

基于图像内容视觉感知的图像质量客观评价方法

姚军财 刘贵忠

Objective assessment method of image quality based on visual perception of image content

Yao Jun-Cai Liu Gui-Zhong

引用信息 Citation: *Acta Physica Sinica*, 67, 108702 (2018) DOI: 10.7498/aps.67.20180168

在线阅读 View online: <http://dx.doi.org/10.7498/aps.67.20180168>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2018/V67/I10>

您可能感兴趣的其他文章

Articles you may be interested in

[天光背景下混浊大气中成像质量的分析方法](#)

Image quality analysis method under background radiation in turbid atmosphere

物理学报.2018, 67(8): 088701 <http://dx.doi.org/10.7498/aps.67.20172625>

[基于主特征提取的 Retinex 多谱段图像增强](#)

Multispectral image enhancement based on Retinex by using structure extraction

物理学报.2016, 65(16): 160701 <http://dx.doi.org/10.7498/aps.65.160701>

[基于声透镜的三维光声成像技术](#)

Three-dimensional photoacoustic imaging technique based on acoustic lens

物理学报.2015, 64(24): 240701 <http://dx.doi.org/10.7498/aps.64.240701>

[基于随机聚类的复杂背景建模与前景检测算法](#)

Complex background model and foreground detection based on random aggregation

物理学报.2015, 64(15): 150701 <http://dx.doi.org/10.7498/aps.64.150701>

[基于照射_反射模型和有界运算的多谱段图像增强](#)

Multispectral image enhancement based on irradiation-reflection model and bounded operation

物理学报.2015, 64(10): 100701 <http://dx.doi.org/10.7498/aps.64.100701>

基于图像内容视觉感知的图像质量客观评价方法*

姚军财¹⁾²⁾ 刘贵忠^{1)†}

1) (西安交通大学电子与信息工程学院, 西安 710049)

2) (陕西理工大学物理与电信工程学院, 汉中 723000)

(2018年1月23日收到; 2018年3月7日收到修改稿)

图像质量客观评价在图像和视频传输、编解码以及服务质量中起着非常重要的作用. 然而现有的方法往往没有考虑图像内容特征及其视觉感知, 使得其质量客观评价与主观感知结果存在一定的差距. 基于此, 本文结合图像内容的复杂性特征和人眼的掩蔽特性、对比敏感度特性以及亮度感知的非线性特性, 提出了一种基于人眼对图像内容感知的图像质量客观评价方法. 该方法首先结合亮度感知的非线性模型将图像进行转换, 得到人眼感知强度图; 再分别以人眼对比敏感度值和图像局部平均对比度值作为权重因子对强度求和, 以求和的数据信息作为人眼感知图像的内容, 并构建图像感知模型; 最后以此模型分别模拟人眼感知参考图像和失真图像, 并计算二者的强度差, 以强度差为评价分数的基础构建图像质量客观评价模型. 采用 LIVE, TID2008 和 CSIQ 三个数据库中的共 47 幅参考图像和 1549 幅测试图像进行仿真实验, 且与 SSIM, VSNR, FSIM 和 PSNRHVS 等典型的图像质量客观评价模型进行对比分析, 同时探讨影响图像质量评价的因素. 结果表明: 所提方法的评价分数与主观评价分数的 Pearson 线性相关性系数和 Spearman 秩相关系数值比 SSIM 的评价结果均有一定程度的提高, 提高幅度分别平均为 9.5402% 和 3.2852%, 比 PSNRHVS 和 VSNR 提高幅度更大. 综合以上表明: 所提方法是一种有效可行的图像质量客观评价方法; 同时, 在图像质量客观评价中, 考虑人眼对图像内容的感知和复杂度的分析有助于提高图像质量主客观评价的一致性, 评价精度可得到进一步的提高.

关键词: 图像质量评价, 图像内容, 人眼视觉特性, 相关系数

PACS: 87.57.C-, 07.05.Pj, 42.30.-d

DOI: 10.7498/aps.67.20180168

1 引言

随着光电、信息和通信技术的快速发展, 图像技术的应用日益广泛. 但是, 在图像的获取、存储、传输和处理过程中均会发生不同程度的改变, 导致图像有一定的失真, 直接影响了观察者的主观感知效果和服务质量. 因此, 需要一种有效而又实用的度量失真图像质量的评价方法. 有效可靠的图像质量客观评价方法不仅能自动评价其优劣, 而且还能对图像进行质量监控, 指导参数设置和优化模块算法, 指导和控制网络传输等. 目前, 虽然图像质量评价方面的研究工作取得了较多的成果, 但还远远

不能满足实际应用的要求和需要, 主要是因为其客观质量评价与主观感知结果存在一定的差距. 因此, 研究尽可能逼近人眼主观感知、有效且方便的图像质量客观评价方法非常必要, 具有重要的理论研究和实际应用价值^[1-3].

近些年, 国内外的许多专家学者对图像质量客观评价方法做了大量的研究工作, 提出了较多的评价模型^[3-12], 其中代表性的有 MSE (mean square error), PSNR (peak signal-to-noise ratio), SSIM (structural similarity index)^[3], VSNR (visual signal-to-noise ratio)^[4], FSIM (feature similarity index)^[5], GMSD (gradient magnitude similarity deviation)^[6] 和 VSI (visual saliency-induced

* 国家自然科学基金 (批准号: 61301237) 和陕西省科技新星计划 (批准号: 2015KJXX-42) 资助的课题.

† 通信作者. E-mail: liugz@mail.xjtu.edu.cn

index)^[7]. 但分析其评价模型, 仍存在些许的不足^[5,12,13], 其主要表现如下.

1) 模型只是计算图像间的数值差异, 完全没有考虑人类感知因素, 致使图像主客观评价结果差异较大. 如PSNR和MSE等基于像素域的评价方法, 但该方法计算简单, 以致目前仍是最为普遍的评价方法之一.

2) 考虑了部分人眼视觉特性, 但主客观评价结果仍存在一定程度的差异. 国内外在此方面做了大量的研究, 得出了较多的成果^[3-10], 其中普遍认同的是Wang等^[3,4]提出的SSIM评价模型. 该模型是利用人眼对图像结构信息的感知特性, 从亮度、对比度和结构三个方面构建的评价模型. 综合考虑模型计算复杂度、评价精度和实时性等因素, 其评价效果优于SNR, PSNR, VSNR, FSIM, GMSD和VSI. 但该模型主要是基于人眼对图像中的结构信息可以自适应这一特征提出的, 没有考虑图像中的视觉冗余^[5], 即人眼对图像的失真存在一定的阈值, 当失真量低于阈值时, 人眼不能感知, 则主观评价分数基本不变; 但对于SSIM来说, 只要有失真, 其评价分数便有变化, 从而导致了主客观评价的不一致性.

3) 没有考虑图像本身内容的感知和复杂度的影响. 目前的评价方法主要是针对压缩和传输等外在因素影响而进行的评价, 如块效应、模糊和振铃效应等, 很少考虑图像内容本身的感知, 并且基本上没有考虑图像内容复杂度的影响, 致使主客观评价结果存在一定程度上的不一致^[10,13-15]. 目前, 这方面的研究已成为热点.

4) 评价模型比较复杂. 此问题是目前结合人眼视觉特性进行图像质量评价研究中的一个难点问题, 结合人眼视觉特性及其数学模型在很大程度上会增加评价模型的计算复杂度, 如VSNR, FSIM和VSI等. 一般评价方法是通过提取和分析图像中引起失真的特征因素、根据视觉特性和权重构建评价模型. 此类方法评价效果虽然较为显著, 但其针对性较强; 且为了提高评价准确度, 需要提取更多的有效特征, 使得计算更加复杂^[4,5].

基于以上的不足, 本文结合图像内容的复杂性特征和人眼对亮度感知的非线性、掩蔽特性以及对对比敏感度特性, 提出了一种基于人眼对图像内容感知的质量客观评价方法, 采用LIVE, TID2008和CSIQ三个数据库中的共47幅参考图像和1549幅测试图像进行仿真实验, 并与PSNR, SSIM和

VSNR等典型的图像质量客观评价模型进行了对比分析. 结果表明: 在图像质量客观评价中, 考虑人眼对图像内容和复杂度的感知有助于提高图像质量主客观评价的一致性, 评价精度可得到进一步的提高.

2 视觉模型及其在图像质量评价中的应用

2.1 对比敏感度视觉模型

用户体验是图像质量客观评价的四大关键因素之一. 目前的评价方法最主要不足之处是没有较好地结合人眼视觉特性, 导致主客观评价的不一致^[3-7]. 人眼对比敏感特性是4个重要的空间视觉特性之一, 国内外专家学者对其进行了大量的研究, 提出了较多的数学模型, 其模型一般称之为“人眼对比敏感度函数”(contrast sensitivity function, CSF), 该函数主要描述人眼对比敏感度值(一般采用人眼对比觉察阈值的倒数来定量描述)与角频率(f_θ)之间的函数关系^[16-18]. 综合考虑图像质量评价算法的复杂性、CSF模型反映感知特性的效果等方面, 分析前人的视觉模型. 本文选用由Stephen等^[16]提出的人眼亮度对比敏感度函数(CSF_L), 以及由Nadenau^[17]提出的人眼红绿(rg)和蓝黄(by)对立色对比敏感度函数(CSF_{rg} 和 CSF_{by}), 其表达如(1)–(3)式:

1) 人眼亮度对比敏感度函数

$$C_L(f_\theta) = a \cdot f_\theta \cdot \exp(-b \cdot f_\theta) [1 + c \cdot \exp(b \cdot f_\theta)]^{0.5}, \quad (1)$$

式中,

$$a = \frac{540 \cdot (1 + 0.7/L)^{-0.2}}{1 + \frac{12}{w \cdot (1 + f_\theta/3)^2}},$$

$$b = 0.3 \cdot (1 + 100/L)^{0.15}, \quad c = 0.06,$$

其中 C_L 表示人眼觉察亮度目标光栅的对比敏感度阈值; L 表示光栅的平均亮度; f_θ 为人眼观察光栅时的角频率, 单位为cpd (cycle per degree); w 表示觉察时空间视角对应的刺激瞳孔的大小^[16].

2) 人眼红绿和蓝黄对立色对比敏感度函数

$$C_{rg}(f_\theta) = a \cdot \exp b(f_\theta)^c$$

$$a = 1, \quad b = -0.152, \quad c = 0.893, \quad (2)$$

$$C_{by}(f_\theta) = a \cdot \exp b(f_\theta)^c$$

$$a = 1, b = -0.2041, c = 0.9, \quad (3)$$

式中 C_{rg} 和 C_{by} 分别是定量描述人眼对红绿和蓝黄对立色光栅的彩色对比敏感度阈值 [17,18].

2.2 视觉模型在图像评价中的应用

在结合人眼视觉特性的图像质量评价中, 最为关键的一点是构建视觉模型与图像数据之间的关系 [12,16], 而两者之间的联系即为频率. 在图像技术中, 图像经变换后得到频谱图, 其频率称之为空间频率. 而在 CSF 模型中的频率是角频率, 其定义为单位视角内刺激人眼的光栅条纹的周期数目, 表达式为 [18]

$$f_{\theta} = \frac{N}{\theta}, \quad (4)$$

式中 N 为视角范围内目标光栅条纹的周期数目, θ 为观察视角. 则要将视觉模型 CSF 应用到图像数据中, 需要构建两频率之间的关系. 具体的计算方法如下.

图像经离散余弦变换 (DCT) 后得到的频谱图上任意一点的频谱坐标 (f_x, f_y) 即为该点的空间频率, 且其任意一点经逆变换后得到空域图. 通过实验发现: 任意一点对应的空域图均是竖直或水平或两者共存的条纹光栅, 且光栅与原频谱图大小相同, 光栅的周期数目为 $(f_x^2 + f_y^2)^{1/2}$. 则可以把图像看成是由若干个光栅合并而成; 而人眼是最好的频谱滤波器, 则人眼观察频谱图像上的点时, 可以等效为观看其点对应的光栅. 基于此特点, 当人眼观察图像时, 人眼即是觉察周期数目为 $(f_x^2 + f_y^2)^{1/2}$ 的光栅, 观察视角为图像大小范围所对应的视角;

结合 (4) 式, 则人眼观察图像时, 图像上任意一点的角频率计算如 (5) 式, 则其较好地将两频率联系起来.

$$f_{\theta} = \frac{\sqrt{f_x^2 + f_y^2}}{\theta}. \quad (5)$$

根据以上关系, 当人眼观察图像时, 图像中任意一点的角频率都可通过 (5) 式计算获得; 再代入到 CSF 视觉模型 (1), (2) 和 (3) 式中, 则可以计算获得人眼觉察图像中任一像素点的对比敏感度阈值, 包括亮度和色度的阈值. 从而将 CSF 用在图像评价中.

3 图像质量客观评价模型构建方法

人眼接收到的图像信息主要是对图像的亮度和色度信息的感知, 则可以把人眼感知到的图像的亮度和色度作为人眼感知图像的初步信息. 而人眼感知图像每一点的亮度和色度信息受到两个方面的影响: 一是该点小范围内其他各点引起的亮度和色度对比对其的影响, 即局部对比度的影响; 二是人眼对目标本身的敏感程度的影响. 考虑二者的影响, 以人眼的敏感程度和局部平均对比度作为感知初步信息的权重来求和, 其信息之和作为人眼感知图像的信息内容 (即感知图像), 以此思想构建图像内容感知模型. 并以此模型求源图像和失真图像之间的强度差, 以该差作为图像质量客观评价的分数, 从而构建图像质量客观评价模型, 模型构建思路流程如图 1.

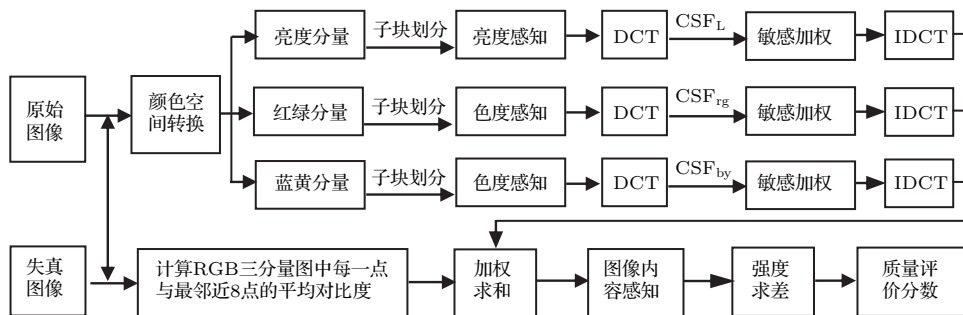


图 1 提出的图像质量客观评价方法流程图

Fig. 1. The architecture for proposed objective evaluation method of image quality.

1) 图像信息强度感知

图像是由亮度和色度混合而成, 但无论是亮度, 还是色度, 对外的表现均是其能量强度的呈现,

人眼对其的感知亦是对其强度的感知. 在目前的图像质量评价方法中, 表现图像差异的方法主要是用强度差来描述. 但人眼对能量强度的感知是按对数

形式变化,从而导致主客观评价的不一致. 基于此,采用人眼对光强度的感知模型((6)式),分别计算人眼对图像的亮度、色度的强度感知值 S , 以此作为人眼感知图像亮度和色度的信息.

$$S = \sum_{i=L,rg,by} K_i \log I_i, \quad (6)$$

式中 I 为彩色图像的亮度(L)、红绿(rg)和蓝黄(by)三个分量图中各点的能量强度值, K 为常数.

具体操作为: 将图像从RGB色空间转换到 $YCbCr$ 色空间描述, 进行亮度和色度分离, 得到亮度、红绿和蓝黄三个分量图; 并分别对其进行 $8 \text{ pixels} \times 8 \text{ pixels}$ 大小子块划分. 利用(6)式, 分别计算三个分量图中所有子块的每一点的亮度对应的主观感知强度值 S , 分别记为 S_L, S_{rg} 和 S_{by} .

2) 人眼对图像信息强度的敏感

同一目标在不同的空间位置或具有不同的强度, 给人眼的感知敏感程度是不同的. 研究表明, 在观察图像时, 最为引起人眼注意的范围是58个像素点的区域^[16,17]. 则在图像技术中, 为了方便处理, 一般将关注的范围定为 $8 \text{ pixels} \times 8 \text{ pixels}$ 大小子块. 结合此特性和人眼对比敏感度特性, 计算图像每一子块中每一像素点的敏感程度. 具体操作如下.

$$[R_{\text{subject}} G_{\text{subject}} B_{\text{subject}}] = T_{\text{YCC-RGB}} \{ T_{\text{refresh}} \{ \text{CAT}_{\text{block}} \{ \text{IDCT} [\text{DCT} (S_{ij}) \times C_i] \} \} \} \quad (i = L, rg, by; j = 1, \dots, M). \quad (7)$$

3) 局部环境对感知图像信息强度的影响

当人眼观看图像时, 人眼感知图像中的每一点都受到该点周围多个像素点的影响. 则为了考虑此影响, 对RGB图像分别以 R, G, B 三分量图中每一点自身为中心, 计算该点与周围最邻近的8个点之间的对比度值, 并求其平均值 C_{mean} . 以此值作为感知该点受到局部周围环境的影响因子.

4) 图像内容感知模型

$$\text{RGB}_{\text{subject}} = \text{CAT}_{\text{RGB}} [C_{m_R} \cdot R_{\text{subject}} \quad C_{m_G} \cdot G_{\text{subject}} \quad C_{m_B} \cdot B_{\text{subject}}], \quad (8)$$

式中 C_{m_R}, C_{m_G} 和 C_{m_B} 为 C_{mean} 对应于 R, G, B 分量图中的局部对比度平均值.

5) 图像质量客观评价模型

按照以上的图像内容感知模型, 分别处理, 计

① 对所有子块进行离散余弦变换, 并将每个子块频谱图的频谱进行平移, 使其零频移至频谱中心. 计算每个子块频谱图中每一点的空间频率 $f = (f_x^2 + f_y^2)^{1/2}$. 根据前面的说明, 则 f 为反离散余弦变换空域图上呈现光栅的条纹的周期数目, 即(4)式中的 N .

② 以 $8 \text{ pixels} \times 8 \text{ pixels}$ 大小子块为每次人眼关注的目标, 计算其空间视角 θ ; 根据①中得出的 N 值, 将 N 和 θ 代入(5)式, 计算所有子块中每一点的角频率 f_θ .

③ 将角频率 f_θ 代入到人眼对比敏感度视觉模型(即(1),(2)和(3)式)中, 分别计算出人眼对图像中每一点的敏感程度值 C_L, C_{rg} 和 C_{by} .

④ 分别对 S_L, S_{rg} 和 S_{by} 进行DCT变换, 以 C_L, C_{rg} 和 C_{by} 为权重因子, 对其变换域值加权, 再对其做反离散余弦变换(IDCT), 并将所有子块(数目为 M) 合成($\text{CAT}_{\text{block}}$), 并排序恢复整幅图像(T_{refresh}), 其结果为人眼主观感知图像亮度和色度的值; 并依据 $YCbCr$ 和RGB之间的经典转换($T_{\text{YCC-RGB}}$)关系式转换到RGB色空间中, 得到人眼主观感知图像的 R, G, B 分量图, 记为 $R_{\text{subject}}, G_{\text{subject}}, B_{\text{subject}}$, 如(7)式:

以 C_{mean} 为权重因子, 分别与主观感知的 $R_{\text{subject}}, G_{\text{subject}}, B_{\text{subject}}$ 三分量图中对应的各点幅值求积. 依据RGB色空间颜色混合法则, 将求积后的 R, G, B 值按比例关系混合(CAT_{RGB})得到RGB彩色图像, 其即为人眼主观感知的图像(表述为感知图像), 记为 $\text{RGB}_{\text{subject}}$. 计算如(8)式, 其即为人眼感知图像内容模型.

算获得原始图像的感知图像 $\text{RGB}_{\text{subject-original}}$ 和失真图像的感知图像 $\text{RGB}_{\text{subject-distortion}}$, 并按对应像素点计算其平均强度差. 基于峰值信噪比(PSNR)的定义式, 用平均强度差代替PSNR中的

均方误差 (MSE) 值计算, 计算结果则可作为图像质量客观评价分数. 其分数值越大, 表明图像质量越好, 反之则越差. 计算如 (9) 式, 其即为基于人眼对图像内容感知的图像质量客观评价模型, 记为 LCCM (image quality evaluation model based on luminance, contrast influence and contrast perception).

$$LCCM = 10 \times \lg \left[\frac{255 \times 255}{\text{abs}(\text{RGB}_{\text{subject-original}} - \text{RGB}_{\text{subject-distortion}})} \right]. \quad (9)$$

4 实验结果与讨论

4.1 实验结果

依据上述的图像质量客观评价方法, 采用 LIVE [19,20], TID2008 [21,22] 和 CSIQ [23,24] 三个数

据库中的共 47 幅参考图像 (如图 2, 图 3 和图 4) 和 1549 幅测试 (失真) 图像进行仿真实验, 对其进行评价, 获得图像质量客观评价分数 LCCM 值. 其中, 测试图像来源于对参考图像进行不同失真类型、失真程度的处理而获得, 3 个图像数据库来源及参数说明如表 1.

为了说明图像质量客观评价结果的准确性, 一般将其与主观评价分数 MOS (mean opinion score) 或 DMOS (difference mean opinion score) 值做相关性分析. 相关性分析一般包括: 1) 做质量主客观评价分数的散点图, 2) 计算相关性参数. 其中, 散点图主要是从直观上反映主客观评价结果的一致性程度; 一般散点图的纵坐标是质量主观评价分数 MOS 或 DMOS 值, 横坐标是采用模型评价得出的客观评价分数; 为了更好地说明主客观评价分数的一致性, 采用最小二乘法对主客观评价分数进行拟

表 1 3 个图像数据库及参数说明
Table 1. Three image databases and the parameters description.

图像库	参考图像	标记方法	参考图像被失真类型	失真程度	失真图像总数
TID2008	5 幅	I01—I05	Additive Gaussian noise (AGN), impulse noise (IN), additive noise in color components (AN), spatially correlated noise (SCN), masked noise (MN), high frequency noise (HFN), quantization noise (QN), Gaussian blur (GB), image denoising (ID), JPEG (JPEG), JPEG2000 (J2K), JPEG transmission errors (JTE), JPEG2000 transmission errors (J2KTE), non eccentricity pattern noise (NEPN), local block-wise distortions (LBWD), mean intensity shift (MIS), contrast change (CC) 共 17 种	4 种	340 幅
LIVE	12 幅	I01—I12	jpeg2k, jpeg, white noise, Gaussian blur, fast fading 共 5 种	5 或 6 种	323 幅
CSIQ	30 幅	I01—I30	Additive Gaussian white noise(agwn), Gaussian blurring(blur), additive pink Gaussian noise(fnoise), jpeg2k, jpeg, global contrast decrements(contrast) 共 6 种	4 或 5 种	886 幅



图 2 LIVE 数据库中提供的 12 幅参考图像
Fig. 2. Twelve reference images proposed by LIVE database.

合, 得出拟合曲线; 则当散点图中的散点越靠近拟合曲线(即离散程度越低), 曲线的单调性越好, 表明主客观评价分数的相关性越好, 一致性越高, 则其评价模型就越好, 反之则越差. 相关性参数主要是从量化上反映主客观评价结果的一致性程度; 相关性参数主要有: Pearson线性相关系数(Pearson linear correlation coefficient, LCC), Spearman秩相关系数(Spearman rank or-

der correlation coefficient, SROCC), 均方根误差(root mean square error, RMSE) 和散点图的离出率(outlier ratio, OR)^[25,26]. 其中LCC和SROCC值越接近于1, 表明相关性越好, 评价模型越好, 反之则越差; RMSE和OR越小, 表明主客观评价结果偏差越小, 评价结果越好, 反之则越差. 此4个参数值能够较好地数值上衡量主客观评价结果的一致性程度.



图3 TID2008数据库中提供的5幅参考图像
Fig. 3. Five reference images provided by TID2008 database.



图4 CSIQ数据库中提供的30参考图像
Fig. 4. Thirty reference images proposed by CSIQ database.

为了说明提出的LCCM图像质量客观评价模型的优劣, 将其评价1549幅图像的结果与LIVE, TID2008和CSIQ三个数据库中提供的主观评价

分数MOS或DMOS值做相关性分析, 且与SSIM, VSNR, PSNRHVS和FSIM模型的评价结果进行对比, 结果如图5, 计算的4个参数值如表2.

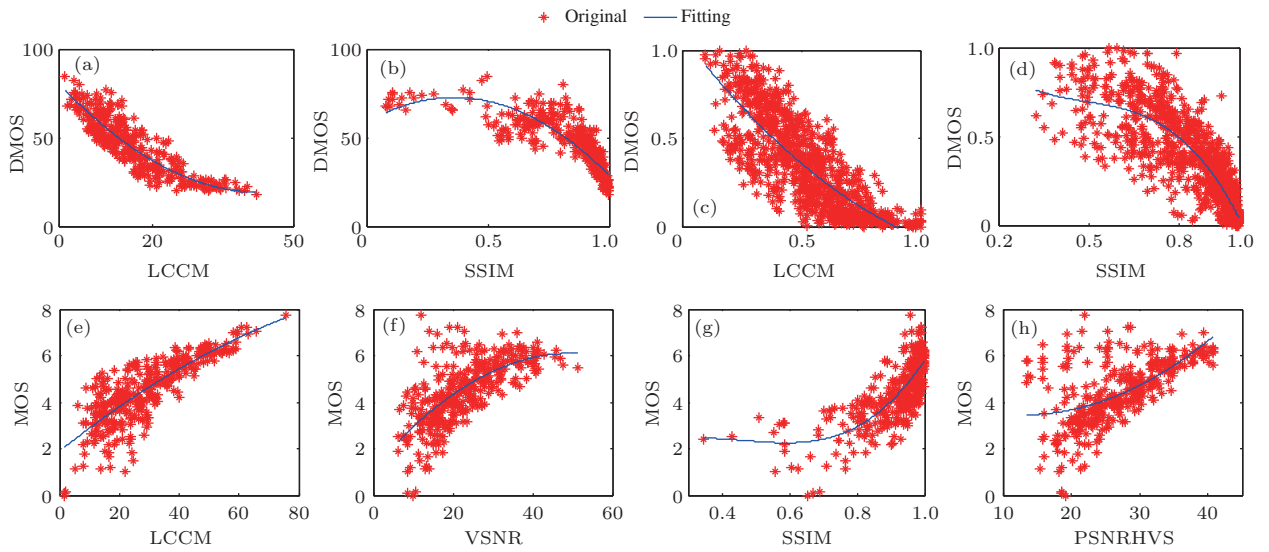


图5 LCCM, SSIM, VSNR和PSNRHVS模型评价的结果与3个数据库提供的主观评价分数MOS或DMOS值之间的散点图 (a) LIVE-LCCM; (b) LIVE-SSIM; (c) CSIQ-LCCM; (d) CSIQ-SSIM; (e) TID2008-LCCM; (f) TID2008-VSNR; (g) TID2008-SSIM; (h) TID2008-PSNRHVS

Fig. 5. Scatter plots of LCCM, VSNR, SSIM and PSNRHVS versus MOS or DMOS provided by the LIVE, CSIQ and TID2008 databases: (a) LIVE-LCCM; (b) LIVE -SSIM; (c) CSIQ-LCCM; (d) CSIQ-SSIM; (e) TID2008-LCCM; (f) TID2008-VSNR; (g) TID2008-SSIM; (h) TID2008-PSNRHVS.

表2 衡量图像主观评价一致性的4个参数值

Table 2. Four parameters to measure the consistency between subjective and objective evaluations of images.

数据库	算法	LCC	SROCC	RMSE	OR
CSIQ (866 幅图)	LCCM	0.8169	0.8805	0.1135	0.1126
	SSIM	0.8090	0.8700	0.1147	0.1215
	VSNR	0.7889	0.8270	0.1215	0.1322
	PSNRHVS	0.7834	0.8304	0.1219	0.1361
	FSIM	0.7996	0.9165	0.1197	0.1294
TID2008 (340 幅图)	LCCM	0.8105	0.8124	0.8449	0.1213
	SSIM	0.7132	0.7649	0.9550	0.1489
	VSNR	0.6694	0.7117	1.0067	0.1584
	PSNRHVS	0.5974	0.6048	1.1191	0.1792
LIVE (323 幅图)	LCCM	0.8647	0.8989	9.1855	0.1080
	SSIM	0.7585	0.8775	9.4497	0.1397

图5中的横坐标为采用LCCM, SSIM, VSNR和PSNRHVS模型评价的质量分数, 纵坐标为主观评价分数. 从图5中可以直观地发现, 相对于

SSIM, VSNR和PSNRHVS的评价分数, LCCM评价结果的离散程度明显偏小, 曲线的单调性好于其他三种方法的结果. 从相关性量化数值上, 表2中的

结果表明, LCCM模型的评价分数与DMOS或MOS的相关性系数LCC和SROCC值比SSIM的均得到了一定程度的提高, 提高幅度分别平均为9.5402%和3.2852%; 比VSNR和PSNRHVS提高幅度更大.

4.2 讨论

为了探讨图像内容复杂性和失真类型对图像质量评价准确度的影响, 以及结合人眼感知特性提高评价精度的途径, 从两个方面将LCCM模型评价结果与3个数据库中提供的主观评价结果进行相关性分析.

4.2.1 图像内容复杂性对质量评价的影响

分析每一幅源图像在不同的失真下的LCCM评价结果与主观评价结果的相关性, 并与SSIM评价结果进行对比分析, 探讨图像内容的复杂性对

质量客观评价准确度的影响. 结果如图6—图10. 其中, 图6是TID2008数据库中5幅源图像(TID-I01—TID-I05)经过17种类型、不同程度失真后, 采用LCCM和SSIM模型对其评价的结果与MOS值之间的散点图; 图7是LIVE数据库中12幅源图像(buildings, fountain, dancers, church, womanhat, sailing1, bikes, carnivaldools, building2, cemetery, lighthouse, caps)经过5种类型、不同程度失真后, 采用LCCM和SSIM模型对每幅源图像失真后的系列图像的评价结果与DMOS值之间的散点图; 图8和图9是CSIQ数据库中30幅源图像(Image01—Image30)经过6种类型、不同程度失真后, 分别采用LCCM和SSIM模型对其评价的结果与DMOS值之间的散点图; 图10是3个数据库中每幅源图像经过不同失真后, 采用LCCM和SSIM模型对其评价的质量分数与主观评价分数之间的相关性参数的对比.

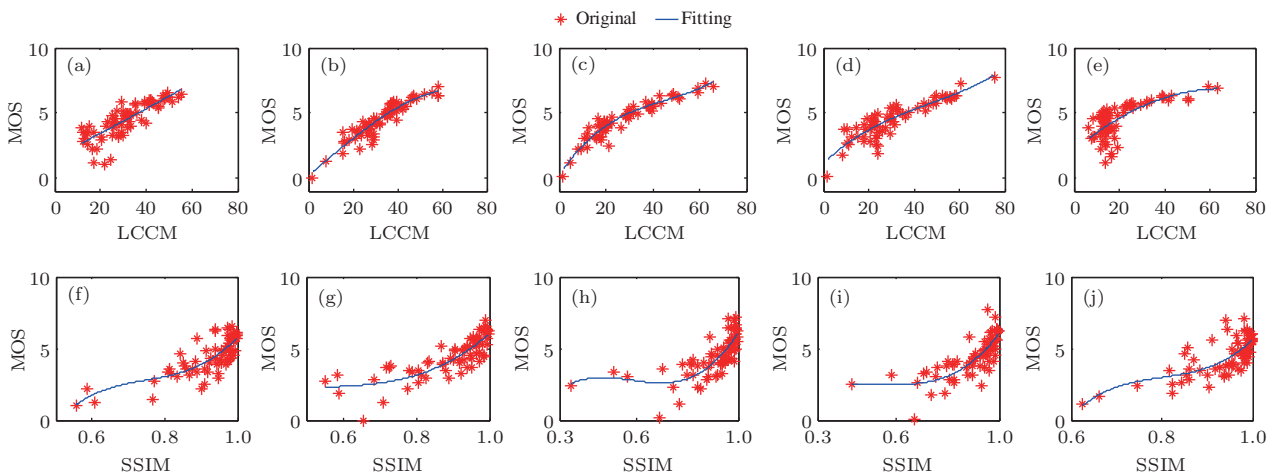


图6 TID2008数据库中5幅源图像每幅经过17种类型的失真后采用LCCM和SSIM模型对其评价的结果与MOS值之间的散点图 (a)—(e) TID-I01-LCCM—TID-I05-LCCM; (f)—(j) TID-I01-SSIM—TID-I05-SSIM

Fig. 6. Scatter plots between MOS and the objective evaluation scores obtained by using the LCCM and SSIM models to evaluate the tested images from any of five reference images in TID2008 database being distorted with 17 types: (a)—(e) TID-I01-LCCM—TID-I05-LCCM; (f)—(j) TID-I01-SSIM—TID-I05-SSIM.

图6—图9中的横坐标为针对三个数据库中的失真视频、采用LCCM或SSIM模型评价的质量分数, 纵坐标为其视频主观评价分数. 通过对比各散点图的离散程度和拟合曲线的单调性, 可以直观地分析主客观评价分数的一致性(即相关性), 从而判断评价模型的优劣. 结合3个数据库中源图像的内容特征, 对比分析图6—图9上的散点分布和图10中的相关性计算结果可以发现: 1) 当源图像本身比较模糊时, 如LIVE库中的图I03和CSIQ

库中的图I05, I10, I24, LCCM评价结果与主观评价结果的相关性明显高于SSIM评价结果的相关性; 2) 当源图像的纹理不复杂、图像内容场景较少、大面积单色调或相近色或平坦区域时, 如LIVE库中的图I05, I06, I08, I11, I12, CSIQ库中的图I06, I11, I14, I26, 以及TID2008库中的I01, I02, I03, LCCM评价精度在一定程度上高于SSIM的评价结果; 而当纹理复杂、场景较多、颜色丰富的图像, LCCM评价效果相近或差于SSIM.

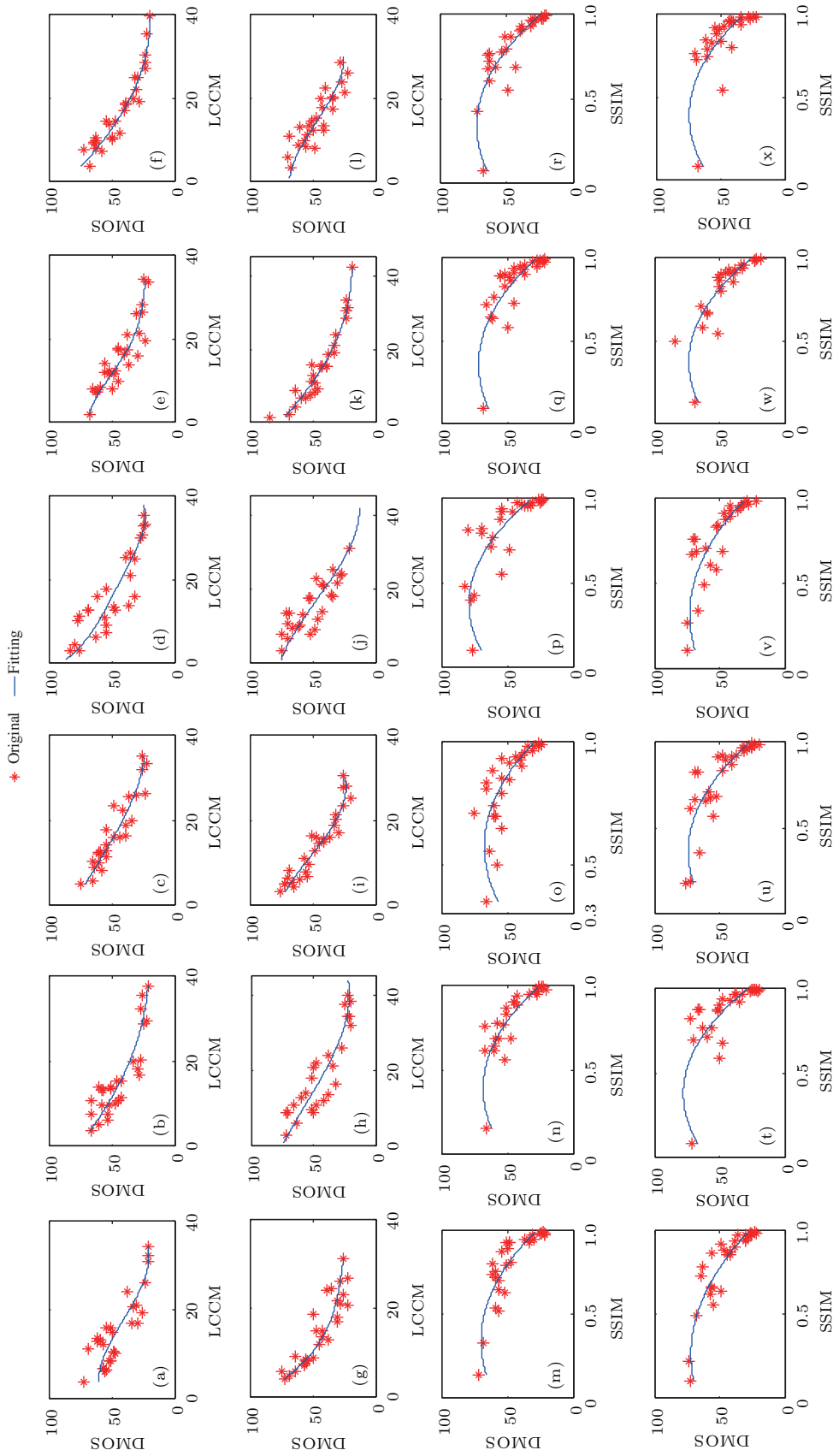


图7 LIVE数据库中12幅源图像每幅经过5种类型的失真后采用LCCM和SSIM模型对其评价的结果与DMOS值之间的散点图 (a) buildings-LCCM; (b) fountain-LCCM; (c) dancers-LCCM; (d) church-LCCM; (e) woman hat-LCCM; (f) sailing1-LCCM; (g) bikes-LCCM; (h) carnival dolls-LCCM; (i) buildings2-LCCM; (j) cemetery-LCCM; (k) lighthouse-LCCM; (l) caps-LCCM; (m) buildings-SSIM; (n) fountain-SSIM; (o) dancers-SSIM; (p) church-SSIM; (q) woman hat-SSIM; (r) sailing1-SSIM; (s) bikes-SSIM; (t) carnival dolls-SSIM; (u) buildings2-SSIM; (v) cemetery-SSIM; (w) lighthouse-SSIM; (x) caps-SSIM

Fig. 7. Scatter plots between DMOS and the objective evaluation scores obtained by using the LCCM and SSIM models to evaluate the tested images from any of 12 reference images in LIVE database being distorted with 5 types: (a) buildings-LCCM; (b) fountain-LCCM; (c) dancers-LCCM; (d) church-LCCM; (e) woman hat-LCCM; (f) sailing1-LCCM; (g) bikes-LCCM; (h) carnival dolls-LCCM; (i) buildings2-LCCM; (j) cemetery-LCCM; (k) lighthouse-LCCM; (l) caps-LCCM; (m) buildings-SSIM; (n) fountain-SSIM; (o) dancers-SSIM; (p) church-SSIM; (q) woman hat-SSIM; (r) sailing1-SSIM; (s) bikes-SSIM; (t) carnival dolls-SSIM; (u) buildings2-SSIM; (v) cemetery-SSIM; (w) lighthouse-SSIM; (x) caps-SSIM.

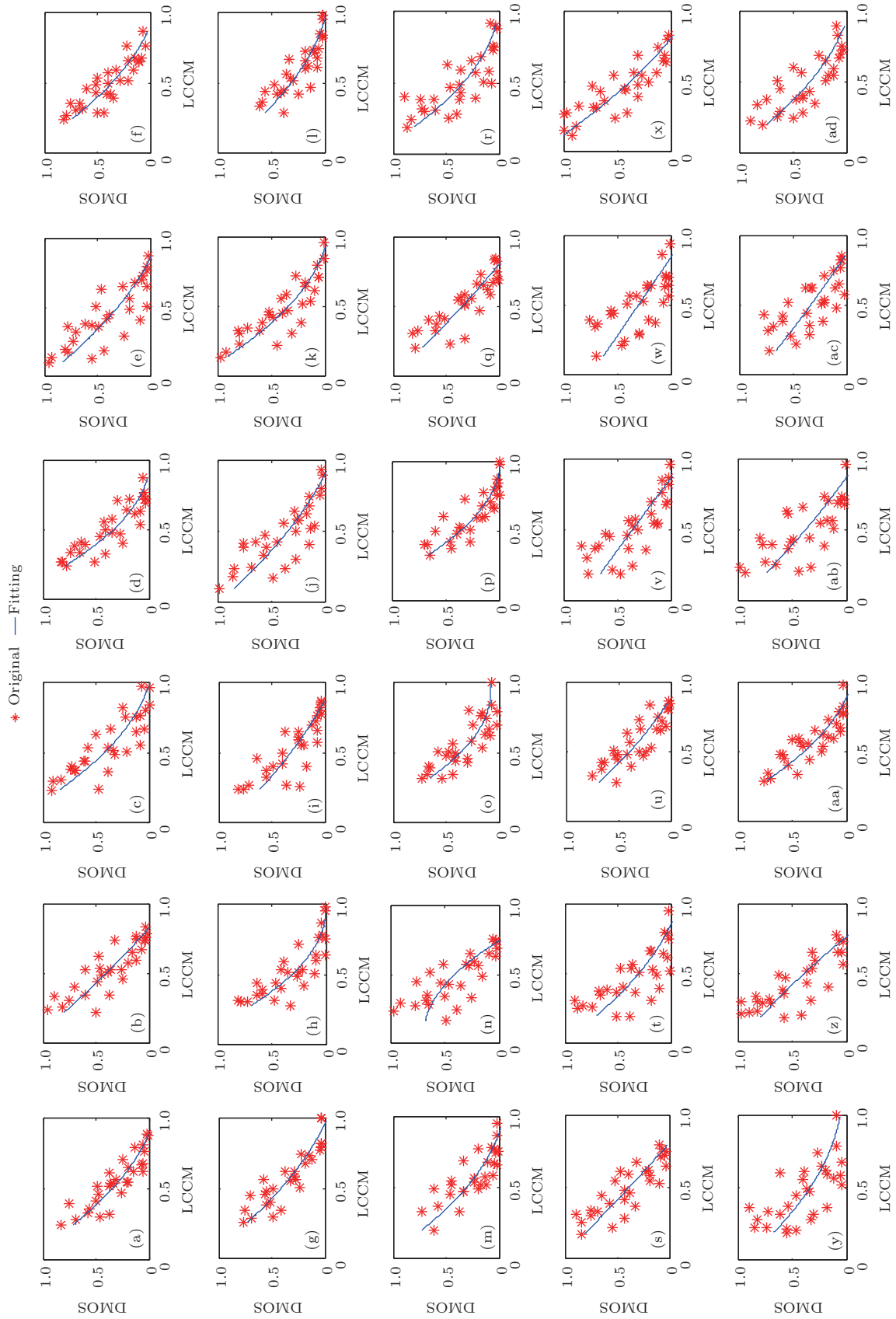


图8 CSIQ 数据库中30幅源图像每幅经过6种类型的失真后采用LCCM模型对其评价的结果与DMOS值之间的散点图 (a)–(ad) Image01—Image30
 Fig. 8. Scatter plots between DMOS and the objective evaluation scores obtained by using the LCCM models to evaluate the tested images from any of 30 reference images in CSIQ database being distorted with 6 types: (a)–(ad) Image01–Image30.

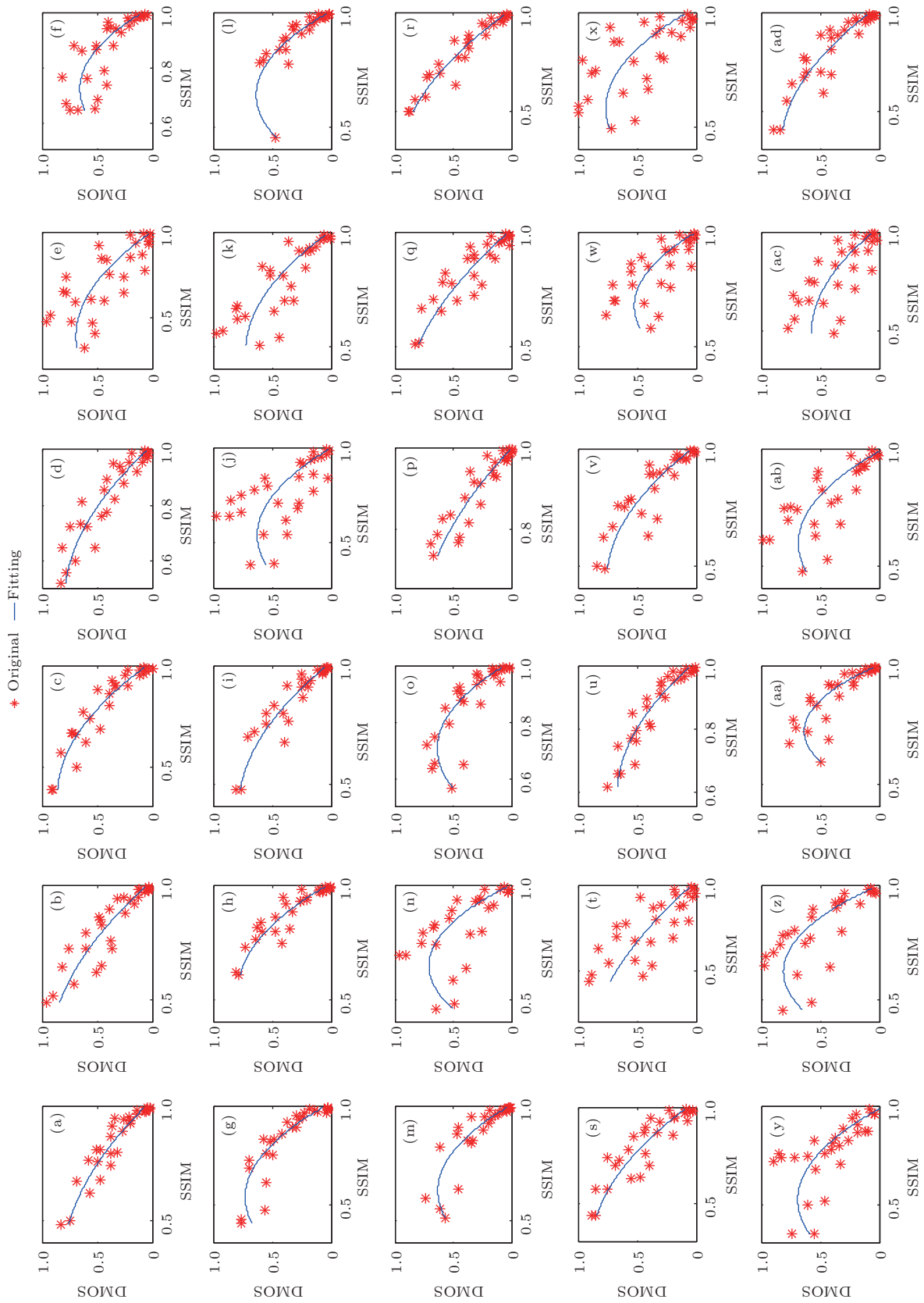


图9 CSIQ 数据库中 30 幅源图像每幅经过 6 种类型的失真后采用 SSIM 模型对其评价的结果与 DMOS 值之间的散点图 (a)—(ad) Image01—Image30 in CSIQ database being distorted with 6 types: (a)—(ad) Image01—Image30.

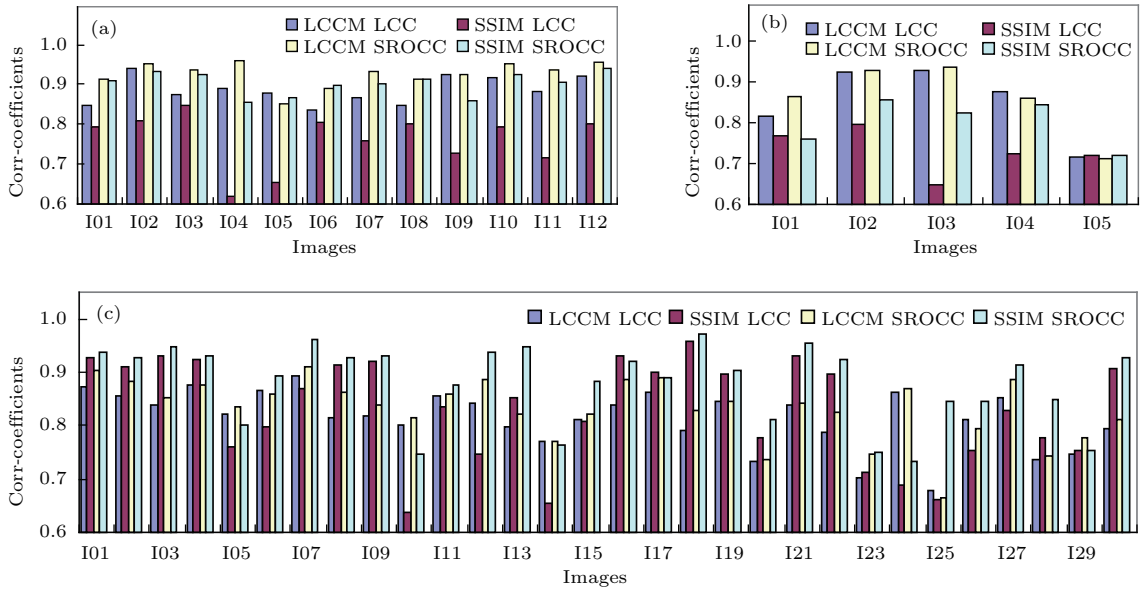


图 10 3 个数据库中每幅源图像经过不同失真后采用 LCCM 和 SSIM 模型评价的质量分数与主观评价分数之间相关性的对比 (a) LIVE; (b) TID2008; (c) CSIQ

Fig. 10. Comparing correlation between the subjective scores and the objective results from evaluating all distorted images of each reference image in there databases with LCCM and SSIM respectively: (a) LIVE; (b) TID2008; (c) CSIQ.

分析其原因主要是对于本身模糊、纹理不复杂、场景较少、大面积平坦区域等特征的源图像, 存在更多的视觉冗余, 人眼感知其失真的阈值更大; 而在图像质量评价时, 只有当图像失真量超过阈值时, 人眼才能感知, 但对于 SSIM 等模型的评价, 只要有失真, 其评价结果便会发生变化; 更大的失真感知阈值增加了 SSIM 等客观评价模型的主客观评价的不一致, 导致其模型的评价误差更大. 这需要采用结合更多视觉特性的评价模型来评价. 而 LCCM 模型考虑了人眼感知亮度的非线性、掩蔽特性和对比度敏感特性, 相对来说, 其结合了更多的人眼视觉特性, 所以对于此类图像的评价效果要优于 SSIM, PSNRHVS 和 MSE 等模型. 具体的原因分析如下.

1) 掩蔽特性. 当以上此类源图像失真时, 因为掩蔽特性使得人眼感知其失真的阈值变大, 原本能够感知的图像失真, 因为感知失真阈值的增加而现在不能被感知; 人眼容忍图像失真的程度增加, 则使得更大范围内的失真图像的主观评价分数基本不发生变化 (图像的失真不超过感知失真阈值); 但对于 SSIM 等模型, 只要有失真, 其评价结果就要发生变化; 而且由于感知失真阈值的增加, 主观分数基本不变的失真图像的客观评价分数变化更大, 所以, 人眼掩蔽特性是导致主客观评价不一致的原因之一. 而 LCCM 模型结合了人眼对比敏感度特

性, 当失真量没有超过感知失真阈值时, 人眼敏感值非常小或者为零, 感知图像上几乎没有差异, 从而 LCCM 模型能够很好地减少因人眼掩蔽特性和图像视觉冗余带来的质量评价结果的误差, 提高了其评价精度.

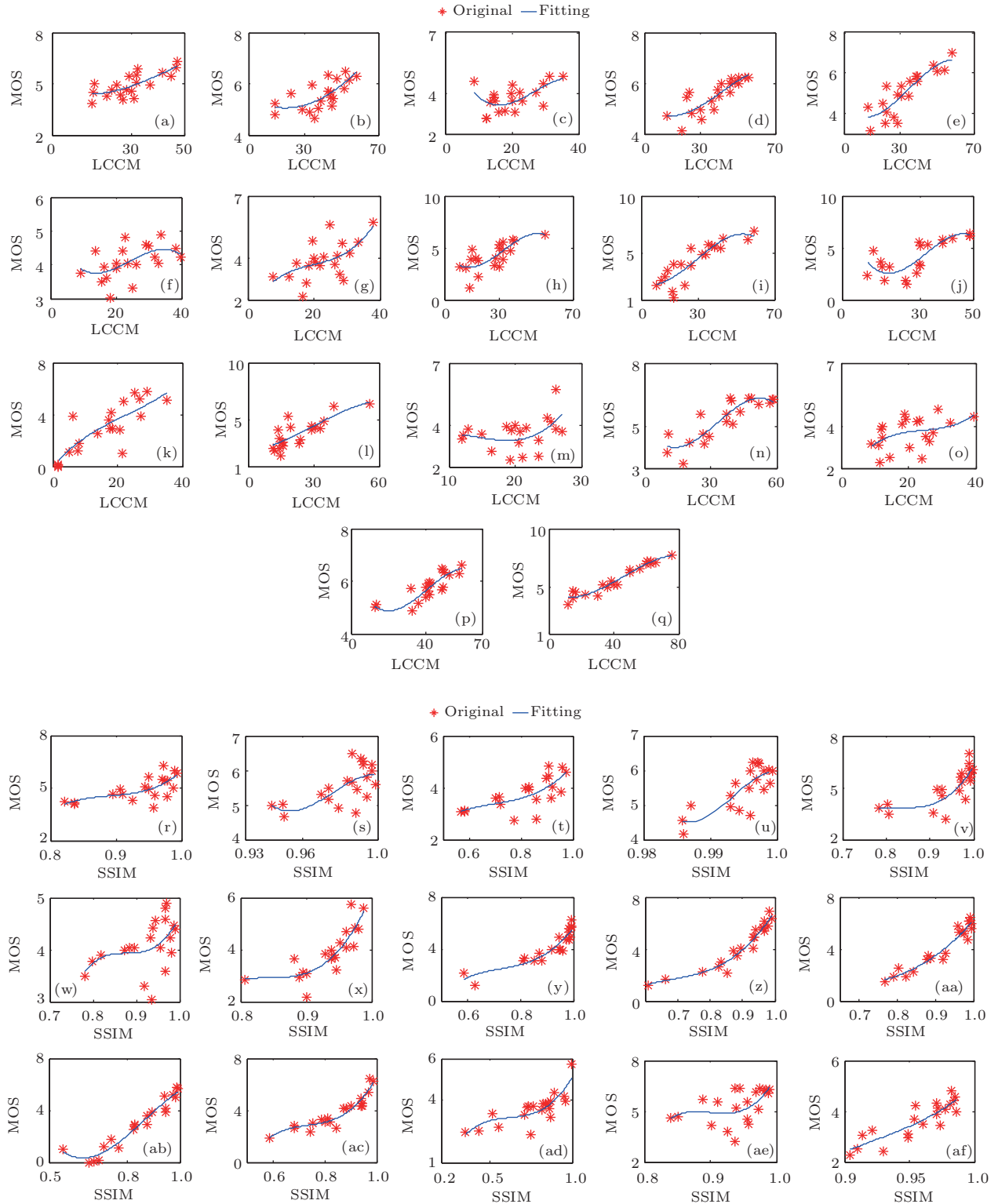
2) 人眼感知亮度的非线性. 人眼对亮度 (或能量强度) 的感知是非线性的, 呈对数变化趋势, 即人眼感知到的失真量比实际的失真量要小. 但对于 SSIM 等模型是按照实际的失真量来评价图像间的差异, 所以其也是导致主客观评价不一致的原因之一. 而 LCCM 模型在构建时, 将图像的亮度和色度的强度均按照人眼感知强度模型 ((6) 式) 进行了转换, 从而使其评价分数能够有效地避免此视觉特性带来的误差.

3) 人眼对比敏感度特性. 人眼对目标的敏感程度主要受到周围环境和空间频率的影响, 即相同的目标在不同的环境中、不同的空间位置, 其人眼感知是不同的. 则基于此特性, 当图像有相同的失真量时, 不同的图像, 人感知的结果是不同的, 但对于 SSIM 等模型, 相同的失真量其评价结果即相同, 从而导致了主客观评价的不一致. 所以, 人眼敏感特性也是导致主客观评价不一致的原因之一. 而在 LCCM 模型的构建中, 是以对比敏感度特性值作为感知亮度的权重因子, 所以能够较好地克服评价过程中其带来的误差, 提高了评价结果的精度.

4.2.2 失真类型对质量评价的影响

分析每一种类型、不同程度的失真下所有图像的LCCM评价结果与主观评价结果的相关性, 并与SSIM进行对比分析, 探讨图像的失真方式对质量客观评价准确度的影响. 结果如图11—图14.

其中, 图11—图13中的横坐标为采用LCCM或SSIM模型评价三个数据库中所有类型、不同失真程度的视频的质量分数, 纵坐标为其相应视频质量主观评价分数. 图中每一个子图代表一种失真类型的评价结果, 子图中的曲线为采用最小二乘法拟合的结果.



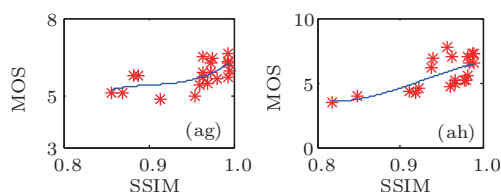


图 11 TID2008 数据库中 17 种失真类型的图像分别采用 LCCM 和 SSIM 模型评价的分数与 MOS 之间的散点图 (a) AGN-LCCM; (b) AN-LCCM; (c) SCN-LCCM; (d) MN-LCCM; (e) HFN-LCCM; (f) IN-LCCM; (g) QN-LCCM; (h) GB-LCCM; (i) ID-LCCM; (j) JC-LCCM; (k) J2K-LCCM; (l) JT-LCCM; (m) J2KT-SSIM; (n) NEPN-SSIM; (o) LBWI-SSIM; (p) MIS-SSIM; (q) CC-SSIM; (r) AGN-SSIM; (s) AN-SSIM; (t) SCN-SSIM; (u) MN-SSIM; (v) HFN-SSIM; (w) IN-SSIM; (x) QN-SSIM; (y) GB-SSIM; (z) ID-SSIM; (aa) JC-SSIM; (ab) J2K-SSIM; (ac) JT-SSIM; (ad) J2 KT-SSIM; (ae) NEPN-SSIM; (af) LBWI-SSIM; (ag) MIS-SSIM; (ah) CC-SSIM

Fig. 11. Scatter plots between the MOS and the objective evaluation scores obtained by evaluating the images with 17 distortions using the LCCM and SSIM models in TID2008 database: (a) AGN-LCCM; (b) AN-LCCM; (c) SCN-LCCM; (d) MN-LCCM; (e) HFN-LCCM; (f) IN-LCCM; (g) QN-LCCM; (h) GB-LCCM; (i) ID-LCCM; (j) JC-LCCM; (k) J2K-LCCM; (l) JT-LCCM; (m) J2KT-SSIM; (n) NEPN-SSIM; (o) LBWI-SSIM; (p) MIS-SSIM; (q) CC-SSIM; (r) AGN-SSIM; (s) AN-SSIM; (t) SCN-SSIM; (u) MN-SSIM; (v) HFN-SSIM; (w) IN-SSIM; (x) QN-SSIM; (y) GB-SSIM; (z) ID-SSIM; (aa) JC-SSIM; (ab) J2K-SSIM; (ac) JT-SSIM; (ad) J2KT-SSIM; (ae) NEPN-SSIM; (af) LBWI-SSIM; (ag) MIS-SSIM; (ah) CC-SSIM.

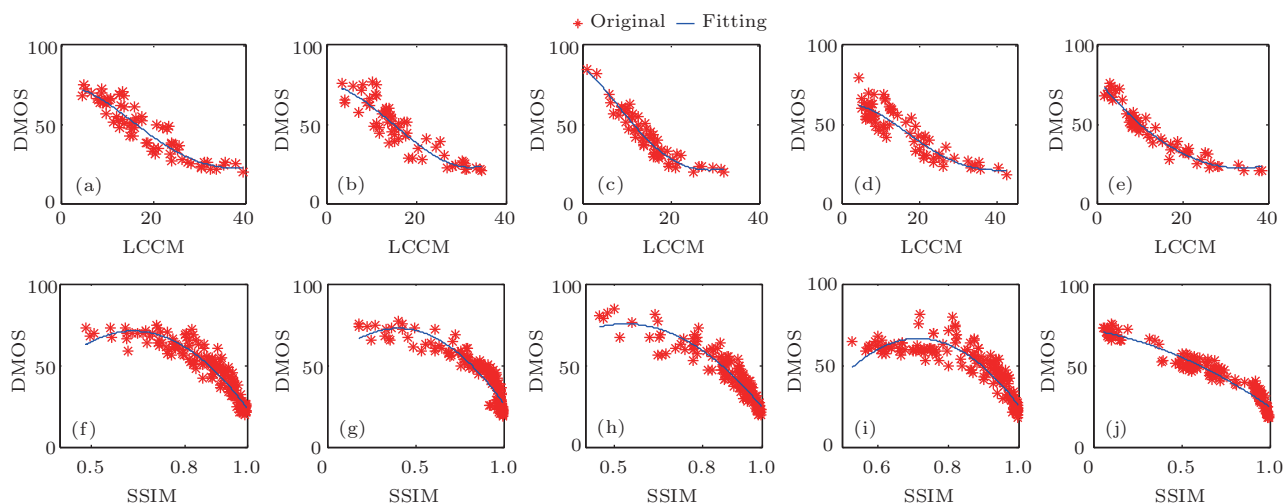


图 12 LIVE 数据库中 5 种失真类型的图像分别采用 LCCM 和 SSIM 模型评价的分数与 DMOS 之间的散点图 (a) jpeg2k-LCCM; (b) fastfading-LCCM; (c) Gaussianblur-LCCM; (d) jpeg-LCCM; (e) whitenoise-LCCM; (f) jpeg2k-SSIM; (g) fastfading-SSIM; (h) Gaussianblur-SSIM; (i) jpeg-SSIM; (j) whitenoise-SSIM

Fig. 12. Scatter plots between the DMOS and the objective evaluation scores obtained by evaluating images with 5 distortions using the LCCM and SSIM models in LIVE database: (a) jpeg2k-LCCM; (b) fastfading-LCCM; (c) Gaussianblur-LCCM; (d) jpeg-LCCM; (e) whitenoise-LCCM; (f) jpeg2k-SSIM; (g) fastfading-SSIM; (h) Gaussianblur-SSIM; (i) jpeg-SSIM; (j) whitenoise-SSIM.

基于散点图的离散程度和拟合曲线的单调性, 对比分析图 11—图 13 中各种不同类型失真下的图像质量评价结果的散点图, 并结合图 14 中计算的相关性参数值, 可以发现: LCCM 对由 AGN, MN, NEPN, HFN, MIS, CC 类型失真的图像的质量评价分数与主观评价分数的相关性明显高于 SSIM 的; 而对由 SCN, QN, JTE, J2 KTE, LBWD 类型失真的图像的评价结果与主观评价结果的一致性低于或接近于 SSIM 评价结果的一致性.

结合不同失真类型的特征和人眼感知特性, 分析其原因主要为: 1) AGN 等 6 种类型的失真均是对图像每个小区域进行了相近程度的失真, 图像中不会出现局部对比度明显变化, 因此, 图像存在较多的视觉冗余, 人眼能够感知到的失真量上限仍比较大, 从而给 SSIM 等模型的评价结果带来了较大的误差; 而 LCCM 模型结合了人眼对比敏感度特性, 较好地克服了由此给质量评价带来的误差, 从而提高了评价精度; 反之, SCN 等 5 种失真对图像局部破坏较大, 致使图像局部对比度改变明显,

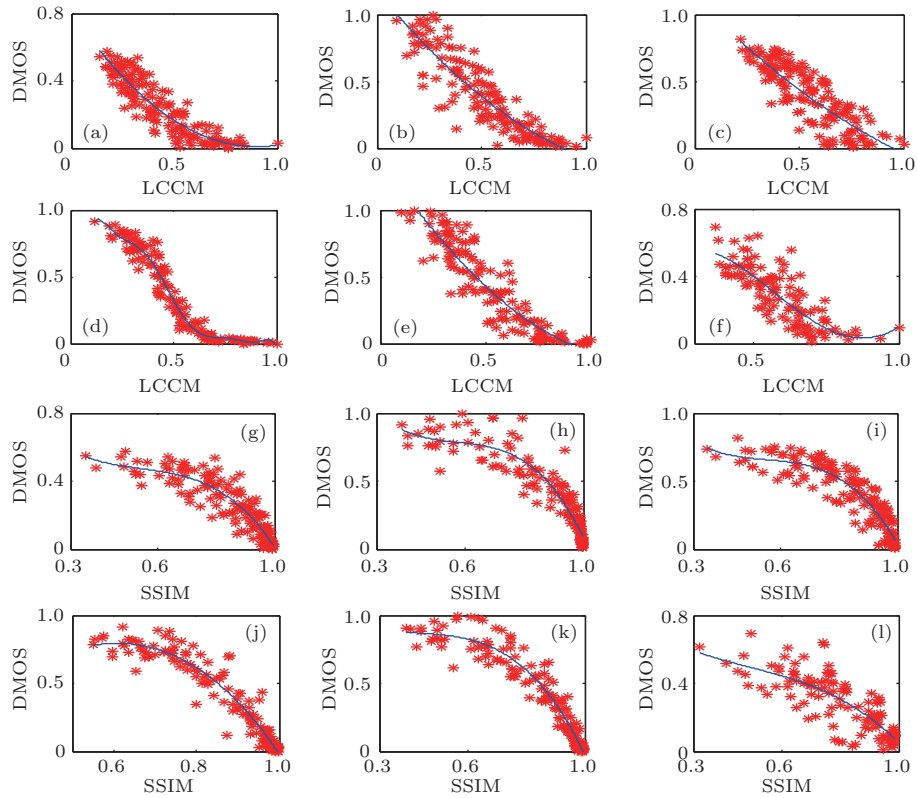


图 13 CSIQ 数据库中 6 种失真类型的图像分别采用 LCCM 和 SSIM 模型评价的分数与 DMOS 之间的散点图 (a) agwn-LCCM; (b) blur-LCCM; (c) fnoise-LCCM; (d) jpeg-LCCM; (e) jpeg2k-LCCM; (f) contrast-LCCM; (g) agwn-SSIM; (h) blur-SSIM; (i) fnoise-SSIM; (j) jpeg-SSIM; (k) jpeg2k-SSIM; (l) contrast-SSIM

Fig. 13. Scatter plots between the DMOS and the objective evaluation scores obtained by evaluating images with 6 distortions using the LCCM and SSIM models in CSIQ database: (a) agwn-LCCM; (b) blur-LCCM; (c) fnoise-LCCM; (d) jpeg-LCCM; (e) jpeg2k-LCCM; (f) contrast-LCCM; (g) agwn-SSIM; (h) blur-SSIM; (i) fnoise-SSIM; (j) jpeg-SSIM; (k) jpeg2k-SSIM; (l) contrast-SSIM.

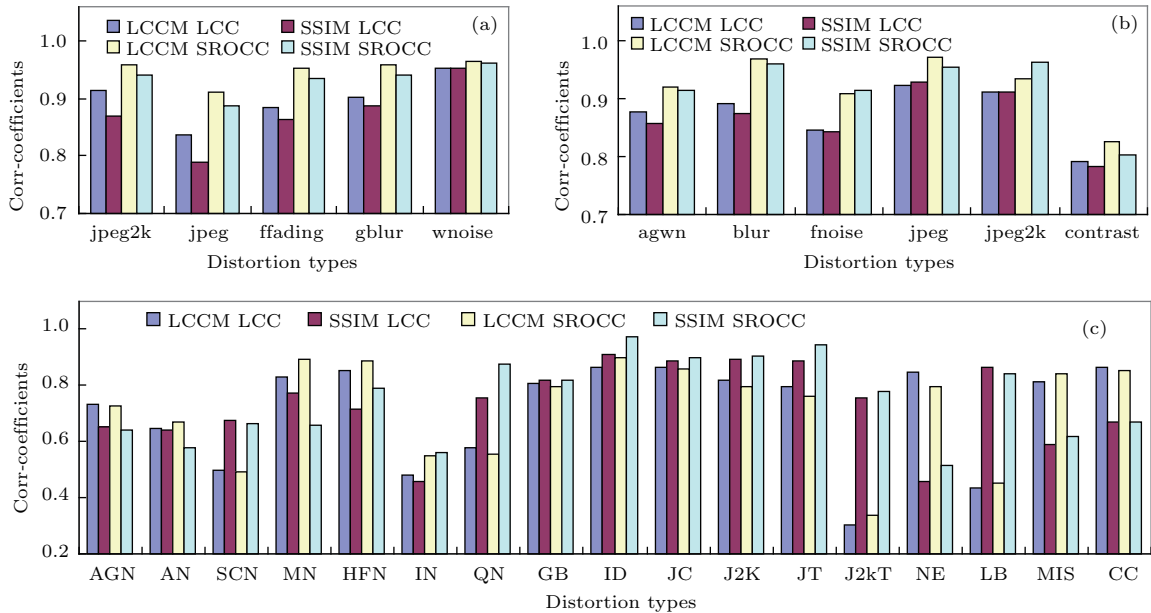


图 14 三个数据库中的图像经过每种类型的失真后分别采用 LCCM 和 SSIM 评价的结果与主观评价结果之间相关性的对比 (a) LIVE; (b) CSIQ; (c) TID2008

Fig. 14. Comparing LCC and SROCC between the subjective evaluation scores and the objective evaluation scores obtained using the LCCM and SSIM models to evaluate the images processed with every type of distortion in LIVE, TID2008 and CSIQ database: (a) LIVE; (b) CSIQ; (c) TID2008.

人眼对其失真非常敏感, 图像的失真量均能被较好地感知, 主客观评价分数已经达到了各自的最佳精度, 一致性很难再得到提升; 2) 人眼对色度的敏感程度低于对亮度的敏感程度, 当颜色有较小的失真时, 人眼不能感知或敏感度较低, 从而也成为评价一致性较差的原因之一。

另外, 提出的LCCM评价方法还能较好地避免因为拍摄焦点的不同而对质量评价结果带来的影响。在拍摄图像时, 对于同一景物, 焦点不同, 图像的表现效果不同, 则其质量评价的结果不同。提出的质量评价算法由于采用 $8 \text{ pixels} \times 8 \text{ pixels}$ 大小的子块的评价结果的平均值, 即每次评价主要聚焦的是图像中一个非常小的部分, 从而能够较好地避免因为拍摄时焦点不同而带来的影响。为了验证所提评价方法的有效性, 对LIVE和TID2008数据库中同一景物但不同的拍摄焦点的3幅图像(LIVE库中的I05, I07, I12与TID2008库中I03, I04, I05是同一景物, 但拍摄的焦点不同)进行计算, 得出LCCM值和SSIM值, 并与其数据库中提供的MOS值进行对比分析, 计算得出线性相关性系数, 结果如图15。

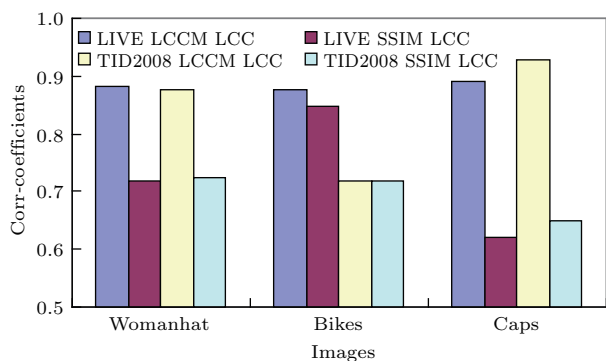


图15 采用LCCM和SSIM模型对3幅不同拍摄焦点的图像的质量评价结果

Fig. 15. Quality evaluation of three images focused differently using the LCCM and SSIM models.

从图15中可以看到: 对于3幅图像, 采用LCCM方法评价得出的结果与主观评价分数的相关性比SSIM方法评价的相关性明显偏高。

5 结 论

结合人眼空间感知特性和图像的复杂性特征, 提出了一种基于人眼对图像内容感知的图像质量客观评价方法。该方法首先结合人眼对亮度感知的非线性、掩蔽特性和对比度敏感特性, 构建了一

种图像内容感知模型; 再以此模型分别模拟人眼感知源图像和失真图像, 并计算其强度差; 最后基于强度差构建图像质量客观评价模型。采用LIVE, TID2008和CSIQ三个数据库中的共47幅参考图像和1549幅测试图像进行仿真实验, 并与SSIM, PSNRHVS和VSNR等典型的图像质量客观评价模型进行对比分析。结果表明: 所提方法的评价分数与主观评价分数的Pearson线性相关性系数和Spearman秩相关系数值比SSIM的评价结果均有一定程度的提高, 提高幅度分别平均为9.5402%和3.2852%, 比PSNRHVS和VSNR提高幅度更大。表明在图像质量客观评价中, 考虑人眼对图像内容的感知和复杂度的分析有助于提高图像质量主观评价的一致性, 评价精度可得到进一步的提高。

参考文献

- [1] Wang Y Q 2014 *J. Nanjing Univ. (Nat. Sci. Ed.)* **50** 361 (in Chinese) [王元庆 2014 南京大学学报(自然科学版) **50** 361]
- [2] Zhuang J Y, Chen Q, He W J, Mao T Y 2016 *Acta Phys. Sin.* **65** 040501 (in Chinese) [庄佳衍, 陈钱, 何伟基, 冒添逸 2016 物理学报 **65** 040501]
- [3] Wang Z, Bovik A C, Sheikh H R, Simoncelli E P 2004 *IEEE Trans. Image Process.* **13** 600
- [4] Chandler D M, Hemami S S 2007 *IEEE Trans. Image Process.* **16** 2284
- [5] Zhang L, Zhang L, Mou X, Zhang D 2011 *IEEE Trans. Image Process.* **20** 2378
- [6] Xue W, Zhang L, Mou X, Bovik A C 2014 *IEEE Trans. Image Process.* **23** 684
- [7] Zhang L, Shen Y, Li H 2014 *IEEE Trans. Image Process.* **23** 4270
- [8] Paudyal P, Battisti F, Sjöström M, Olsson R, Carli M 2017 *IEEE Trans. Broadcast.* **63** 507
- [9] Bae S H, Kim M 2016 *IEEE Trans. Image Process.* **25** 2392.
- [10] Gu K, Wang S, Zhai G, Ma S, Yang X 2016 *Signal Image Video Process.* **10** 803
- [11] Wang X L, Wu D W, Li X, Zhu H N, Chen K, Fang G 2017 *Acta Phys. Sin.* **66** 230302 (in Chinese) [王湘林, 吴德伟, 李响, 朱浩男, 陈坤, 方冠 2017 物理学报 **66** 230302]
- [12] Akamine W Y L, Farias M C Q 2014 *J. Electron. Imaging* **23** 061107
- [13] Li C F, Bovik A C 2010 *J. Electron. Imaging* **19** 143
- [14] Guo J, Hu G, Xu W, Huang L 2017 *J. Vis. Commun. Image Represent.* **43** 50
- [15] Hou W, Mei F H, Chen G J, Deng X W 2015 *Acta Phys. Sin.* **64** 024202 (in Chinese) [侯旺, 梅风华, 陈国军, 邓喜文 2015 物理学报 **64** 024202]
- [16] Stephen W, Huw O, Vien C, Iain P S 2006 *Color Res. Appl.* **31** 315

- [17] Nadenau M 2000 *Ph. D. Dissertation* (Lausanne: École Polytechnique Fédérale de Lausanne)
- [18] Yao J C, Shen J, Wang J H 2008 *Acta Phys. Sin.* **57** 4034 (in Chinese) [姚军财, 申静, 王剑华 2008 物理学报 **57** 4034]
- [19] Sheikh H R, Wang Z, Cormack L LIVE Image Quality Assessment Database Release 2 <http://live.ece.utexas.edu/research/quality> [2017-12-17]
- [20] Sheikh H R, Sabir M F, Bovik A C 2006 *IEEE Trans. Image Process.* **15** 3440
- [21] Nikolay N P Tampere Image Database 2008 TID2008, version 1.0 <http://www.ponomarenko.info/tid2008.htm> [2017-12-17]
- [22] Nikolay P, Vladimir L, Alexander Z, Karen E, Jaakko A, Marco C, Federica B 2009 *Adv. Modern Radioelectron.* **10** 30
- [23] Larson E C The CSIQ image database <http://vision.ok-state.edu/?loc=csiq> [2017-12-17]
- [24] Larson E C, Chandler D M 2010 *J. Electron. Imaging* **19** 011006
- [25] Zhang F, Bull D R 2013 *Proceedings of the 20th IEEE Interatinonal Conference on Image Processing (ICIP)* Melbourne, Australia, September 15–18, 2013 p39
- [26] Gu K, Wang S, Zhai G, Lin W, Yang X, Zhang W 2016 *IEEE Trans. Image Process.* **62** 446

Objective assessment method of image quality based on visual perception of image content*

Yao Jun-Cai¹⁾²⁾ Liu Gui-Zhong^{1)†}

1) (School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

2) (School of Physics and Telecommunication Engineering, Shaanxi University of Technology, Hanzhong 723000, China)

(Received 23 January 2018; revised manuscript received 7 March 2018)

Abstract

Objective image quality assessment (IQA) plays a very important role in transmission, encoding, and quality of service (QoS) of the image and video data. However, the existing IQA methods often do not consider image content features and their visual perception, so there is a certain gap between the objective IQA scores and the subjective perception. To solve this problem, in the study, we propose an objective IQA method based on the visual perception of image content, which combines the complexity characteristics of image content, and the properties of masking, contrast sensitivity and luminance perception nonlinearity of human visual system (HVS). In the proposed method, the image is first transformed using a nonlinear model of luminance perception to obtain the intensity perception image. Then, the intensity information is summed using the contrast sensitivity values of HVS and the average contrast values of the local image as a weighting factor of the intensity. The summed data information is taken as the content of human perceiving image, and an image perception model is constructed. Finally, the reference images and distorted images are perceived by simulating the HVS with this model. Moreover, the difference in intensity between two perceived images is calculated. Based on the intensity difference and peak signal-to-noise ratio model, an objective IQA model is constructed. Further, the simulation with 47 reference images and 1549 test images in the LIVE, TID2008, and CSIQ databases is conducted. Moreover, the experimental results are compared with those of four typical objective IQA models, namely SSIM, VSNR, FSIM, and PSNRHVS. In addition, we explore the factors that affect the IQA accuracy and a way to improve assessment accuracy by combining HVS characteristics, through analyzing the correlation between IQA results of the proposed model and the subjective mean opinion scores (MOSs) provided in the three image databases from the following two aspects. Namely, (1) all reference images in three image databases are distorted by multiple types, and the distorted

* Project supported by the National Natural Science Foundation of China (Grant No. 61301237) and the Natural Science Foundation of Shaanxi Province, China (Grant No. 2015KJXX-42).

† Corresponding author. E-mail: liugz@mail.xjtu.edu.cn

images of each reference image are taken as a test sequence. Then, the proposed model is used to evaluate each test sequence to obtain the IQA scores. By analyzing the correlation between the IQA scores of each test sequence and the subjective MOSs and comparing them with the assessment results of SSIM, we explore the influence of the image content complexity on the objective IQA accuracy. (2) The test images which are distorted by each type and many distortion degrees are used as another sequence, and they are evaluated by the proposed IQA model. By analyzing the correlation between the subjective MOSs and the IQA results of each test sequence, and comparing them with assessment results of SSIM, we discuss the influence of image distortion mode on the IQA accuracy. The experimental results show that the coefficient values of Pearson linear correlation and Spearman rank order correlation between the objective IQA scores obtained by the proposed method and the subjective MOSs have been averagely improved by 9.5402% and 3.2852%, respectively, in comparison with IQA results from the SSIM method. Also, they are enhanced more significantly than those from the PSNRHVS and VSNR methods. In summary, it is shown that the proposed IQA method is an effective and feasible method of objectively assessing the image quality; moreover, it is shown that in the objective assessment of image quality it is very helpful to improve the consistency of subjective and objective assessment of image quality by considering the content perception and complexity analysis of the images.

Keywords: image quality evaluation, image content, human visual system characteristics, correlation coefficient

PACS: 87.57.C-, 07.05.Pj, 42.30.-d

DOI: [10.7498/aps.67.20180168](https://doi.org/10.7498/aps.67.20180168)