

DOI: 10.11992/tis.201805027

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180716.1307.012.html>

一种特征字典映射的图像盲评价方法研究

王伟^{1,2}, 刘辉², 杨俊安²

(1. 军事科学院 评估论证研究中心, 北京 100091; 2. 国防科技大学 电子对抗学院, 安徽 合肥 230037)

摘要: 图像质量评价对于许多计算机视觉任务来说, 是至关重要的一环。传统的方法往往聚焦于人类直观打分, 其最大不足就是打分数据的庞大性。为了解决这个难题, 本文提出了一种图像质量盲评价框架。首先分别提取图像的局部结构特征和全局统计特性, 在学习阶段, 提出了一种基于字典池的映射策略来加速打分的进程。实验结果显示, 本文所提方法准确度和鲁棒性相比较时下其他算法, 取得了更加令人满意的结果。

关键词: 客观评价; 盲评价; 图像质量评价; 局部结构特征; 全局统计特性; 特征提取; 字典; 池化映射

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2018)06-0989-05

中文引用格式: 王伟, 刘辉, 杨俊安. 一种特征字典映射的图像盲评价方法研究[J]. 智能系统学报, 2018, 13(6): 989-993.

英文引用格式: WANG Wei, LIU Hui, YANG Jun'an. Blind quality evaluation with image features codebook mapping[J]. CAAI transactions on intelligent systems, 2018, 13(6): 989-993.

Blind quality evaluation with image features codebook mapping

WANG Wei^{1,2}, LIU Hui², YANG Jun'an²

(1. Center for Assessment and Demonstration Research, Academy of Military Science, Beijing 100091, China; 2. Institute of Electronic Warfare, National University of Defense Technology, Hefei 230037, China)

Abstract: Image quality assessment is crucial to many computer vision tasks. Traditional approaches concentrate on human perceptual scoring. The biggest hurdle to these subjective efforts is the difficulty of collecting the enormous human scored data. To solve this difficulty, we propose a blind image quality assessment framework. Starting with local structural characteristics and global statistics characteristics of images, we utilize a codebook-based pooling strategy to accelerate the scoring stage. Experimental results show that by comparison with other algorithm, an effective performance in accuracy and robustness was achieved using the proposed approach.

Keywords: objective assessment; blind assessment; image quality assessment; local structural feature; global statistics characteristics; feature extraction; codebook; pooling mapping

数字视频和图像缩小了人类和自然界的差距。然而不幸的是, 由于各种各样的图像退化和畸变, 有用的信息可能会被丢失, 因此如何评价失真图像的质量成为了一种紧迫需要。理论上, 人类视觉系统 (HVS) 是最有效和直接的方式, 但是其收集大量人类打分数据的过程极其复杂、耗时。

关于图像质量评价的研究已经有较长时间。通常图像质量可以通过直接计算原始图像和退化图像之间的距离来衡量。然而在大多数情况下,

我们往往只有失真后的图像, 并没有原始图像做参考。这种问题的出现影响了图像质量评价的进程, 进而可能会给后续分析带来麻烦, 因此怎样设计出一个合适的盲图像质量评价系统是人们迫切的愿望。

1 相关工作

建立一个典型盲图像质量评价系统通常需要两个成分: 一系列特征以及学习型的回归模型。典型全参考图像质量评价过程流程图如图 1 所示。之前的方法通常将特征提取和模型训练分开考虑。对于这些特定类型的失真图像, 相应的特

收稿日期: 2018-05-22. 网络出版日期: 2018-07-17.

基金项目: 中国博士后科学基金项目 (2015M572722); 安徽省自然科学基金项目 (1408085MKL46).

通信作者: 王伟. E-mail: wwei009@mail.ustc.edu.cn.

征被用来衡量它们的质量,比如空间域的边缘宽度^[1]以及在某些转换域的峭度^[2]。然而当退化种类未知时,上述论文中所提的特定特征将无法使用。

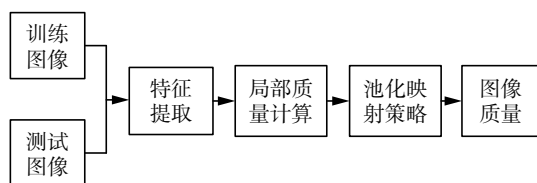


图1 典型全参考图像质量评价过程流程图
Fig. 1 The flowchart of typical FR_IQA process

还有一些其他方法没有使用人工手工选取的局部描绘算子,而是基于特征学习。监督式滤波器学习已被 Jain 和 Karu 在文献[3]中采用作为纹理分类,其中特征提取和分类任务被神经网络执行。文献[4]采用了无监督特征学习,其中局部算子被编代码然后用一种无监督的方式来学习。为了学习一个更加简洁紧凑的和有区分性的字典,学习类的方法优化包括采用一个优化过的具有重建性和区分性的标准^[5-6]。

2 本文方法

相较这些传统方法,一个具有区分性的特征集和灵活的学习策略可以提高表现。为了在图像质量预测阶段改善联合统计特征和学习策略,介绍了两种不同的特征提取成分:1)一个局部特征提取器;2)一个汇总了局部特征分布的全局特征提取器。

为了不失一般性,采用了一个基于字典类译码本的方法来规避标准信息的优势。传统的字典学习方法聚焦于信号的重构,同时需要字典中的学习原子,应该能很好地代表图像块,本文所提方法则没有这个限制。事实上,后续显示用于图像质量分类的字典将有很大不同。

2.1 图像特征提取

特征提取的第1步是从增强的图像块中捕捉合适的特征。在这个部分中,讨论怎么使用一系列线性滤波器来获取局部特征。受人类视觉系统发现和感知物体的启发,尝试设计出一个自然的、面向对象的检测器,在不同种类间具有一般性。首先,整幅图像被分割成互相重叠的图像块,然后从这些区域中提取两种类型的特征:一类是局部特征,另一类则是统计整幅图像局部特征分布情况的全局特征。

2.1.1 局部特征

为了方便,将原始图像块归一化成一系列局部描述算子。本节将提取以下算子来代表整幅图像的局部特征。

1) 显著性地图

图像的亮度变化传达着图像自身许多有用的信息。带通图像响应,尤其是高斯滤波器响应,可以用作描绘多种图像的语义结构,比如直线、边缘、角点以及斑点等,这些都与人类主观认知密切相关。有了这个发现,我们开发了一种语义显著性准则来感知图像中的目标^[7]:

$$S(x, y) = \|I_{\mu} - I_{\omega_{hc}}(x, y)\| \quad (1)$$

式中: I_{μ} 是图像特征向量的均值, $I_{\omega_{hc}}(x, y)$ 是原始图像对应的高斯模糊图像像素值(使用一个 5×5 可分离的二项式内核), $\|\cdot\|$ 是 L_2 范数。

2) 梯度幅值图

梯度算子通常使用卷积模板来表达,本文实现中,梯度幅值的定义为

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

式中 G_x 和 G_y 分别是图像相应像素位置的水平和垂直梯度。自然界梯度统计分布在图像分析中扮演着关键角色,梯度相似性可被用作建立图像语义结构的基本要素,这些语义结构与人类可感知的自然图像质量密切相关。

2.1.2 全局统计特性

为了克服传统像素级评价方法的不足,我们更加关注局部特征细节和全局统计特征之间的平衡。

统计作为一种全局描述算子,是对所提取局部特征分布的概述。特别指出的是,我们使用了一种 BRISQUE^[8] 参数模型,模型对像素值进行了 GGD 归一化,其中形状和尺度参数都被用作特征。虽然滤波响应的峭度和峰度值可能不能正确描述分布的形状,但对于质量高低不等的图像,它们是相对好的指示器,如图2所示。

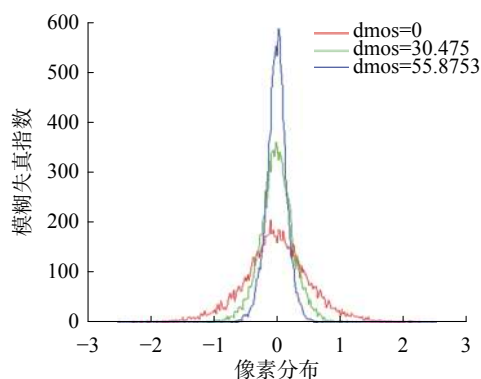


图2 滤波响应后不同等级模糊失真示例(高 DMOS 值表示低质量无量纲)

Fig. 2 Examples of filter responses for different levels of blur distortion (high DMOS indicates low quality)

此外峭度值 S 和峰度值 K 的计算非常方便,定义为

$$K = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^4 / SD^4 - 3$$

$$S = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^3 / SD^3 \quad (3)$$

式中: \bar{x} 和 SD 分别代表着整幅图像像素灰度值的均值和方差, n 为像素数目。

2.2 图像分类打分

本节讨论这些局部特征和全局描述算子是如何通过来学习准则预测不同图像的质量, 轻松解决图像质量盲评价的病态任务问题。

2.2.1 学习阶段

为了方便训练一个值得可信的图像质量盲评价模型通常需要一个很大的人类打分训练集, 通过主观测试获得这样一个模型的代价显得昂贵耗时。因此在训练阶段开发一个不需要如此大数据量的人类主观打分的图像质量盲评价模型变得迫不及待。本节求助于聚类学习和池化映射策略。

1) 池化策略

如果手里没有任何人类主观质量打分, 我们瞄准学习图像质量盲评价的中心集。为了方便起见, 退化图像和原始图像都被分割成重叠块。

所提方法的关键问题是怎样分配感知质量。本文我们求助于特征相似性指数^[9](FSIM)来计算参考图像和变形图像之间的相似度。通过这种方式, 可以去除对人类主观打分的依赖。退化图像的FSIM得分定义如下:

$$s_i = S(r_i, d_i) = \frac{2P(r_i)P(d_i) + \alpha}{P(r_i)^2 + P(d_i)^2 + \alpha} \times \frac{2G(r_i)G(d_i) + \beta}{G(r_i)^2 + G(d_i)^2 + \beta} \quad (4)$$

式中: d_i 和 r_i 代表退化图像和参考图像, $P(r_i)$ 和 $G(r_i)$ 分别代表着在中心点 r_i 的相位一致性和梯度幅值。为了避免0除计算, α 和 β 是正值常量。

相似度得分 s_i 在某种意义上反映了退化图像 d_i 的质量。为了使得这些相似度得分的均值接近真实人类主观打分质量, 我们将 s_i 归一化, 提出一种百分数的池化策略来完成归一化过程。

假设全部图像块的集合为 Ω , Φ_p 表示其中质量最低的20%局部块集合。归一化因子的计算为

$$W = \frac{\sum_{i \in \Phi} s_i}{c \sum_{i \in \Phi_p} s_i} \quad (5)$$

式中: $c=20$, 每个退化质量块的最终得分归一化为

$$\omega_i = s_i / W \quad (6)$$

因此整幅图像的质量可表达为平均质量, 这就与百分池化结果保持一致了。

2) 聚类

通过块质量归一化策略, 将相似度得分按分值分成多个组, 再把那些降质块按各自的局部构造分为不同类。

归一化后的块得分 ω_i 是个介于0~1之间的正值, 首先统一将 ω_i 量化成 L 个等级, 然后将具有相近质量等级的图像块归合成同一组, 记作 Q_l 。因此有:

$$Q_l = \begin{cases} \{d_i | q_{l-1} < \omega_i < q_l, & l = 2, 3, \dots, L \\ \{d_i | \omega_i < q_l, & l = 1 \end{cases} \quad (7)$$

对于每组 Q_l , 应用聚类过程来获得质量。为了提高聚类准确度, 每个质量层应该基于相同的结构特征。在预处理阶段, 使用高斯高通滤波器来提取块特征, 藉此提高图像的细节结构。具体实现中, 使用的是高斯金字塔的方案在不同尺度提取图像细节特征, 并形成一特征向量。QAC通过对滤波特征 f_i 应用K均值聚类算法, 得到 Q_l 组类心 $m_{l,k}$:

$$\min_{m_{l,k}} \sum_{k=1}^K \sum_{d \in Q_{l,k}} \|f_i - m_{l,k}\|^2 \quad (8)$$

式中: $Q_{l,k}$ 是 Q_l 组中第 k 个类。值得注意的是, 也可以使用其他相似度准则。然而考虑到计算复杂度, 为了方便使用欧式距离, 为了有效求解上述最小化问题, 应用了谱聚类^[10]的方法。

2.2.2 测试阶段

有了这些学习得到的不同质量层的中心集, 可以通过把这些退化图像特征映射到相应的质量得分池, 轻易地推断出感知质量, 如图3所示。

通过应用滤波算法, 可以获得新的增强图像及其相应的质量等级。

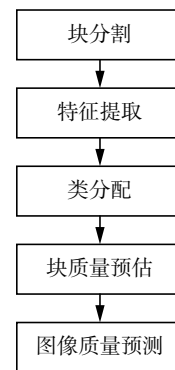


图3 图像质量估计流程图

Fig. 3 The process of image quality prediction

通过式(9)所示的均值策略来推断新测试图像最终的质量得分。

$$Q(A) = \frac{1}{n} \sum_{i=1}^n q_i \quad (9)$$

式中: n 是所有图像块的数量, q_i 是每个图像块的质量, $Q(A)$ 是待估图像 A 的最终质量得分。

3 实验

3.1 实验设置

在通用基准数据集 LIVE 数据库上评估了我们所提算法。LIVE 数据库包含了 29 幅参考图像及其对应的 779 幅 5 种不同类型的退化图像。为了代表性和简洁性考虑, 仅仅选择其中两个典型的退化类型: 白噪声 (WN)、高斯模糊 (BLUR)。对于每个退化图像, 都有一个主观感知得分 DMOS 值, 分值范围从 0 ~ 100 不等。较低的 DMOS 得分代表着较好的视觉质量。实验中, 随机选取 80% 的参考图像及其退化图像作为训练集样本, 剩余 20% 用作测试集。

3.2 质量预测

为了验证本文所提方法的有效性, 选择 LIVE 数据库中一个数据集来分别预测它们的质量得分, 如图 4 所示。

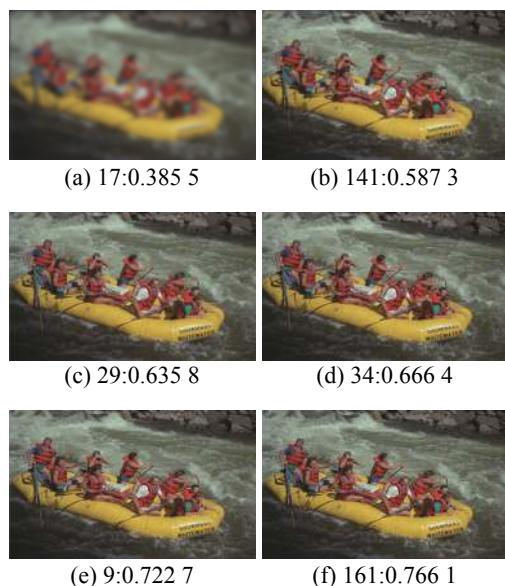


图 4 LIVE 数据库中同一幅基准图像不同退化程度后的图像质量预测

Fig. 4 Quality prediction of different degraded degree image from LIVE database using proposed method

从结果可以看出, 使用本文算法预测出的分数和人类视觉系统的结果完全一致。

3.3 对比评估

用植入线性关联系数 (LCC) 准则来评估本文所提算法的表现。LCC 通过衡量真实分数和预测分数之间的线性关系来预测分数。假如有 n 幅退化图像, 每幅图像有一个人类感知分数等级和一个预测的分数等级。LCC 的计算为

$$LCC = \frac{\sum_{i=1}^n d_i}{\sqrt{\sum_{i=1}^n d_i^2}} \quad (10)$$

式中 $d_i = v_i - p_i$ 是二者之间的差距。

作为对比, 选择集中具有代表性的 NR-IQA 方法来评价 LIVE 数据库中特定退化类型和非特定退化类型的实验。对于前者, 随机选择特定类型的退化图像来训练和测试, 对于后者所有退化类型的图像将放在一起训练和测试。

前人几种表现良好的无参考图像质量评价和全参考图像质量评价方法用来作为对比: 如表 I 所列, BRISQUE^[8]、CORNIA^[4]和 CNN^[11]都是从原文中提取的方法。除了 CNN, 所有这些方法在作对比时, 取 LIVE 数据库中的 80% 做训练, 剩余 20% 做测试。对于 CNN 取 60% 做训练, 20% 做更新, 剩余的 20% 做测试。

表 1 LIVE 数据库上的 LCC
Table 1 LCC on LIVE

LCC	WN	BLUR	FF	ALL
PSNR	0.926	0.779	0.870	0.856
PSIM	0.976	0.978	0.912	0.960
SSIM	0.982	0.893	0.939	0.906
BRISQUE	0.985	0.951	0.903	0.942
CORNIA	0.987	0.968	0.917	0.935
CNN	0.984	0.953	0.933	0.953
OURS	0.987	0.968	0.944	0.957

4 结束语

本文提出了一个简单有效的框架来自动评估图像质量。全文的创新性在于集成使用了局部特征和全局统计特性描绘图像。本文的贡献是两方面的: 首先, 本文用到了一个包括局部特征和全局统计特性有区分性的特征集来表征图像块; 其次, 研究了怎样在无标记的数据集上构建字典, 使用了一个跟退化类型无关的池化策略来加速学习过程。

相比前人的方法, 本文所提算法有了很大的提高, 实验表明本文算法能有效处理某种程度上不可预见退化类型的图像。未来将考虑将目标检测融合到我们的框架中去, 以此代替对所有检测区域质量的均值化处理。

参考文献:

- [1] WANG Wei, LIU Hui, ZHENG Jinjin, et al. Integrated blur image quality assessment based on human visual perception[C]//Proceedings of the International Conference on Computer Science and Applications. Wuhan, China, 2017: 119–124.
- [2] FERZLI R, KARAM L J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)[J]. IEEE transactions on image processing, 2009, 18(4): 717–728.
- [3] HASSEN R, WANG Zhou, SALAMA M. No-reference image sharpness assessment based on local phase coherence measurement[C]//Proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. Dallas, USA, 2010: 2434–2437.
- [4] CHEN Mingjun, BOVIK A C. No-reference image blur assessment using multiscale gradient[C]//Proceedings of 2009 International Workshop on Quality of Multimedia Experience. San Diego, USA, 2009: 3.
- [5] LI Chaofeng, BOVIK A C. Content-partitioned structural similarity index for image quality assessment[J]. Signal processing: image communication, 2010, 25(7): 517–526.
- [6] CIANCIO A, DA COSTA A L, DA SILVA E A, et al. No-reference blur assessment of digital pictures based on multifeature classifiers[J]. IEEE transactions on image processing, 2010, 20(1): 64–75.
- [7] YAO Hongxun, HUSEH M Y, YAO Guilin, et al. Image evaluation factors[C]//KAMEL M, CAMPILHO A. Image Analysis and Recognition. Berlin, Heidelberg: Springer, 2005: 255–262.
- [8] HASLER D, SÜSSTRUNK S E. Measuring colourfulness in natural images[J]. Electronic imaging, 2003, 5007: 87–95.
- [9] ZHANG Qiang, HAN Yu, CAI Yunze. Novel full-reference image quality assessment metric based on entropy fusion[J]. Optik-international journal for light and electron optics, 2013, 124(21): 5149–5153.
- [10] GU Ke, ZHAI Guangtao, YANG Xiaokang, et al. A new reduced-reference image quality assessment using structural degradation model[C]//Proceedings of 2013 IEEE International Symposium on Circuits and Systems. Beijing, China, 2013: 1095–1098.
- [11] ROTH S, BLACK M J. Fields of experts: a framework for learning image priors[C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA, 2005: 860–867.

作者简介:



王伟, 男, 1987 年生, 博士, 主要研究方向为计算机视觉、模式识别、智能信息处理。获得国家发明专利授权 1 项。发表学术论文 18 篇, 被 SCI 检索 3 篇, EI 检索 12 篇。



刘辉, 男, 1983 年生, 博士, 主要研究方向为智能信息处理、通信对抗技术。获得国防发明专利授权 1 项。发表学术论文 25 篇, 被 SCI 检索 4 篇, EI 检索 16 篇。



杨俊安, 男, 1965 年生, 教授, 博士生导师, 主要研究方向为机器学习、智能信息处理、通信对抗技术。获得国防发明专利授权 1 项。发表学术论文 70 余篇, 被 SCI 检索 8 篇, EI 检索 30 余篇。