# 无参考图像空间质量评价

## Anish Mittal, Anush K. Moorthy and Alan C. Bovik

## Department of Electrical and Computer Engineering University of Texas at Austin Austin, Texas 78712

### 摘要：

我们提出了一种基于自然场景统计的无参考图像空间质量评估器(BRISQUE)，它提取局部归一化亮度信号的逐点统计信息，并基于与自然图像模型的偏差来测量图像的自然度（或缺乏自然度）。我们还建立了邻域成对统计量的分布模型，提供归一化亮度信号在失真方向上的信息。虽然具有多尺度，但该模型易于计算的特点使它的计算快速并且省时。这一框架结构表现出统计性能，优于其他提出的无参考算法和全参考结构相似性指数(SSIM)。

### 1. 介绍

个人数字助理(PDA)，手持显示器和能够上网的智能手机带来了通讯和网络的革命技术。这为服务供应商带来了新的挑战，即在有限带宽的情况下优化诸如文本，图像和视频之类的结构化多媒体数据相关的网络。传统的交付方式主要集中在优化吞吐量、缓冲区长度和容量上。虽然这些方法表现出了出色的性能，我们可以预期，考虑到多媒体数据的性质和对最终消费者行为的建模，可能会带来明显的性能提高。这就提出了使用人类感知模型进行服务质量优化的想法。基于这种应用思维，我们提出了一种基于模型的方法来预测失真图像的视觉质量。这是一种无参考方法，不需要任何参考图像来比较失真图像的质量。虽然相关参考图像的辅助信息确实有用并使问题更简单，但通过网络发送该信息会导致额外的带宽开销，并且在实际场景中可能很难检索这些信息。另外,关于原始参考是否存在的假设也是值得怀疑的，因为在拍摄高质量照片的过程中所有捕获的图像通常都受到失真的影响。

在无参考质量评估领域的初步工作是通过建模模糊、噪声或压缩伪影来解决图像中的模糊、颗粒、块状和振铃现象[4、10、19、13、3、18]。然而，建立自然图像统计模型（或缺乏自然图像统计模型）的双重方法是非常有意义的。这样可以避免对失真进行详细建模，使得到的方法独立于图像受到的失真。最近开发了两种基于DCT统计（BLINDS）-II[12]和基于失真识别的图像真实性和完整性评估（DIIVINE）的无图像标记算法，对失真介质图像的类型没有任何假设。然而，这些建议的算法存在一些相关的缺陷，这可能使他们对某些应用程序无效。DIIVINE[9]很贵，因为它涉及到一个相对大量特性的计算。BLIINDS-II[12]计算了离散余弦变换系数的块统计数据，这也很耗时。离散余弦变换系数也很耗时间。DIIVINE和BLINDS-II都提供顶级的NRQA性能(到目前为止)，但值得考虑的是是否可以开发出纯空间的无需局部转换的NR QA算法，以提高性能和计算效率。

本文介绍了一种利用空间域归一化亮度系数以及相邻归一化亮度系数的对积来提供失真方向信息的模型。然后演示了如何参数化各自的分布。这些系数可以作为与人类对图像质量判断相关的统计特征。这些特征用于失真识别和失真度特定质量评价，两者相结合时产生与失真无关的图像质量度量。

### 2. 自然场景统计模型的空间域

我们的方法是建立在自然场景中观察到的统计规律模型的基础之上。“自然”的场景包含使用光学相机拍摄的任何图像无需人工处理。因此，自然图像不一定是诸如树木或天空的户外环境的图像。心理物理学家在空间域[11]和小波中域[17]中分析了自然场景统计(NSS)。在目前的发展中，我们专注于空间域的统计数据。Ruderman观察到对数对比值的局部均值[11]减法和方差归一化产生的去相关系数服从高斯分布。这个过程类似于中心环绕机制，随后是自适应增益控制。给定一个亮度图像I，我们计算:

(1)

式中，为空间指标，M、N分别为图像的高度和宽度，C=1当分母趋于0时用来防止不稳定发生的常数。例如在图像块对应于平面坐标的情况下

(2)

(3)

式中ω={ωk, l | k= -K,…,K， l= -L,…,L }是一个循环对称高斯滤波器采样到3个标准差。在我们的实现中，K = L = 3。在这些值在2到4时性能上变化没有多大影响。在我们的算法中，这种归一化亮度系数(公式1)称为均值减去对比度归一化(MSCN)系数，并用于所有的进一步处理。

我们的模型基于这样一个前提，即MSCN系数具有特征统计特性，在存在畸变时会受到干扰[7]，并且这些系数建模使得可以预测失真的类型和感知严重性[9]。图1说明了MSCN系数分布随失真类型而变化。一个重要的观察结果是原始图像如何呈现如Ruderman[11]所示的高斯统计特性。相比之下，每个失真都以独特的方式修改自然场景统计。例如，与白噪声引起的颗粒状伪影相比，JPEG产生的阻塞和模糊伪影引起类似拉普拉斯统计量的统计量，从而增加了分布的方差。

为了量化自然图像的统计量以及随失真程度的变化，我们使用AGGD(不对称广义高斯模型)的分布来拟合原始图像和失真图像的MSCN统计量。对每一幅图像，我们估计2个参数(α,(σl2+σr2)/2)由MSCN系数的AGGD拟合，形成我们的第一组特征。

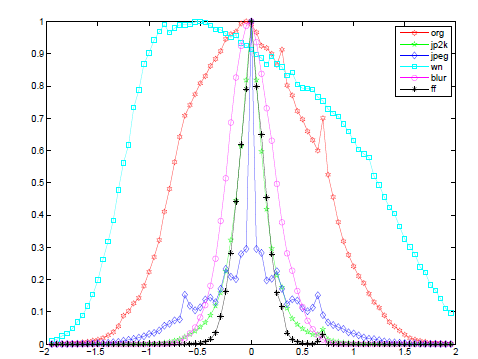


图1所示。自然未失真图像和各种失真版本的MSCN系数柱状图。来自实时IQA数据库[15]- JPEG2000 (Jp2K)和JPEG的失真压缩，加性高斯白噪声(WN)，高斯模糊(blur)，和瑞利快速衰落信道仿真(FF)。

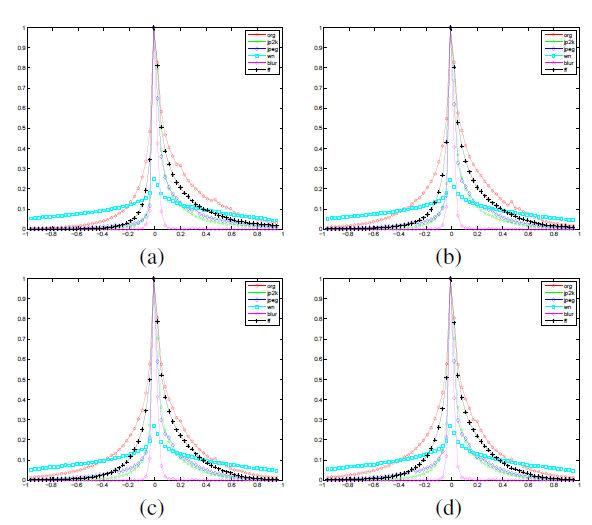
虽然MSCN系数明显更均匀，但系数的符号呈现出规整的结构[16]。为了量化邻域之间的依赖关系，我们通过相邻MSCN系数之间的成对乘积，沿四个方向：水平（hz）、垂直（v t）、主对角线（dg1）和次对角线（dg2），在距离为1像素的情况下建立方向关系模型。

(4)

(5)

(6)

(7)



**图2所示。MSCN系数对积的直方图对于自然不失真的图像及其各种失真版本。(a)水平，(b)垂直，(c)主对角线，(d)次对角线。来自实时IQA的扭曲数据库[15]- JPEG2000 (Jp2k)和JPEG压缩，添加剂白高斯噪声(WN)，高斯模糊(blur)，和瑞利快衰落信道仿真(FF)。**

图2显示了在存在参考图像的失真和每个失真版本的情况下，沿四个方向中的每个方向的配对产物是如何变化的。

配对后的分布也使用AGGD参数化，AGGD有效地捕获其分布的不对称性，如图2所示。因此,除了AGGD的参数(α,σl2，σr2)，我们也计算它配对后分布的平均值。对于AGGD，其均值为:

(8)

因此，对于每个成对的乘积，计算(α,σl2，σr2),这16个参数(4个参数/方向\*4个方向)形成另一组特征。由于图像固有的多尺度特性和失真对图像结构的跨尺度影响，我们计算了多个尺度上的BRISQUE特征。此外，研究人员已经证明，多尺度QA算法在与人类感知的相关性方面具有良好的性能 [21,12]。我们观察将量表数量增加到2个以上并没有改善性能。因此，我们总共有36个特性——每个尺度18个。

### 3.非对称广义高斯模型

零模式的AGGD定义为：

(9)

其中***α***是形状参数， and 是尺度参数

(10)

(11)

***γ(.)***是***γ***函数：

(12)

α控制分布的“形状”， 和分别确定模式左右偏差。当左右方差相等时，AGGD（非对称广义高斯分布）包含广义高斯分布(GGD)。研究人员利用这种不对称分布对偏态和重尾态分布进行了建模，用于纹理分析。采用[6]中提出的基于矩量的方法估计AGGD的 和

### 4. 两级质量评估框架

一旦我们有了这些特性，一个类似于DIVIINE的2阶段框架来进行质量推断。首先，计算的特征用于训练概率支持向量分类（SVC）以找出图像中每个失真的出现概率，然后利用训练后的支持向量回归(SVR)函数对每一种失真进行特定失真图像质量的评价。一旦我们从每个回归量得到质量，我们计算失真概率向量和失真质量向量的点积。我们使用LIBSVM包[2]来实现SVM和SVR以及用于分类和回归的径向偏置函数(RBF)内核。

这个框架背后的动机是每一个失真以不同的方式影响自然图象。因此,一组不同的特征对每个失真都很重要。我们使用通用功能集来表示图像的“自然性”。虽然这些特性不会模拟任何特定的伪象，例如模糊、振铃或阻塞，但它们仍然可以捕获各种失真的特点，从而可以训练分类器，该分类器可以将模型参数的偏差映射到图像受到影响的失真类型。

### 5. 效果评估

质量评估算法的性能通过与人类感知的相关性来衡量，其中人类对视觉质量的看法通常是来自大规模人类研究，使用由人类观察者评定的大量失真(和可能参考)信号。平均意见得分(MOS)或差异平均意见分数(DMOS)是通过对人体中的每个视觉信号进行平均得到的。得到的MOS / DMOS代表视觉信号的感知质量。质量评估（QA）算法的目标是模拟这些信号的质量得分，以便它们与人类对质量的看法（MOS / DMOS）相关联。

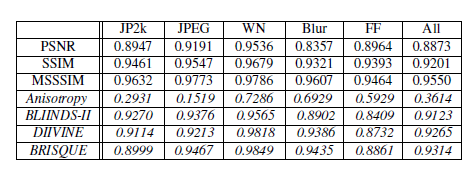
#### 5.1 与人类知觉的关系

我们使用LIVE IQA数据库[15]来测试算法的性能。它由29幅参考图像和779张失真图像组成，涵盖了五种不同的失真类别—— JPEG和jpeg2000压缩、白噪声、高斯模糊和瑞利快速衰落信道失真。其中每个失真图像都有一个相关的平均意见差值(DMOS)。训练分类器来预测失真的发生，并学习对应于各失真的回归函数。质量是使用前面解释过的信息来计算的。

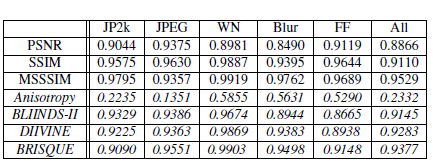
我们将数据集分为80%的训练和20%的测试，以确保训练内容和测试内容之间是没有重叠的。为了确保报告的结果不依赖于用来训练框架的空间内容选择，我们随机选取80%的空间内容进行训练，剩下的20%进行测试，并且在LIVE数据库上重复这个标准交叉执行1000次。最终我们报告在迭代中得分的中位数。

所采用的性能指标是斯皮尔曼秩序相关系数(SROCC)与皮尔逊线性相关系数(PLCC)在预测算法分数和DMOS之间的关系。在计算PLCC（皮尔逊线性相关系数）之前，使用逻辑非线性分数将算法分数映射到DMOS，如[15]中描述。SROCC和PLCC的值接近1表示与人类意见的相关性方面表现良好。这些性能指标分别列于表1和表2。

我们列出了全参考指标的表现:峰值信噪比(PSNR)和结构相似性指数(SSIM)[20]和几种无参考算法:基于各向异性的NR IQA[5]，基于DCT统计(BLIINDS) - II[12]的无图像标记器以及基于失真识别的图像真实感和完整性评价(DIIVINE)指数的性能进行了表列，进行比较。



**表1。在LIVE IQA数据库中，1000个试验组合的斯皮尔曼秩有序相关系数中位数(SROCC)。斜体表示无参考算法。**

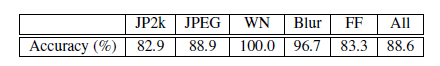


**表2。在LIVE IQA数据库中，1000个试验组合的皮尔逊线性相关系数中值(PLCC)。斜体表示无参考算法。**

从表1和表2可以看出，BRISQUE在与人类感知的相关性方面表现的相当不错，超过了目前全参考和无参考图像质量评价指标。

#### 5.2 分类精度

虽然使用了一个概率框架来进行失真识别，但是我们也展示了我们的特性对于显式失真识别时如何有用的。在表3中，我们报告了分类器对LIVE数据库中每种失真以及所有失真的分类精度中值。



**表3.1000次训练测试的分类准确率中值**

### 6. 结论

我们提出了一种基于自然场景统计的快速、简单的空间域无参考图像质量模型，评估证明，它不仅优于现有的NR IQA方法，而且它的性能也比全参考峰值信号表现要好——噪声比(PSNR)与结构相似度指标(SSIM)。除了这里考虑到的失真外，BRISQUE可以很容易地扩展，使其适用于通用目的无IQA问题。此外，BRISQUE的计算十分有效，这使其成为一个在实际应用中有吸引力的选择，例如图像恢复和质量网络评估。

### 7. 参考文献

**[1]** A.C. Bovik. Perceptual image processing: Seeing the future. IEEE, 98(11):1799–1803, 2010.

**[2]** C.C. Chang and C.J. Lin. LIBSVM: A library for support vector machines. 2001.

**[3]** J. Chen, Y. Zhang, L. Liang, S. Ma, R. Wang, and W. Gao. A no-reference blocking artifacts metric using selective gradient and plainness measures. Proceedings of Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing, pages 894–897,2008.

**[4]** R. Ferzli and L.J. Karam. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). IEEE Transactions on Image Processing, 18(4):717–728, 2009.

**[5]** S. Gabarda and G. Crist´obal. Blind image quality assessment through anisotropy. Journal of Optical Society of America, 24(12):42–51, 2007.

**[6]** N.E. Lasmar, Y. Stitou, and Y. Berthoumieu. Multiscale skewed heavy tailed model for texture analysis. In International Conference on Image Processing, pages 2281–2284. IEEE, 2009.

**[7]** A.K. Moorthy and A.C. Bovik. Statistics of natural image distortions. In International Conference on Acoustics Speech and Signal Processing, pages 962–965.

**[8]** A.K. Moorthy and A.C. Bovik. A two-stage framework for blind image quality assessment. In International Conference on Image Processing, pages 2481–2484, 2010.

**[9]** A.K. Moorthy and A.C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Transactions on Image Processing, 2011.

**[10]** N.D. Narvekar and L.J. Karam. A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection. In International Workshop on Quality of Multimedia Experience, pages 87–91, 2009.

**[11]** D.L. Ruderman. The statistics of natural images. Network computation in neural systems, 5(4):517–548,1994.

**[12]** M. Saad, A.C. Bovik, and C. Charrier. Model-Based Blind Image Quality Assessment Using Natural DCT Statistics . IEEE Transactions on Image Processing,2011.

**[13]** N.G. Sadaka, L.J. Karam, R. Ferzli, and G.P. Abousleman. A no-reference perceptual image sharpness metric based on saliency-weighted foveal pooling. In International Conference on Image Processing, pages 369–372,2008.

**[14]** K. Sharifi and A. Leon-Garcia. Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. IEEE Transactions on Circuits and Systems for Video Technology, 5(1):52–56, 1995.

**[15]** H.R. Sheikh, M.F. Sabir, and A.C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Transactions on Image Processing, 15(11):3440–3451, 2006.

**[16]** E P Simoncelli. Statistical modeling of photographic images. In Alan Bovik, editor, Handbook of Image and Video Processing, chapter 4.7, pages 431–441. Academic Press, May 2005. 2nd edition.

**[17]** A. Srivastava, A.B. Lee, E.P. Simoncelli, and S.C. Zhu. On advances in statistical modeling of natural images. Journal of mathematical imaging and vision, 18(1):17–33, 2003.

**[18]** S. Suthaharan. No-reference visually significant blocking artifact metric for natural scene images. Journal of Signal Processing, 89(8):1647 – 1652, 2009.

**[19]** S. Varadarajan and L.J. Karam. An improved perception-based no-reference objective image sharp sharpness metric using iterative edge refinement. In International Conference on Image Processing, pages 401–404,2008.

**[20]** Z. Wang, A.C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004.

**[21]** Z. Wang, E. P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In Asilomar Conference on Signals, Systems and Computers, volume 2, pages 1398–1402, 2003.

**[22]** 刘书琴.图像质量评价算法研究及应用[D].北京工业大学,2012.

**[23]** 张淑芳,张 聪,张 涛等. 通用型无参考图像质量评价算法综述[J]. 《计算机工程与应用》唯一官方网站, 2015, 51(19): 13-23.

**[24]** 王志明. 无参考图像质量评价综述. 自动化学报, 2015, 41(6): 1062-1079.

**[25]** 楼斌, 沈海斌, 赵武锋, 等. 基于自然图像统计的无参考图像质量评价[J]. J4, 2010, 44(2): 248-252.

**[26]** 杨璐,王辉,魏敏.基于机器学习的无参考图像质量评价综述[J].计算机工程与应用,2018,54(19):34-42.

**[27]** 陈勇,帅锋,樊强.基于自然统计特征分布的无参考图像质量评价[J].电子与信息学报,2016,(7).

**[28]** 王啸晨.基于机器学习的图像质量评价研究[D].天津大学,2015.