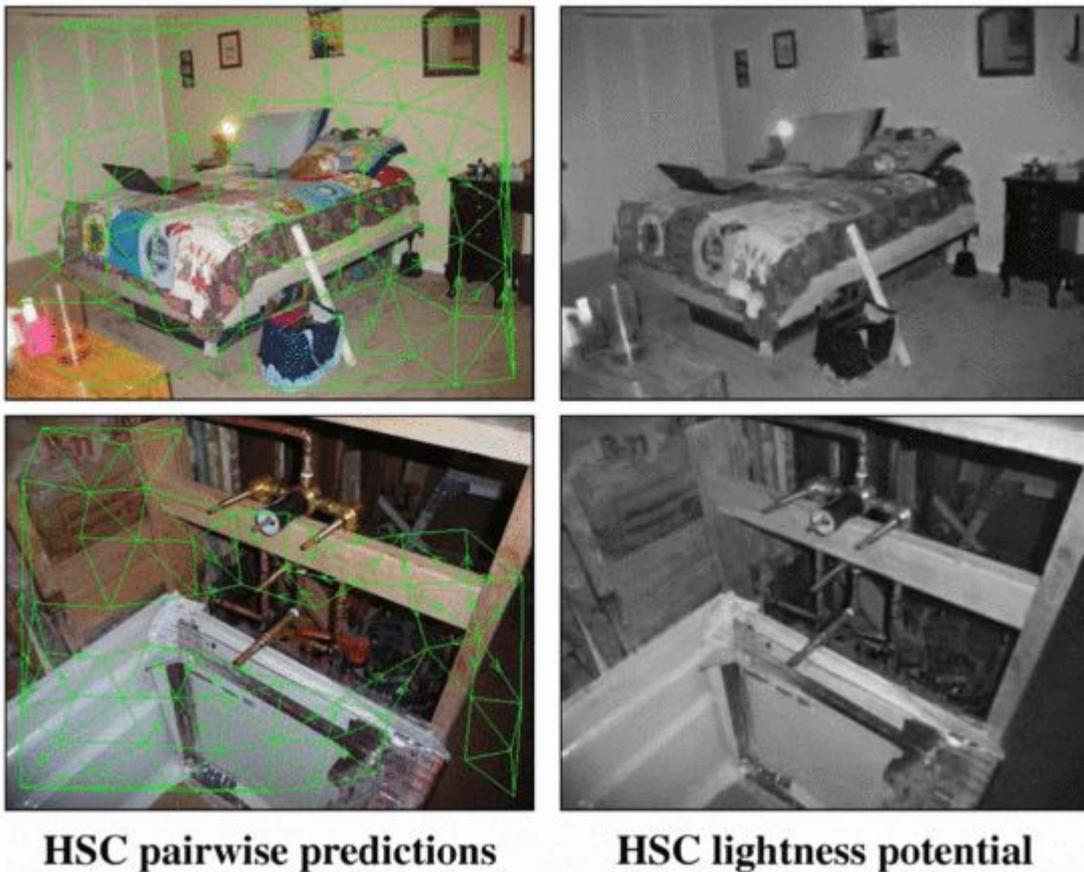


图 6.全局亮度等级。从训练的相对亮度判断一个线性差分模型，我们提取每个像素的亮度电位（全局亮度等级）。左：我们的位于所述图形节点的像素之间的相对亮度的 HSC 模型的成对的结果。从箭头指向对模型进行预测。右：密集使用预测功能 HSC 相对轻盈的部分。

这个往往是有效的，我们可以在每个像素局部地评价亮度潜力来恢复图像中的所有像素的光亮度。

在 HSC 的描述符的情况下，这样的评价是廉价的而且是极其稀少的。对 [ 17 ] 的 HSC 产生密集的图像描述符的实现，图 6 表示出采用 HSC 特征获得示例电位。它们是类似于在传统的固有图像分解的反射率信道的缩放版本。

尽管密集的评价也可以用 CNN 的特点，CNN 实现，我们用 [12] 有针对性地在任意大小的输入情况下做完全卷积，这中独立的图像中的补丁运行十分昂贵。因此，我们专注于匹配人判断和补丁匹配的任务。

用在标准数据集的基准，在两个点反射，一个离散的判断从较轻，颜色较深区域开始，同样呈现基于反射比的试验：

$$\hat{J}_{ij}(R; \delta) = \begin{cases} 1, & \text{if } \frac{R_i}{R_j} > 1 + \delta \\ -1, & \text{if } \frac{R_j}{R_i} > 1 + \delta \\ 0, & \text{otherwise} \end{cases}$$

在  $\delta$  在一定阈值内，相对反射率的变化被认为是微不足道的。

我们的线性分类器可以被解释为执行反射率比测试，根据变换后的反射域日志进行差异测试：

$$g(z) = w^T z = \alpha \log R \tag{14}$$

$$\begin{aligned} f(z_i, z_j) &= g(z_i) - g(z_j) = \alpha \log R_i - \alpha \log R_j \\ &= \alpha \log \frac{R_i}{R_j}. \end{aligned} \tag{15}$$

请注意， $f(z_i, z_j)$  不改变其正负号，0 或  $\pm 1$  不管是什么，它可以被认为是内部的光。放参数控制的速率感知反射率涉及亮度的潜力，并且可以认为是内部差到外部感知亮度差的灵敏度的参数。

现在给出预定义的阈值的反射率测试例证，我们调整使得对于可感知的亮度差异最为匹配的训练集人的判断：

$$\hat{J}_{ij}(f; \alpha, \delta) = \begin{cases} 1, & \text{if } f(z_i, z_j) > \alpha \log(1 + \delta) \\ -1, & \text{if } f(z_j, z_i) > \alpha \log(1 + \delta) \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

注意  $\alpha$  (在整个  $\alpha \log(1 + \delta)$ ) 训练集上) 不是每一个图像或每对查询点进行了优化。

以前的工作是分解以前的图像 [ 10 ]，[ 4 ] 没有类似地专门为成对亮度判断任务调整。虽然它们的反射率是直接建模，并不需要像我们这样的解释，以确保公平基准的比较，我们把他们的输出反射图和在算法上做了最优性能报告 (重标度的一个等价形式)。

## 第四部分

### 结果分析

我们使用的内在形象在野外 (IIW) 数据集 [ 4 ] 进行训练并测试我们的亮度进行分类。在总计 5230 个图像中每一个图像 (图 3 中的节点和边) 之间都有查询点。在关于“哪一点具有较深的表面颜色?” 的多次问询中 [ 4 ] 共获得 4880372 个回复。1381 亚马逊的 Mechanical Turk 工人汇总个人数据共 875833 比较 5230 张照片, 用于测试结果是否一致。

这个小组两两比较的已被用来基准几种反射率模型 [ 4 ]。提出了加权人的分歧率 (whdr) 来衡量人类的判断率, 算法是通过点对每个判断的置信度:

$$WHDR_{\delta}(J, R) = \frac{\sum_{ij} C_{ij} (J_{ij} \neq \hat{J}_{ij}(R; \delta))}{\sum_{ij} C_{ij}} \quad (17)$$

在这里,  $J_{ij}$  和  $\hat{J}_{ij}$  是人类地面实况和预测。请注意, 我们分别计算每个图像, 然后将所有照片横跨在数据集中以产生整体基准数。

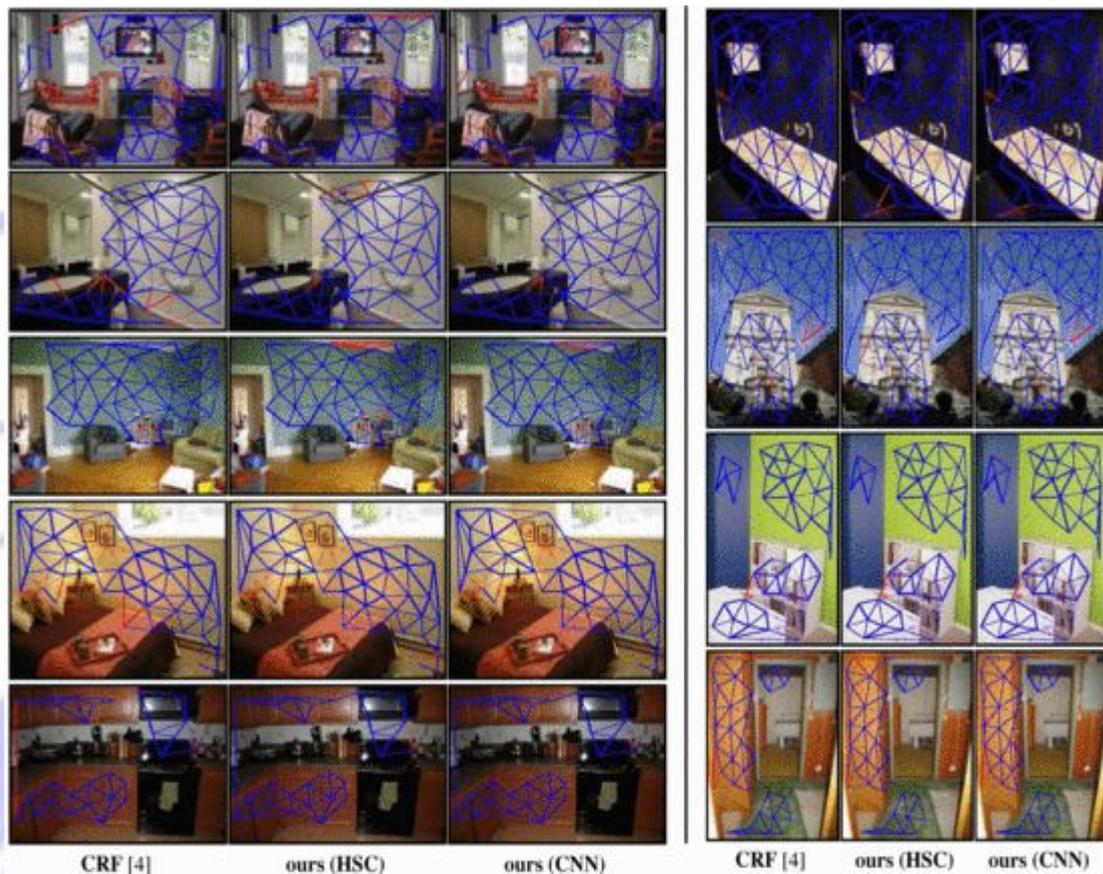


图 7. 良好的亮度预测。我们想象我们的模型 (HSC 和 CNN) 的两个版本的基准亮度成对预测的 CRF 模型 [4]。如在图 3 中, 上图的边箭头指向根据地面实况 (朝向较暗端点) 和颜色表示分类正确性 (红色失误)。在这里, 我们展示了所有模型执行的例子。

国家的最先进的结果是由贝尔等人完成的。与相关文献[4]，探讨反射和遮光等多种形式先验固有图像分解致密 CRF 模型。注意，人类受试者不一定彼此一致，或者每个单独的所有查询对之间。根据 WHDR 人的一致性达到 7.5% [4]。

我们引入一个额外的度量，分类错误率 (1, ACC)，在每个查询的边缘衡量算法的性能，而不是利用所有查询边/图像平均的精度 (或误差 whdr)。我们定义的分类精度 (ACC) 为：

$$ACC_{\delta}(J, R; \delta) = \text{mean}(J_{ij} = \hat{J}_{ij}(R; \delta),$$

对于我们的实验中，我们把数据集分成 80% 规范培训实例与 20% 个测试例

子。我们排序图像 ID 的数据集，然后以每 5 个为一组，并使用第一个作为测试而其他的作为训练。我们重新评估了 [4] 报道了这种分裂几种方法，以确保分数的匹配 [4] 和特定的训练测试分裂的影响可以忽略不计的影响。

表 1 比较了我们的结果与国家的最先进的 CRF 方法，以及其他竞争的方法。我们的模型进行培训，发现约 80% 的数据并没有进一步的优化。表 1 显示的是，尽管我们的本地分类模型的简单，我们的 CNN 的表现与上测试集的最佳方法的标准杆，并领先其他所有的。在这里，CNN 从随机初始权训练，但是我们没有看到他们之间的显著性能差异。



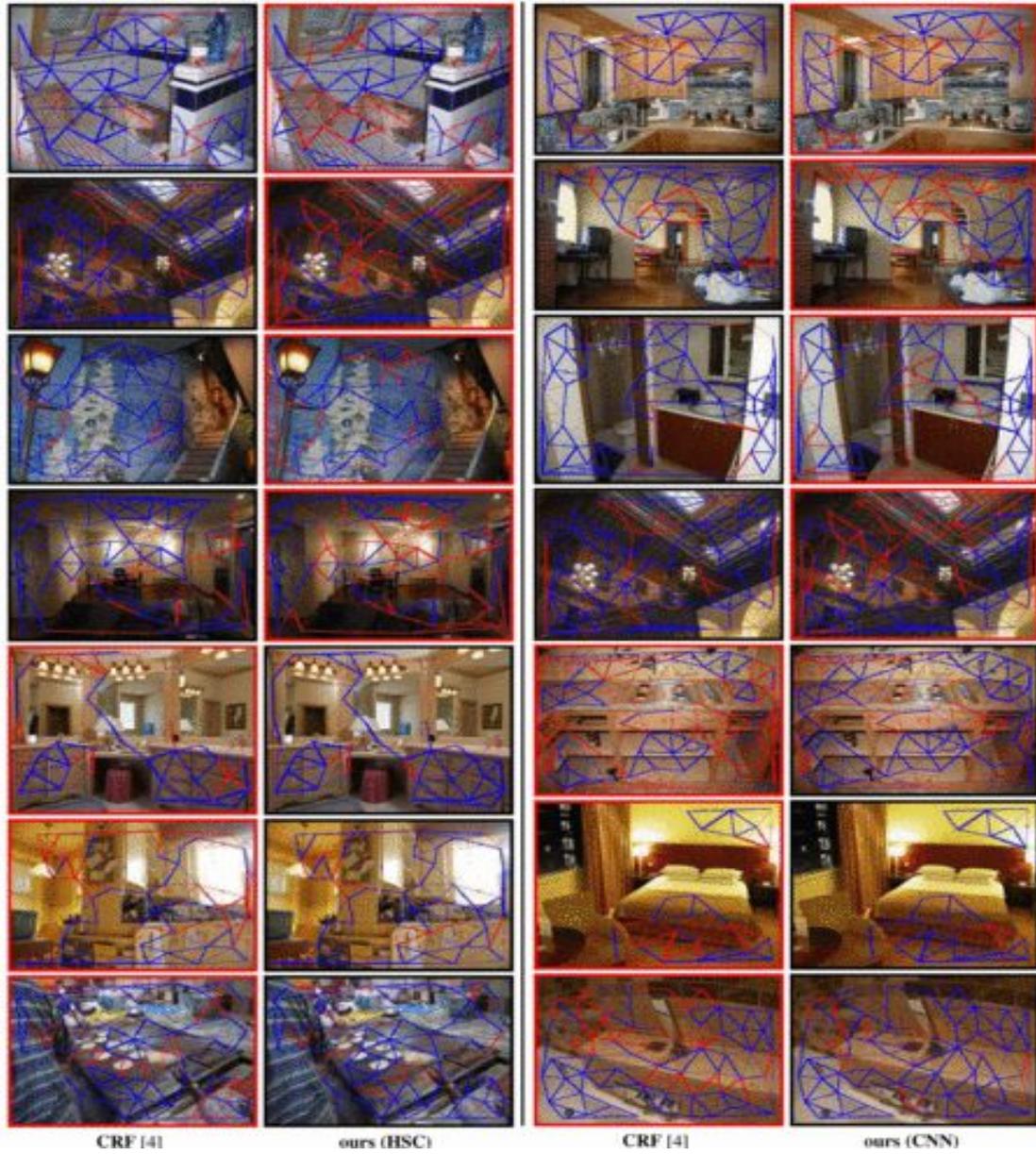


图 8.亮度预测的误区。在图 7 中，我们想象我们的模型（HSC 和 CNN 的版本）的预测和的 CRF 模型[4]。在这里，我们展示一个它们之间有很大的性能差距的例子。左：慢性肾功能衰竭和 HSC 模型之间显著的性能差距的例子。右：用实例和 CRF 和 CNN 模型之间显著的性能差距。对于每一对，对模型分析结果不佳的结果会在自己的预测图像周围出现红色边框。

	WHDR (%)	Error Rate (%)
Ours (HSC)	20.9	24.5
Ours (CNN)	18.3	22.3
Ours (CNN-ImageNet)	<b>18.1</b>	<b>22.0</b>
CRF [4] (rescaled)	18.6	22.3
Retinex-Color [10] (rescaled)	19.5	23.3
Retinex-Gray [10] (rescaled)	19.8	23.8
Shen and Yeo [22] (rescaled)	23.2	26.1
Zhao <i>et al.</i> [26] (rescaled)	22.8	26.4
CRF [4]	20.6	25.6
Retinex-Color [10]	26.9	32.4
Retinex-Gray [10]	26.8	32.3
Shen and Yeo [22]	32.5	35.1
Zhao <i>et al.</i> [26]	23.8	28.2

表 1. 在野外测试结果固有的图像。对于每种算法，我们显示了加权分歧率（whdr，越低越好），以及分类的点标记在地面真理之间的明度变化的符号错误率。它们包括我们自己重新评估竞争的方法，这是密切配合的表现在[4]。此外，我们报告的相互竞争的方法的放大版本，其特异性地优化其用于成对分类任务输出性能。我们的算法是由[4]国家的最先进的性能开发的 CRF 的做法相提并论。我们建议读者参考[4]进行比较，能够扩展当前的工作。

图 7 显示了用我们的方法处理的示例图像，以及具体的做法。在这些图像中的查询边缘往往存在大致均匀的补丁，导致判断往往是不明确的。图 8 显示了一种方法在一种技术上远远优于另一种的示例图像。我们的模式和 CRF 的方法有不同的故障处理模式。

我们与 HSC 功能的方法相比似乎使比 CRF 在低光和高光条件下区分点具有更多的错误，而错误的特征是：它往往误以为不同的亮度（带箭头的边缘）在低光一样轻，而同样亮度（没有箭头的边缘）却认为存在差异。CNN

的方法提供了一个更统一的改进，是的 CRF 照明条件改善得更均匀。图 9 提供的可视化的分类性能的函数可以提高程序的性能，如功能（平均强度）和纹理（像素密度局部方差）修补。

CRF 方法似乎在小结构的纹理区域还是有许多的错误。这个错误模式可以由 CRF 模型所使用的反射率稀疏（每图象约 20 反射），并分段恒定先验地进行说明。当图像包含许多不同的反射率，如每个小结构呈现不同的反射率，CRF 会给出违反假设的不准确的分解结果，因此产生更多的错误。

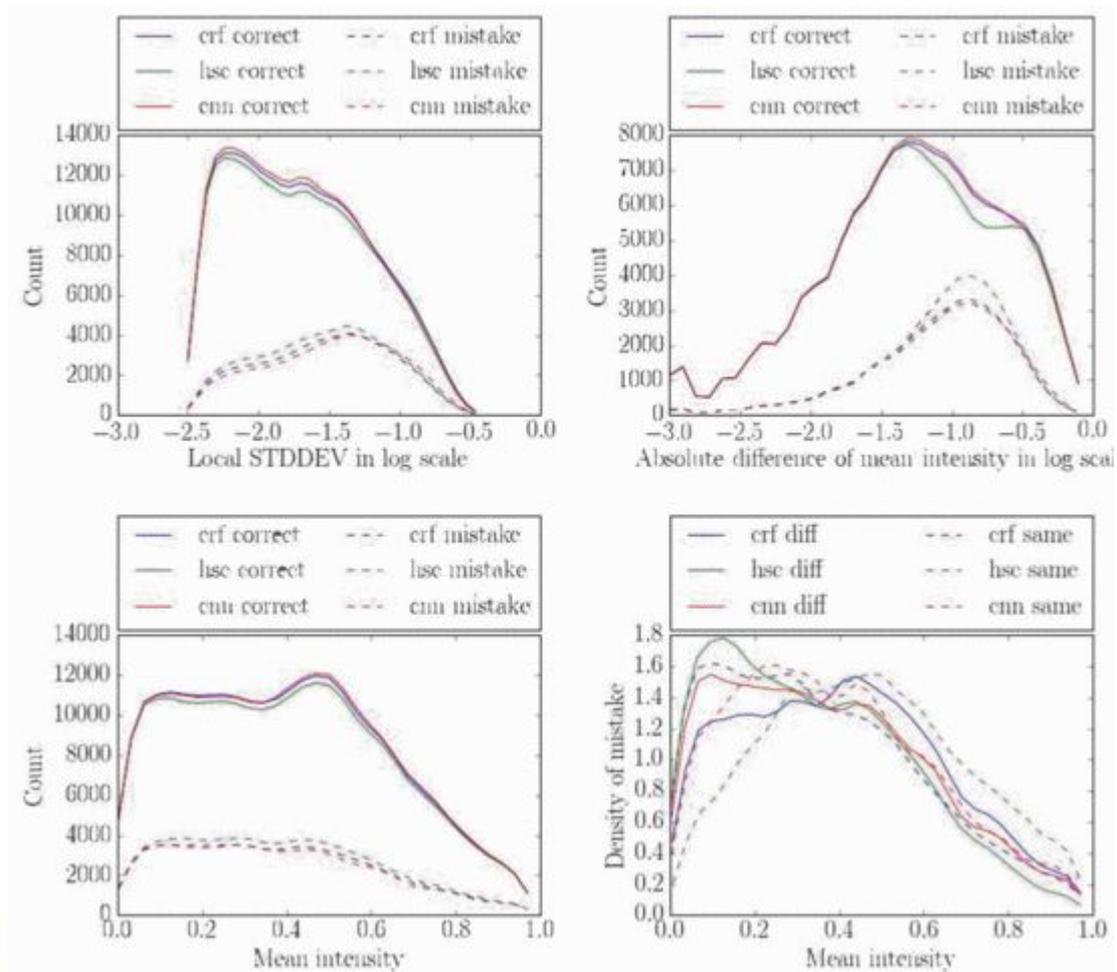


图 9.地方补丁统计功能的错误。左上：密对补丁强度方差的绘图错误显示，揭示了 CNN 的模型在平滑区域较低的错误计数（低像素强度方差）。右上：HSC 模型使得更多的错误时比较两片大大不同强度。左下：我们比较模型误差相对于两个补丁的平均强度。我们的 CNN 模型比别人多补丁平均亮度范围内均匀则越好。右下：我们按类型组织的错误（地面实况不同或相同的亮度贴片对正确分类）。HSC 使得在低光照地面存在更多的失误，在高光条件下也会产生错误。

## 第五部分

### 总结

我们展示了一个本地的分类模型，执行更复杂的步骤以完成恢复相对亮度的任务。使用丰富的补丁，通过分层稀疏编码或卷积神经网络，获得大量的训练数据，使我们能够学习比以前一个更好的本地模型。这样的模式

开拓新领域的探索，可以用于内在图像的分解这样的经典问题。

### 致谢

感谢和 Ayan Chakrabarti 所进行的有价值的探讨。

## REFERENCES 参考文献

1. E. H. Adelson and M. Gazzaniga

*pp. 339-51, 1999, MIT Press*

2. J. T. Barron and J. Malik

"Intrinsic scene properties from a single RGB-D image"

*IEEE Conference on Computer Vision and Pattern Recognition, 2013,*

3. H. G. Barrow and J. M. Tenenbaum

"Recovering intrinsic scene characteristics from images"

*Computer Vision Systems, pp. 3-26, 1978*

4. S. Bell, K. Bala and N. Sanvoly

"Intrinsic images in the wild"

*ACM Trans. on Graphics, 2014,*

5. B. Blakeslee, W. Pasioka and M. E. McCourt

*Vision Research*

*no. 45, pp. 607-15, 2005*

6. L. Bo, X. Ren and D. Fox

"Multipath sparse coding using hierarchical matching pursuit"

*CVPR, 2013*

7. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei

"Imagenet: A large-scale hierarchical image database"

*IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255,*

8. E. Garces, A. Munoz, J. Lopez-Moreno and D. Gutierrez

"Intrinsic images by clustering"

*Computer Graphics Forum (Eurographics Symposium on Rendering),*

9. P. Gehler, C. Roth, M. Kiefel, L. Zhang and B. Scholkopf

"Recovering intrinsic images with a global sparsity prior on reflectance"

*Neural Information Processing Systems, 2011*

10. R. Grosse, M. K. Johnson, E. H. Adelson and W. T. Freeman

"Ground truth dataset and baseline evaluations for intrinsic image algorithms"

*International Conference on Computer Vision,*

11. B. Horn

*Computer Graphics and Image Processing*

*vol. 3, pp. 277-99, 1974*

12. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Gir-shick, S. Guadarrama and

T. Darrell

"Caffe: Convolutional architecture for fast feature embedding"

*arXiv preprint arXiv:1408.5093, 2014*

13. F. Kelly and S. Grossberg

"Neural dynamics of 3-d surface perception: figure-ground separation and lightness

perception"

*Perception and Psychophysics, no. 62, pp. 1596-618, 2000*

14. A. Krizhevsky, S. Ilya and G. E. Hinton

"Imagenet classification with deep convolutional neural networks"

*Neural Information Processing Systems,*

15. E. H. Land and J. J. McCann

"Lightness and retinex theory"

*Journal of Optical Society of America, vol. 61, no. 1, pp. 1-11, 1971*

16. Z. Liao, J. Rock, Y. Wang and D. Forsyth

"Non-parametric filtering for geometric detail extraction and material representation"

*IEEE Conference on Computer Vision and Pattern Recognition, 2013,*

17. M. Maire, S. X. Yu and P. Perona

"Reconstructive sparse code transfer for contour detection and semantic labeling"

*ACCV, 2014*

18. I. Omer and M. Werman

"Color lines: image specific color representation"

*IEEE Conference on Computer Vision and Pattern Recognition,*

19. Y. C. Pati, R. Rezaifar and P. S. Krishnaprasad

"Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition"

*Asilomar Conference on Signals, Systems and Computers, 1993,*

20. W. D. Ross and L. Pessoa

"Lightness from contrast: a selective integration model"

*Perception and Psychophysics*, no. 62, pp. 1160-81, 2000

21. R. Rubinstein, M. Zibulevsky and M. Elad

Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit

2008

22. L. Shen and C. Yeo

"Intrinsic images decomposition using a local and global sparse representation of reflectance"

*IEEE Conference on Computer Vision and Pattern Recognition*,

23. Y. Tang, R. Salakhutdinov and G. Hinton

"Deep lambertian networks"

*International Conference on Machine Learning*,

24. M. Tappen, W. Freeman and E. Adelson

"Recovering intrinsic images from a single image"

*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005,

25. S. X. Yu

"Angular embedding: from jarring intensity differences to perceived luminance"

*IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2302-9,

26. Q. Zhao, P. Tan, Q. Dai, L. Shen, E. Wu and S. Lin

"A closed-form solution to retinex with nonlocal texture constraints"

*IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012*

