

5 月 30 日-6 月 5 日工作汇报

Ku Jui

May 2023

Contents

1	资料调研	2
1.1	局限性	2
1.1.1	Retinex	2
1.1.2	评价指标	2
1.1.3	数据集	3
1.2	客观评价指标	3
1.2.1	峰值信噪比 (Peak-Signal to Noise Ratio, PSNR)	4
1.2.2	结构相似性 (Structural Similarity, SSIM)	4
1.2.3	均方误差 (Mean Square Error, MSE)	5
1.2.4	信息熵 (Information Entropy, IE)	5
1.2.5	标准偏差 (Standard Deviation, STD)	5
1.2.6	亮度顺序误差 (Lightness Order Error, LOE)	6
1.3	主观评价指标	6
1.3.1	差分平均意见分数 (Differential Mean Opinion Score, MOS)	6
2	个人工作进展	6
2.1	思考	6
3	下周工作计划	7

1 资料调研

1.1 局限性

现有的方法仍有很大的改进空间，例如如何在提高亮度的同时消除产生的噪声，如何避免颜色失真现象等。一些现有的方法可以有效地解决一个问题，但往往会忽略了其他问题。不同的方法在不同的数据集上往往具有不同的优势，即在不同的评估标准下有不同的优势。

例如，在 LOL 数据集（目前应用最广泛的数据集）下 URetinexNet 具有十分卓越的性能，因为其在 PSNR、SSIM、MSE 指标上目前排名第一，但是 URetinexNet 在 SCIE 数据集下，MBLLEN 方法在 PSNR、SSIM、MSE 指标中排名第一。



Figure 1: Visual results of different algorithms on low-illumination images sampled from the VE-LOL-L dataset.

Fig. 1显示了作为一个室内图像的 VE-LOL-L 数据集的增强效果。很明显，Fig. 1c和 Fig. 1d是红色的，但 Fig. 1b和 Fig. 1e呈现的颜色更为橙色。同样，Fig. 1c显示的图像中桌面物品的颜色偏差也很大，且图像的整体颜色不一致。

1.1.1 Retinex

与直接端到端网络学习的结果相比，基于深度学习和 Retinex 理论的方法有时具有更好的增强效果。

Retinex 模型方法将低照度图像分解为照度图像和反射图像，通过特定的算法减少甚至消除入射图像的影响，保留物体本质的反射特性图像。然而，该算法也存在一定的局限性。以反射贴图作为增强结果，会导致细节丢失、颜色失真等问题。此外，由于反射图大量噪声产生，该模型忽略了噪声。反射映射本身的计算是一个不适定问题，只能通过近似估计来计算。

1.1.2 评价指标

PSNR、MSE、SSIM 和 IE 是比较经典的、流行的图像评价指标，但这些指标还远未达到人类视觉感知的评价能力。同时，虽然这些评价指标可以评价图像质量，但它们不能

表达低照度增强图像与实际图像之间的关系。因此，在低照度图像增强的评价指标中，在实现人类视觉效果与机器感知之间的平衡还有相当的局限性。

但是，主观视觉效果较好的增强算法在客观评价中并没有取得良好的效果，这正是目标评价指标需要改进的地方。

同时，现有的客观评价指标用于其他研究领域，如图像除雾、除雨、降噪等。这些指标不是为微光图像增强而设计的，这些评价指标远未达到人类的自然感知效果。

1.1.3 数据集

LOL 和 LE-LOL-L 数据集的测试图像的客观评价指标值比 SCIE 和 LIME 数据集更适合人类视觉效果。用成对数据集训练的算法将得到更明显的增强效果。从数据中可以看出，基于监督学习的方法比其他方法更能取得更好的效果。

然而，无监督学习、Zero-Shot 学习和半监督学习是当前的发展趋势，主要是因为低光照环境下很难获得成对的图像。与其他方法相比，基于完全监督学习的方法泛化能力较低，适用性较差。此外，无监督学习方法可以通过设计适当的损失函数和网络结构，引入更多反映环境特征的先验知识。

1.2 客观评价指标

图像的客观评价指标很多，不同角度的评价标准也不尽相同。每种评价标准都有其相应的优缺点。到目前为止，还没有明确设计用于微光图像增强的评估指标。现有的图像质量评价方法可分为完全参考评价指标和无参考评价指标。

完全参考指标

完全参考评价指标需要有一个完整的参考图像（即原始图像或专家标注的高质量图像）作为参照。通过将待评价的图像与参考图像进行比较，计算它们之间的差异来评估图像的质量。常见的完全参考评价指标包括：均方误差（Mean Square Error, MSE）、峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）、结构相似性指数（Structural Similarity Index, SSIM）

无参考评价指标

无参考评价指标不需要参考图像，仅通过待评价图像自身的特征来评估图像质量。这些指标主要基于图像的统计特性、信息熵等，无需外部参考信息。常见的无参考评价指标包括：1) 图像清晰度评估：通过图像的锐度、对比度等特征来评估图像的清晰度和质量。2)

图像亮度评估：基于图像的亮度直方图、灰度均值等特征，评估图像的亮度质量。3) 图像块失真评估：通过比较图像中局部块之间的差异，来评估图像的块状失真程度。

Table 1 列出了最新的几个经典关键的质量评价指标。

Abbreviation	Full-/Non-Reference
MSE (Mean Square Error)	Full-Reference
MAE (Mean Absolute Error)	Full-Reference
SNR (Signal to Noise Ratio)	Full-Reference
PSNR (Peak-Signal to Noise Ratio)	Full-Reference
LPIPS (Learned Perceptual Image Patch Similarity)	Full-Reference
IFC (Information Fidelity Criterion)	Full-Reference
VIF (Visual Information Fidelity)	Full-Reference
SSIM (Structural Similarity Index)	Full-Reference
IE (Information Entropy)	Non-Reference
NIQE (Natural Image Quality Evaluator)	Non-Reference
LOE (Lightness Order Error)	Full-Reference
PI (Perceptual Index)	Non-Reference
MUSIQ (Multi-scale Image Quality Transformer)	Non-Reference
NIMA (Neural Image Assessment)	Non-Reference
SPAQ (Smartphone Photography Attribute and Quality)	Non-Reference

Table 1: Abstract of objective quality evaluation index of the image.

1.2.1 峰值信噪比 (Peak-Signal to Noise Ratio, PSNR)

峰值信噪比 (PSNR) 是应用最广泛的指标之一。单位为 dB。它用于测量两个图像之间的差异。如压缩图像与原始图像、压缩图像质量评估、恢复图像与实际图像、恢复算法性能评估等。PSNR 值越高，失真越小。PSNR 的公式为 Eq: 1

$$PSNR = 10 \times \log \frac{MaxValue^2}{MSE} \quad (1)$$

其中 MSE 是两个图像的均方误差；MaxValue 是图像像素的最大值。

1.2.2 结构相似性 (Structural Similarity, SSIM)

结构相似性 (SSIM) 是参考图像质量评价中应用最广泛的标准。SSIM 用于突出值范围为 0-1 的两幅图像之间的亮度、对比度和结构相似性；越接近 1，两个图像就越相似。假设 x 和 y 是两个输入图像，SSIM 的公式为 Eq: 2

$$SSIM = [l(x, y)]^\alpha [C(x, y)]^\beta [S(x, y)]^\gamma \quad (2)$$

其中， $l(x, y)$ 是亮度比较， $C(x, y)$ 是对比度比较， $S(x, y)$ 是结构比较。 α, β, γ 大于 0，用于调节三部分比重。 $l(x, y)$ 、 $C(x, y)$ 和 $S(x, y)$ 分别为 Eq.3, Eq.4, Eq.5

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (3)$$

$$C(x, y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c^2} \quad (4)$$

$$(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (5)$$

其中 μ_x 和 μ_y 表示两幅图像的平均值, σ_x 和 σ_y 分别表示两幅图的标准差。 σ_{xy} 表示两个图像的协方差。 c_1 、 c_2 和 c_3 是常数, 以避免分母为 0。

1.2.3 均方误差 (Mean Square Error, MSE)

均方误差 (MSE) 也是衡量图像质量最常用的指标之一。它是指估计值与真实值之间的平方差的期望值。在图像处理算法中, 它是被处理后的图像像素值与原始像素值之间的平方差的平均值。MSE 表达式如 Eq.6

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [g(x, y) - \hat{g}(x, y)]^2 \quad (6)$$

其中, M 是图像的高度, N 是图像的宽度。 $g(x, y)$ 和 $\hat{g}(x, y)$ 分别表示原始图像和增强图像。MSE 的值越小, 图像质量越好。

1.2.4 信息熵 (Information Entropy, IE)

信息熵 (IE) 反映了图像所携带的信息量。信息熵越大, 图像信息越丰富, 质量越好。IE 用于比较不同图像中信息内容的差异。例如, 在同一区域拍摄的图像也会由于额外的拍摄时间而具有不同的信息内容。IE 表达式 Eq.7

$$H = - \sum_{k=0}^M p(k) \log_2 p(k) \quad (7)$$

其中 $p(k)$ 是灰度级 k 的概率密度, M 是最大灰度级

1.2.5 标准偏差 (Standard Deviation, STD)

标准差 (STD), 也称为标准差, 表示每个数据项与平均值的平均距离。STD 也是方差的平方根。STD 可以反映数据集的分散程度。与图像相比, STD 反映了插图与平均值之间的分散程度, 是特定范围内图像对比度的衡量标准。标准偏差越大, 图像中包含的信息越多, 视觉效果越好。STD 的表达式 Eq.8

$$STD = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N f(i, j) (f(i, j) - \delta)^2}{M \times N}} \quad (8)$$

其中 M 是图像的高度, N 是图像的宽度, $f(i, j)$ 表示图像的 (i, j) 处的像素的灰度值。 δ 表示图像的平均灰度值, 其表达式为 Eq.9

$$\delta = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N f(i, j) \quad (9)$$

1.2.6 亮度顺序误差 (Lightness Order Error, LOE)

亮度顺序误差 (LOE) 是图像亮度的顺序差异, 并且通过评估图像在邻域中的亮度的顺序变化过程来评估图像的照度变化。LOE 反映了图像的自然保持能力。较小的值表示图像具有更好的亮度顺序并且看起来更自然。表达式为 Eq.10

$$RD(i, j) = \sum_x^M \sum_y^N U(L(i, j), L(x, y)) \oplus U(L_e(i, j), L(x, y)) \quad (10)$$

式中, M 是图像的高度, N 是图像的宽度, 而 \oplus 是 XOR 算子, $L(i, j)$ 和 $L_e(x, y)$ 分别表示三个颜色通道中的最大值。LOE 的值越小, 图像保持的亮度顺序越好。

1.3 主观评价指标

1.3.1 差分平均意见分数 (Differential Mean Opinion Score, MOS)

差分平均意见得分 (MOS) 或差分平均看法得分 (DMOS)。其中, MOS 是应用最广泛的主观 IQA 方法。不同的人在原始图像和增强图像之间进行主观比较, 以获得 MOS 分数, 最后获得平均分数。MOS 评分范围从 1 到 5, 评分越高表示主观增强越好。MOS 评分公式为 Eq.??

$$MOS = \frac{\sum_{n=1}^N R_n}{N} \quad (11)$$

其中 R 表示评价者对图像的满意度得分的数量, N 表示评价者的总数。

2 个人工作进展

2.1 思考

具有融合和模型框架的算法往往具有更好的泛化能力, 并且无监督学习方法比有监督学习方法更稳健。

低照度图像增强算法的目的是为更高级别的视觉任务做准备。因此，当前的研究旨在制定通用的低照度图像增强算法，以服务于更高级别的视觉任务，所以低照度图像增强算法的速度应该尽可能的快，这样可以使得应用在高帧率的低照度视频增强上成为一种可能。

无监督学习方法上述方法有两个限制。首先，数据集中的成对图片是有限的。其次，在成对数据集上的训练模型会导致过拟合问题。

3 下周工作计划

(1) 无监督学习方法上述方法有两个限制。首先，数据集中的成对图片是有限的。其次，在成对数据集上的训练模型会导致过拟合问题。目前已有的解决方案是采用启发性学习方法，另一种是采用深度 GAN，结合注意力机制捕获全局/局部的特征。

References

- [1] Liang Dai, Wenchao Chen, Yanwei Liu, Antonios Argyriou, Chang Liu, Tao Lin, Penghui Wang, Zhen Xu, and Bo Chen. Switching gaussian mixture variational rnn for anomaly detection of diverse cdn websites. In *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, pages 300–309. IEEE, 2022.