

结合空间域与变换域特征提取的盲立体图像质量评价

陈 勇^{*①} 金曼莉^① 朱凯欣^① 刘焕淋^② 陈 东^①

^①(重庆邮电大学工业物联网与网络化控制教育部重点实验室 重庆 400065)

^②(重庆邮电大学通信与信息工程学院 重庆 400065)

摘 要: 针对立体图像质量预测准确性不足的问题, 该文提出了一种结合空间域和变换域提取质量感知特征的无参考立体图像质量评价模型。在空间域和变换域分别提取输入的左、右视图的自然场景统计特征, 并在变换域提取合成独眼图的自然场景统计特征, 然后将其输入到支持向量回归(SVR)中, 训练从特征域到质量分数域的预测模型, 并以此建立SIQA客观质量评价模型。在4个公开的立体图像数据库上与一些主流的立体图像质量评价算法进行对比, 以在LIVE 3D Phase I图像库中的性能测试为例, Spearman秩相关系数、皮尔逊线性相关系数和均方根误差分别达到0.967, 0.946和5.603, 验证了所提算法的有效性。

关键词: 立体图像; 无参考; 空间域特征; 变换域特征

中图分类号: TN911.73; TP391.41

文献标识码: A

文章编号: 1009-5896(2021)10-2958-09

DOI: [10.11999/JEIT200694](https://doi.org/10.11999/JEIT200694)

Blind Stereo Image Quality Evaluation Based on Spatial Domain and Transform Domain Feature Extraction

CHEN Yong^① JIN Manli^① ZHU Kaixin^① LIU Huanlin^② CHEN Dong^①

^①(Key Laboratory of Industrial Internet of Things & Network Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

^②(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: For the problem of insufficient accuracy of stereo image quality prediction, a blind stereoscopic image quality assessment model combining spatial domain and transform domain to extract quality-aware features is proposed. Firstly, the statistical features of the natural scenes in the left and right views are extracted respectively in space domain and transformation domain, and statistical features of natural scenes from synthetic monocular images is extracted in transformation domain. Finally, Support Vector Regression (SVR) is used to train a stereoscopic image quality evaluation model from the feature domain to the quality score domain, so as to establish SIQA objective quality evaluation model. The performance of the proposed method is compared with some state-of-the-art full-reference, reduced-reference and no-reference stereoscopic image quality evaluation algorithms on the four public stereo image databases, taking the performance test in live 3D phase I image library as an example. SROCC of 0.967, PLCC of 0.946 and RMSE of 5.603 are achieved, which verifies the effectiveness of the proposed algorithm.

Key words: Stereoscopic image; No reference; Spatial domain characteristics; Transform domain features

1 引言

图像在获取、传输和存储的过程中不可避免地引入失真, 影响观看者的视觉体验质量。因此,

设计与人眼主观评分一致的立体图像质量评估 (Stereoscopic Image Quality Assessment, SIQA) 模型对于3D图像质量优化和处理至关重要^[1]。

人类视觉系统 (Human Visual System, HVS) 可以将左眼和右眼获得的不同视图转换为单幅视图, 在转换过程中, 并不是简单将左右眼刺激相加就可以获得单一视图^[2,3]。与2D图像质量评价相比, 3D除了需要考虑深度信息和双目视觉特性对立体图像质量评价的影响^[4]。Ding等人^[5]通过加

收稿日期: 2020-08-06; 改回日期: 2021-07-23; 网络出版: 2021-08-27

*通信作者: 陈勇 chenrong@cqupt.edu.cn

基金项目: 国家自然科学基金 (51977021)

Foundation Item: The National Natural Science Foundation of China (51977021)

权组合的方式来模拟复杂的双目视觉机制；Ding等人^[6]将log-Gabor滤波器和局部二值算子(Local Binary Pattern, LBP)应用到立体图像的左右视图上来获取其频率幅度和纹理图；Hachicha等人^[7]利用伯努利广义高斯(Bernoulli Generalized Gaussian, BGG)分布对其最终的小波子带进行建模；文献^[8]采用不同的双目组合策略构建两幅独眼图，并分别提取其双目特征，从而使该模型的预测分数与人眼主观评分更加接近。由于自然图像(Natural Scene Statistics, NSS)在空间域和频域上都表现出特定的统计规律^[9]，并且随自然图像NSS中不同类型的失真的引入，其统计规律会发生变化。

对此，本文提出了一种结合空间域和变换域提取质量感知特征的无参考立体图像质量评价模型。先将输入的左、右视图进行分块并进行分块判别，然后构建模拟左右视图在大脑融合过程中的双目组合模型，并分别在空间域和变换域提取自然统计特征。利用支持向量回归(Support Vector Regression, SVR)训练从特征域到质量分数域的立体图像质量评价模型。

2 SIQA模型整体框架的构建

图1给出了本文SIQA模型建立的整体框架。其主要原理为：将左、右视图分割成大小为 $N \times N$ 图像块，通过比较左、右图像块的结构相似度及所含信息量大小，构建双目组合模型。由于失真并不会均匀地分布在整幅图像上，且不同类型和程度的失真会引起图像的信息量和结构产生不同的变化，导致左、右视图不同区域的结构相似度发生变化。对此，在空间域和变换域分别提取输入的左、右视图

的自然场景统计特征，并在变换域提取合成独眼图的自然场景统计特征。将提取的特征输入到SVR中，训练从特征域到质量分数域的预测模型，从而建立SIQA客观质量评价模型。

3 特征的提取

3.1 空间域特征的提取

(1) 自然场景统计特征的提取

自然图像遵循特定的统计特性，其统计特性会随失真的存在而发生改变^[10]。现以LIVE 3D Phase I数据库^[11]中一幅立体图像的左视图和右视图，以及相应的合成独眼图为例如图2(a)–图2(c)，分别提取左视图、右视图和独眼图上亮度统计分布特征，并分别进行均值对比度归一化操作(Mean Sub-

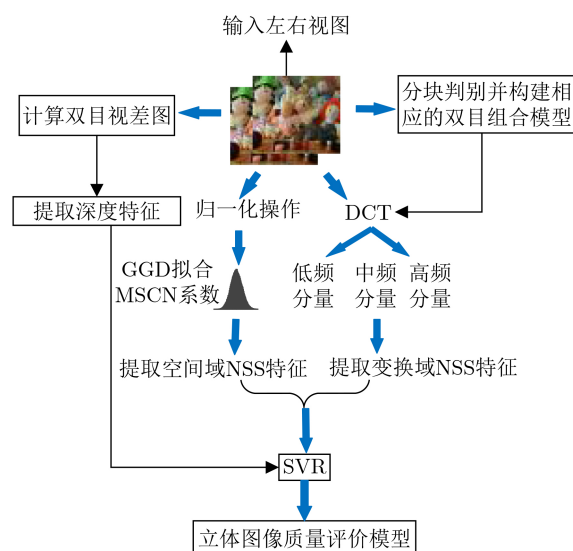


图1 SIQA模型的整体框架

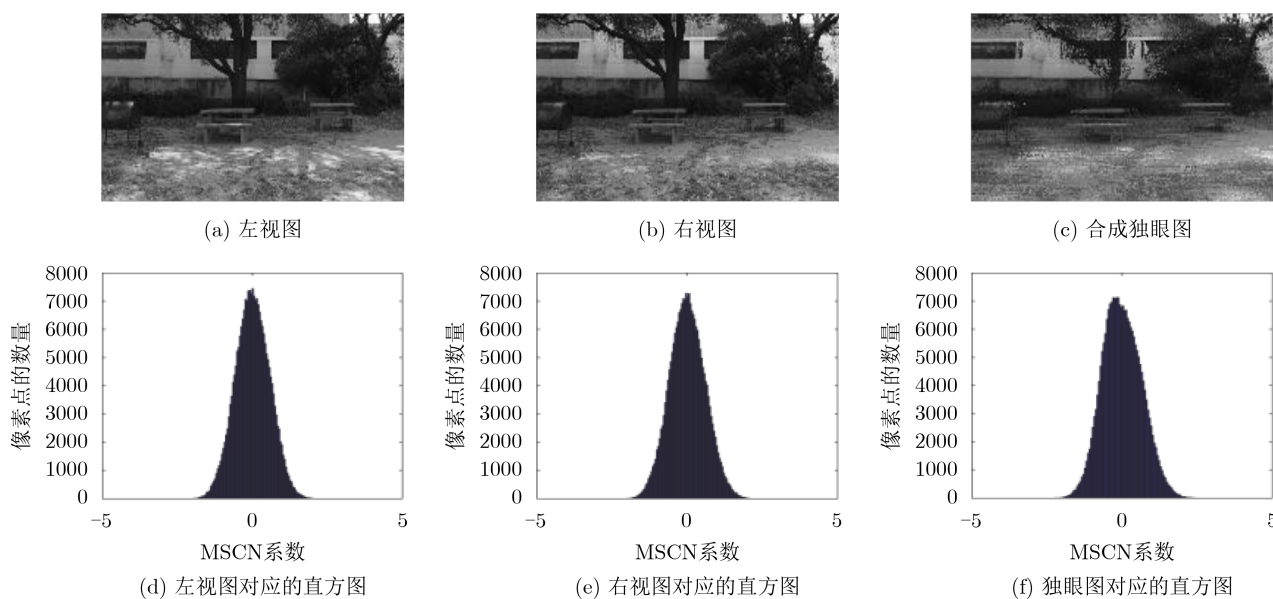


图2 左视图、右视图、合成独眼图及其相应的MSCN系数的统计分布直方图

tracted Contrast Normalized, MSCN)计算亮度分布统计特征,即计算图像的MSCN系数,以此得到其对应MSCN系数的统计分布直方图,如图2(d)–图2(f)。从图中可以看出自然图像的MSCN系数分布与广义高斯分布(Generalized Gaussian Distribution, GGD)具有一定的一致性。因此,可采用GGD模型来拟合左视图、右视图和独眼图的MSCN系数的统计分布,并将拟合形状的参数 α 和方差 v^2 作为第1组质量感知特征。

(2)深度特征的提取

由于包含更多信息量的视图将会更加吸引HVS的注意^[12]。对此,利用式(1)计算左视图的信息量,并用其对视差图进行加权以获得加权视差图 D_e 。由于失真会使视差图的统计特征发生变化,因此,从加权视差图上提取峰度 K 和偏斜度 S ,并将其作为深度特征 f_3 。其加权视差图 D_e 、偏斜度 S 和峰度 K 的计算分别如式(2)、式(3)和式(4)所示:

$$EL = 1 - \frac{1}{N} \cdot \sum_{k=1}^N \sum_{i=0}^L p(x_{ki,l}) \cdot \log_2 \left(\frac{1}{p(x_{ki,l})} \right) \quad (1)$$

其中,EL是左视图中包含的信息量; N 和 L 分别表示RGB颜色空间的通道数和最大像素值;下标 k 和 i 表示RGB颜色空间的第 k 个颜色通道和第 i 灰度级; $p(x_{ki,l})$ 表示左视图中的像素值 x 等于 i 时的概率。用同样的方法可以计算得到右视图中所包含的信息量ER。

$$D_e(i, j) = w_e(i, j) \times d(i, j) \quad (2)$$

其中, $w_e(i, j)$ 是 (i, j) 处视差值的权重,将式(1)应用到以 (i, j) 为中心的 $N \times N$ 大小的图像块上,计算图像块的信息量作为像素点 (i, j) 处的权重 $w_e(i, j)$; $d(i, j)$ 是 (i, j) 处的视差值。

$$S = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \frac{(D_e(i, j) - m)^3}{v^{\frac{3}{2}}} \quad (3)$$

$$K = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \frac{(D_e(i, j) - m)^4}{v^2} \quad (4)$$

其中, S 、 K 、 m 和 v 分别表示加权视差图的偏斜度、峰度、均值和方差; $M \times N$ 是加权视差图的尺寸。并将加权视差图的偏斜度、峰度和均值作为水平特征。

由于非对称失真导致的左右视图纵向信息不匹配现象也会影响HVS的深度感知,并导致双眼视觉不适,从而影响人眼对立体图像深度质量的判断^[13]。对此,本文计算左右视图之间纵向变化的线性相关程度,以此作为对水平深度信息的补充。根据式(5)和式(6)分别计算左右视图的视差补偿图。

$$I_{dl}(i, j) = I_l((i, j) + d_r(i, j)) \quad (5)$$

$$I_{dr}(i, j) = I_r((i, j) + d_l(i, j)) \quad (6)$$

其中, $d_r(x, y)$ 表示以左视图为主视图,从右视图上找到与左视图目标像素点相似度最大的对应点,计算两个目标点之间的水平距离; $d_l(x, y)$ 表示以右视图为主视图,从左视图上找到与右视图目标像素点相似度最大的对应点,即两个目标点之间的水平距离。

在上述基础上,对左、右视图及其对应的视差补偿图进行去均值归一化操作,并计算其差补偿图之间的纵向相关系数:

$$C = \text{corr} \{N[I_l(i, j)], N[I_{dl}(i, j)]\} \quad (7)$$

其中, $\text{corr}(\mathbf{X}, \mathbf{Y})$ 是MATLAB中的一个相关函数,用于计算两个矩阵 $\mathbf{X}$ 和 $\mathbf{Y}$ 列与列之间的成对相关系数; $N(\cdot)$ 是去均值对比度归一化操作。

现以图3(a)–图3(b)为LIVE 3D Phase II图像库^[14]中的一幅左、右视图为例,通过上述方法计算的左视图与右视差补偿图之间的纵向相关系数的统计分布图和右视图与左视差补偿图之间的纵向相关系数的统计分布图(如图3(c)、图3(d))。从图中看出,纵向相关系数的统计分布与非零均值非对称广义高斯分布(Asymmetric Generalized Gaussian Distribution, AGGD)具有一致性。

其非零均值非对称广义高斯分布AGGD模型计算:

$$f(x; \alpha, \beta_l, \beta_r, m) = \begin{cases} \frac{\alpha}{(\beta_l + \beta_r)\Gamma(1/\alpha)} \exp\left(-\left(\frac{x-m}{\beta_l}\right)^\alpha\right), & x < m \\ \frac{\alpha}{(\beta_l + \beta_r)\Gamma(1/\alpha)} \exp\left(-\left(\frac{x-m}{\beta_r}\right)^\alpha\right), & x \geq m \end{cases} \quad (8)$$

其中:

$$\beta_l = \sigma_l \sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)} \quad (9)$$

$$\beta_r = \sigma_r \sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)} \quad (10)$$

其中, α 是形状参数,控制AGGD的形状; β_l, β_r 控制两边的扩散程度; m 是均值,当 $m = 0$ 时,该分布就是零均值AGGD。

3.2 变换域特征的提取

失真类型分为信息加性失真和丢失性失真。信息加性失真和丢失性失真在整幅图像上的不均匀分布会导致失真图像不同区域的信息量和结构产生不同的变化^[12]。因此,本文将分块判别左右视图在HVS融合过程中可能发生的双目行为,并构建独眼图,然后在变换域提取反映图像失真程度的双目特征,从而使客观评分更加接近人眼主观评分。对此,提取离散余弦变换(Discrete Cosine Transform,

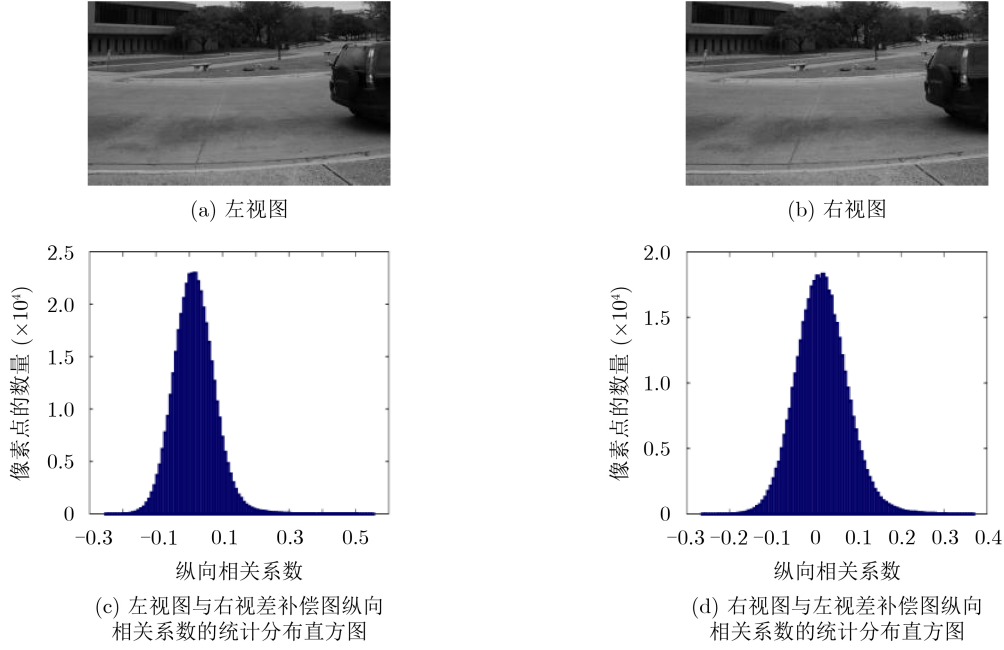


图3 左右视图及纵向相关系数图的统计分布直方图

DCT)域特征, 即形状参数 α 、频率变化系数 z 、不同频率子带的平均能量差异比。

首先, 将左、右视图分割为 $N \times N$ 大小的图像块, 分块比较左、右图像块的结构相似度及所含信息量大小, 构建相应的双目组合模型, 即独眼图。当立体图像的失真类型为非对称信息丢失失真 (Information-Loss Distortion, ILD) 时, 立体图像的感知质量受左、右视图中图像质量较高的那幅视图支配; 对于非对称高斯噪声之类的信息加性失真 (Information-Additive Distortion, IAD), 立体图像的感知质量受左、右视图中质量较低的视图支配^[2]。正如文献^[15,16]所述立体图像的感知质量由包含更多信息的视图主导, 因此选较多信息的视图作为双目组合过程中的主导视图。所提方法与文献^[17]相比, 不需要对失真类型进行分类, 从而降低了算法的复杂度, 同时避免了可能由分类引起的误差。

为了判别和模拟HVS中不同的双目行为, 需要构建不同的独眼视图来模拟双目视觉机制。对此采用结构相似度模型SSIM^[18]计算左右两幅视图图像块之间每对像素点的相似度值, 然后将所有像素点的相似度值求平均的方式来计算图像块之间的结构相似度 $S_{L,R}$, 并通过比较左视图和右视图包含的信息量来选择双目组合过程中的主导视图。通过式(1)计算熵来衡量左右视图图像块的信息量, 同时根据文献^[19]的双目刺激之间的相似/不相似程度来区分不同的双目行为。同时, 参考文献^[4]设置了结构相似度阈值 T_1 和信息量阈值 T_2 来构建不同的双目组合模型。

(1) 双目融合

当 $S_{L,R} \geq T_1$, 选择左视图作为主导视图:

$$CI_f(x, y) = W_L(x, y) \cdot I_L(x, y) + W_R(x+d, y) \cdot I_R(x+d, y) \quad (11)$$

(2) 双目抑制

当 $S_{L,R} < T_1$ 且 $EL - ER \geq T_2$ 时, 左视图主导立体图像的感知质量, 其双目抑制模型:

$$CI_{SL}(x, y) = W_L(x, y) \cdot I_L(x, y) + W_R(x+d, y) \cdot I_R(x+d, y) \quad (12)$$

当 $S_{L,R} < T_1$ 且 $ER - EL \geq T_2$ 时, 右视图主导立体图像的感知质量, 其双目抑制模型:

$$CI_{SR}(x, y) = W_L(x+d, y) \cdot I_L(x+d, y) + W_R(x, y) \cdot I_R(x, y) \quad (13)$$

(3) 双目竞争

当 $S_{L,R} < T_1$ 且 $EL - ER < T_2$ 时, 会出现左、右视图交替主导融合过程的现象。基于此, 双目竞争模型:

$$CI_R(x, y) = W_{SL}(x, y) \cdot CI_{SL}(x, y) + W_{SR}(x+d, y) \cdot CI_{SR}(x+d, y) \quad (14)$$

其中, W_L 和 W_R 分别为左、右视图权重, W_{SL} 和 W_{SR} 分别以左、右视图为主视图的权重, 其计算为

$$W_L(x, y) = \frac{E_L(x, y)}{E_L(x, y) + E_R(x+d, y)} \quad (15)$$

$$W_R(x+d, y) = \frac{E_R(x+d, y)}{E_L(x, y) + E_R(x+d, y)} \quad (16)$$

$$W_{SL}(x, y) = \frac{E_{SL}(x, y)}{E_{SL}(x, y) + E_{SR}(x+d, y)} \quad (17)$$

$$W_{SR}(x, y) = \frac{E_{SR}(x+d, y)}{E_{SL}(x, y) + E_{SR}(x+d, y)} \quad (18)$$

其中, $I_L(x, y)$ 和 $I_R(x, y)$ 分别表示左、右视图; $d(x, y)$ 是根据文献[14]计算的水平视差值; $E_L(x, y)$ 和 $E_R(x, y)$ 分别表示左右视图的2D-Gabor能量响应; $E_{SL}(x, y)$ 和 $E_{SR}(x, y)$ 分别表示以左、右视图为主视图的两幅抑制图的2D-Gabor能量响应。2D-Gabor滤波器定义:

$$g(x, y, \lambda, \theta) = \frac{1}{2\pi\sigma} \exp \left[-\frac{1}{2} \left(\frac{x^2 + y^2}{\sigma^2} \right) \right] \times \exp \left[i \left(\frac{2\pi x}{\lambda} \right) \right] \quad (19)$$

其中, $x = x_0 \cos q + y_0 \sin q$ 和 $y = -x_0 \sin q + y_0 \cos q$; (x_0, y_0) 是滤波器的中心点; λ 是波长, 控制滤波器的尺度; q 是方向角度; s 是椭圆高斯包络沿 x 轴和 y 轴的标准偏差。由于初级视觉皮层中的简单细胞和复杂细胞具有不同尺度的感受野, 因此, 使用具有5个频域尺度和4个方向的多尺度Gabor滤波器, $\sigma = 0.5\lambda$, $\lambda \in (1, \sqrt{2}, 2\sqrt{2}, 4)$, $\theta \in (1, \pi/4, \pi/2, 3\pi/4)$ 。

2D-Gabor能量响应定义:

$$GE_{\theta, \lambda}(x, y) = I(x, y) \otimes g(x, y, \lambda, \theta) \quad (20)$$

其中, $\otimes$ 表示卷积操作; $I(x, y)$ 是输入图像。

然后, 对合成独眼图图像块进行DCT, 并利用GGD模型拟合独眼图图像块的DCT系数, 提取拟合的形状参数 α 作为图像块特征。并通过平均组合的方式, 将所有图像块形状参数的平均值作为整幅独眼图的质量感知特征。令 X 表示输入图像的DCT系数, 则下一个DCT域特征频率变化系数 z 可以通过DCT系数幅度的标准差 $s|X|$ 和均值 $m|X|$ 之比计算得到:

$$z = \frac{s|X|}{m|X|} \quad (21)$$

提取独眼图图像块的频率变化系数 z 特征, 同样采用平均池化的方式将所有图像块频率变化系数的平均值作为整幅独眼图的第2个变换域特征。图4(a)~图4(d)分别给出了LIVE 3D Phase I 图像库[11]中一幅未失真、模糊失真的右视图以及经过DCT后的能量分布图。从图4中可以看出, 图像经过DCT后能量集中分布在左上角。经失真和未失真图像的能量分布图对比, 失真会使图像的能量分布发生变化, 为了度量这一变化, 对此定义了局部DCT能量子带比。

为了便于说明计算和量化不同频率子带的DCT系数, 以视图分割为 5×5 大小的图像块为例构成图5所示中 5×5 的矩阵, 并按颜色划分为3个不同频率的子带, DCT系数矩阵从左上角向右下角移动, 表示空间频率越来越高。令 n 表示不同的频带, 其中 $n = 1, 2, 3$ (低, 中, 高)。然后将GGD拟合频带 n 的模型方差 s_n^2 定义为对应于频带 n 的平均能量, 计算式为

$$E_n = s_n^2 \quad (22)$$

则不同频率子带的平均能量差异比 R_n ($n = 2, 3$) 的计算式:

$$R_n = \frac{\left| E_n - \frac{1}{n-1} \sum_{j < n} E_j \right|}{E_n + \frac{1}{n-1} \sum_{j < n} E_j} \quad (23)$$

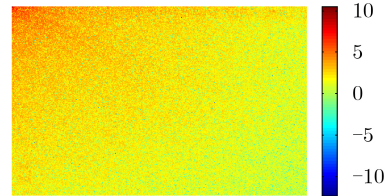
比率越大则表明频带 n 的频率能量与较低频率的频带中的平均能量之间的差异越大。特征 R_n 可用于测量受失真影响时较低和较高频段中能量的相对分布。计算 R_2 和 R_3 的平均值作为第3个DCT域特征。同样地, 计算所有图像块的该特征, 并将其均值作为整幅独眼图的第3个DCT域特征。



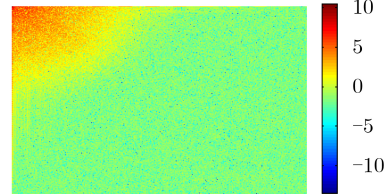
(a) 参考图像



(c) 模糊失真图像



(b) 参考图像对应的能量分布图



(d) 失真图像对应的能量分布图

图4 参考图像及失真图像经DCT后的能量分布图

DC	C_{11}	C_{12}	C_{14}	C_{15}
C_{21}	C_{22}	C_{23}	C_{24}	C_{25}
C_{31}	C_{32}	C_{33}	C_{34}	C_{35}
C_{41}	C_{42}	C_{43}	C_{44}	C_{45}
C_{51}	C_{52}	C_{53}	C_{54}	C_{55}

图5 不同频率子带的DCT系数

对于左、右视图，直接输入整幅图像，使用GGD拟合其DCT系数，并分别提取左、右视图相应的形状参数 α 、频率变化系数 z 和不同频率子带的平均能量差异比作为其DCT域特征。

3.3 多尺度特征的提取

自然图像具有多尺度特性^[20,21]，提取多尺度特征有利于处理图像分辨率的变化以及从图像显示器到观察者的距离或观察者视觉系统的敏锐度变化，从而提高模型性能。对此，对输入图像进行两次下采样，所建立的SIQA模型在3个尺度上基于NSS场景进行空间域和变换域特征提取，并在单尺度上提取立体图像的深度特征，从而为捕捉不同分辨率下图像失真度的变化提供依据。

4 SIQA模型的验证

4.1 立体图像质量评价模型的构建

首先，对输入图像经过两次下采样，从3个尺度上分别提取左、右视图的 $2 \times 3 = 6$ 维空间域NSS特征， $3 \times 3 = 9$ 维DCT域特征，即从左、右视图上分别获得15维单目特征；在3个尺度上提取独眼图 $3 \times 3 = 9$ 维DCT域特征。另外，从加权视差图上提取3维水平深度特征，相关系数图上提取4维纵向深度特征，即共提取46维质量感知特征。

在训练阶段，首先提取训练集的质量感知特征，然后利用支持向量回归SVR训练从特征向量到主观评分的预测模型。在测试阶段，只需将提取到待测图像的特征向量输入到训练好的预测模型中即可得到待测图像的质量分数。图像库中的主观评分是以DMOS(Differential Mean Opinion Score)或MOS(Mean Opinion Score)的形式呈现，其中，DMOS值越大，表示图像质量越差，而MOS值越大，表示图像质量越好。

4.2 实验与分析

本文所建立模型分别在LIVE 3D Phase I数据库^[11]、LIVE 3D Phase II数据库^[14]、Waterloo-IVC 3D Phase I数据库^[3]以及Waterloo-IVC 3D Phase II数据库^[22]中进行测试。

4.2.1 算法性能评估

本文采用Spearman秩相关系数(Spearman's RankOrdered Correlation Coefficient, SROCC)、皮尔逊线性相关系数(Pearson Linear Correlation Coefficient, PLCC)以及均方根误差(Root Mean Squared Error, RMSE)来验证所提方法的性能。SROCC和PLCC分别评估了模型预测分数的单调性以及预测分数和客观评分的一致性；RMSE反映了预测分数和客观评分的误差。PLCC和SROCC值越接近1，且RMSE值越小，则表示主观一致性越高，模型的预测性能越好，即模型预测分数和人眼主观评分更加接近。实验中，训练集选80%图像库的图像和测试集选另外20%的图像。为避免偏差，将随机训练-测试重复1000次，取中值作为结果。

4.2.2 一致性实验

所提模型在4个立体数据库中预测图像的质量分数和主观评分的散点图如图6所示。从图6中可以看出，所提模型预测分数与人眼主观评分的散点图比较聚集且非常接近中心线，特别是图6(a)、图6(b)和图6(d)中的所有散点非常接近中心线，表明所提算法具有良好的预测性能，与人眼主观评分具有较高的一致性。

4.2.3 模型性能评估

(1)不同图像库中的测试

选用LIVE 3D Phase I和LIVE 3D Phase II两个立体图像库对所提算法针对每种失真类型立体图像质量预测的准确性进行测试，其中图像库包含JPEG2000(JP2K)、JPEG压缩失真，高斯白噪声(Gaussian White Noise, WN)，高斯模糊(Gaussian blur, Gblur)和快速衰落(Fast Fading, FF)等5种失真类型。测试结果列于表1中(括号中为LIVE 3D Phase II数据)，分析可知所提算法与人眼主观观测结果具有较好的一致性。

(2)对比算法的测试

对本文算法的整体性能评估。将本文所提出的方法与一些主流的SIQA方法进行比较，如全参考Lin^[29]、Khan^[23]、Chen FR^[14]、Jiang^[24]和SSIM^[18]；半参考Ma^[25]；无参考SINQ^[8]、Zhou^[8]、Karimi^[26]、Yang-SAE^[27]、Fezza^[17]、BRISQUE^[28]和全参考Lin^[29]等。表2给出了所提算法与一些主流的SIQA算法在公开立体图像库中整体性能比较的结果。与对比算法相比，所提模型的整体性能在4个SIQA数据库中表现优异。由于所提SIQA模型考虑到非均匀分布失真对图像质量评价的影响，同时结合空间域和变换域提取质量感知特征，从而提高了模型性能。

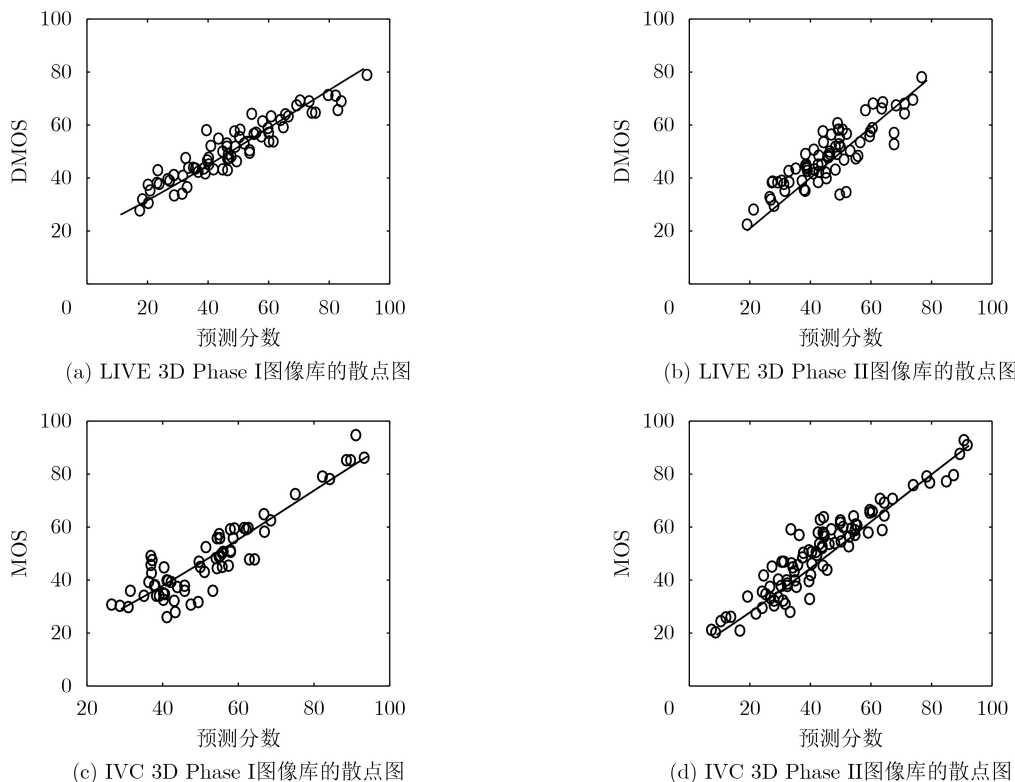


图6 所提模型在4个数据库中图像预测值和主观评分的散点图

4.2.4 泛化性能测试

本文分别将LIVE 3D Phase I和Phase II数据库用作训练集和测试集, LIVE 3D Phase I/Phase II表示所提模型在LIVE 3D Phase I图像库中训练,然后在LIVE 3D Phase II图像库中进行测试; LIVE 3D Phase Ii/Phase I表示所提模型在LIVE 3D Phase II图像库中训练,然后在LIVE 3D Phase I图像库中进行测试。

表1 LIVE 3D Phase I和II图像库中的性能测试

失真类型	PLCC	SROCC	RMSE
WN	0.964(0.988)	0.958(0.964)	6.004(4.954)
JP2K	0.924 (0.927)	0.907(0.915)	6.518(5.361)
JPEG	0.730 (0.903)	0.785(0.900)	7.010(5.337)
Gblur	0.972 (0.979)	0.937(0.962)	5.412(3.478)
FF	0.943 (0.945)	0.913(0.921)	5.433(3.679)
All	0.967 (0.949)	0.946(0.935)	5.603(3.501)

表2 LIVE 3D Phase I和II图像库中整体性能比较

对比算法		LIVE 3D Phase I			LIVE 3D Phase II		
		PLCC	SROCC	RMSE	PLCCS	ROCCR	MSE
Lin ^[29]	FR	0.936	0.931	5.744	0.911	0.893	4.647
Khan ^[23]	FR	0.927	0.916	—	0.932	0.922	—
Chen FR ^[14]	FR	0.833	0.915	6.268	0.770	0.888	4.892
Jiang ^[24]	FR	0.945	0.933	5.276	0.916	0.903	4.523
Ma ^[25]	RR	0.930	0.929	6.024	0.921	0.917	4.390
SINQ ^[8]	NR	0.955	0.935	4.781	0.936	0.931	3.959
Zhou ^[8]	NR	0.941	0.921	5.540	0.923	0.919	4.262
Karimi ^[26]	NR	0.956	0.940	4.998	0.923	0.913	4.436
Yang-SAE ^[27]	NR	0.961	0.949	—	0.938	0.928	—
Fezza ^[13]	NR	—	—	—	0.925	0.908	3.018
BRISQUE ^[28]	NR	0.910	0.901	6.793	0.782	0.770	7.038
本文算法	NR	0.967	0.946	5.603	0.949	0.935	3.501

实验中, 对比算法为: Yang-SAE^[27]、BRISQUE^[28]、CNN-based^[30], 并利用PLCC和SROCC指标评估所提模型与对比算法的泛化性能。表3给出了本文算法与一些SIQA算法跨库实验的对比结果, 与其

他SIQA算法相比, 所提模型在LIVE 3D Phase I和Phase II图像库上测试的PLCC值和SROCC值基本达到最大值, 表明本文算法训练的预测模型针对不同的测试数据库具有良好的通用性。

表 3 跨库性能对比实验

方法	LIVE 3D Phase I/Phase II		LIVE 3D Phase II/Phase I	
	PLCC	SROCC	PLCC	SROCC
Yang-SAE ^[27]	0.861	0.846	0.860	0.845
BRISQUE ^[28]	0.595	0.458	0.572	0.556
CNN-based ^[30]	0.227	0.208	0.743	0.741
本文算法	0.853	0.849	0.883	0.868

5 总结

针对失真并不会均匀地分布在整幅视图上的问题, 将输入的左、右视图分块, 然后分块判别并构建模拟左右视图在大脑中融合过程的双目组合模型, 并在变换域提取双目特征。最后将空间域和变换域特征同时输入到SVR, 训练从特征域到质量分数域的立体图像质量评价模型。并在4个公开的立体图像数据库上与一些主流的全参考、半参考和无参考立体图像质量评价算法进行对比, 验证了所提算法具有良好的性能。

通过实验验证了所提建立SIQA图像质量评价模型与人眼主观视觉具有高度一致性, 表明所提方法具有一定的参考价值。

参 考 文 献

- [1] 高新波, 路文, 查林, 等. 超高清视频画质提升技术及其芯片化方案[J]. 重庆邮电大学学报:自然科学版, 2020, 32(5): 681–697. doi: [10.3979/j.issn.1673-825X.2020.05.001](https://doi.org/10.3979/j.issn.1673-825X.2020.05.001).
GAO Xinbo, LU Wen, ZHA Lin, et al. Quality elevation technique for UHD video and its VLSI solution[J]. *Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition*, 2020, 32(5): 681–697. doi: [10.3979/j.issn.1673-825X.2020.05.001](https://doi.org/10.3979/j.issn.1673-825X.2020.05.001).
- [2] 张敏辉, 杨剑. 评价SAR图像去噪效果的无参考图像质量指标[J]. 重庆邮电大学学报:自然科学版, 2018, 30(4): 530–536. doi: [10.3979/j.issn.1673-825X.2018.04.014](https://doi.org/10.3979/j.issn.1673-825X.2018.04.014).
ZHANG Minhui and YANG Jian. A new referenceless image quality index to evaluate denoising performance of SAR images[J]. *Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition*, 2018, 30(4): 530–536. doi: [10.3979/j.issn.1673-825X.2018.04.014](https://doi.org/10.3979/j.issn.1673-825X.2018.04.014).
- [3] 徐弦秋, 刘宏清, 黎勇, 等. 基于RGB通道下模糊核估计的图像去模糊[J]. 重庆邮电大学学报:自然科学版, 2018, 30(2): 216–221. doi: [10.3979/j.issn.1673-825X.2018.02.009](https://doi.org/10.3979/j.issn.1673-825X.2018.02.009).
XU Xianqiu, LIU Hongqing, LI Yong, et al. Image deblurring with blur kernel estimation in RGB channels[J].

Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2018, 30(2): 216–221. doi: [10.3979/j.issn.1673-825X.2018.02.009](https://doi.org/10.3979/j.issn.1673-825X.2018.02.009).

- [4] CHEN Yong, ZHU Kaixin, and LIU Huanlin. Blind stereo image quality assessment based on binocular visual characteristics and depth perception[J]. *IEEE Access*, 2020, 8: 85760–85771. doi: [10.1109/ACCESS.2020.2992746](https://doi.org/10.1109/ACCESS.2020.2992746).
- [5] DING Jian and SPERLING G. A gain-control theory of binocular combination[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(4): 1141–1146. doi: [10.1073/pnas.0509629103](https://doi.org/10.1073/pnas.0509629103).
- [6] DING Yong and ZHAO Yang. No-reference stereoscopic image quality assessment guided by visual hierarchical structure and binocular effects[J]. *Applied Optics*, 2018, 57(10): 2610–2621. doi: [10.1364/AO.57.002610](https://doi.org/10.1364/AO.57.002610).
- [7] HACHICHA W, KAANICHE M, BEGHDAI A, et al. No-reference stereo image quality assessment based on joint wavelet decomposition and statistical models[J]. *Signal Processing: Image Communication*, 2017, 54: 107–117. doi: [10.1016/j.image.2017.03.005](https://doi.org/10.1016/j.image.2017.03.005).
- [8] LIU Lixiong, LIU Bao, SU Chechun, et al. Binocular spatial activity and reverse saliency driven no-reference stereopair quality assessment[J]. *Signal Processing: Image Communication*, 2017, 58: 287–299. doi: [10.1016/j.image.2017.08.011](https://doi.org/10.1016/j.image.2017.08.011).
- [9] DAKIN S C and BEX P J. Natural image statistics mediate brightness ‘filling in’[J]. *Proceedings of the Royal Society B: Biological Sciences*, 2003, 270(1531): 2341–2348. doi: [10.1098/rspb.2003.2528](https://doi.org/10.1098/rspb.2003.2528).
- [10] 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, 38(7): 1645–1653. doi: [10.11999/JEIT151058](https://doi.org/10.11999/JEIT151058).
CHEN Yong, SHUAI Feng, and FAN Qiang. A no-reference image quality assessment based on distribution characteristics of natural statistics[J]. *Journal of Electronics & Information Technology*, 2016, 38(7): 1645–1653. doi: [10.11999/JEIT151058](https://doi.org/10.11999/JEIT151058).
- [11] MOORTHY A K, SU Chechun, MITTAL A, et al.

- Subjective evaluation of stereoscopic image quality[J]. *Signal Processing:Image Communication*, 2013, 28(8): 870–883. doi: [10.1016/j.image.2012.08.004](https://doi.org/10.1016/j.image.2012.08.004).
- [12] RYU S and SOHN K. No-reference quality assessment for stereoscopic images based on binocular quality perception[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, 24(4): 591–602. doi: [10.1109/TCSVT.2013.2279971](https://doi.org/10.1109/TCSVT.2013.2279971).
- [13] ZHOU Jun, WANG Ling, YIN Haibing, *et al.* Eye movements and visual discomfort when viewing stereoscopic 3D content[J]. *Digital Signal Processing*, 2019, 91: 41–53. doi: [10.1016/j.dsp.2018.12.008](https://doi.org/10.1016/j.dsp.2018.12.008).
- [14] CHEN Mingjun, SU Chechun, KWON D K, *et al.* Full-reference quality assessment of stereopairs accounting for rivalry[J]. *Signal Processing:Image Communication*, 2013, 28(9): 1143–1155. doi: [10.1016/j.image.2013.05.006](https://doi.org/10.1016/j.image.2013.05.006).
- [15] MEEGAN D V, STELMACH L B, and TAM W J. Unequal weighting of monocular inputs in binocular combination: Implications for the compression of stereoscopic imagery[J]. *Journal of Experimental Psychology:Applied*, 2001, 7(2): 143–153. doi: [10.1037/1076-898X.7.2.143](https://doi.org/10.1037/1076-898X.7.2.143).
- [16] SMITH III E L, FERN K, MANNY R, *et al.* Interocular suppression produced by rivalry stimuli: A comparison of normal and abnormal binocular vision[J]. *Optometry and Vision Science:Official Publication of the American Academy of Optometry*, 1994, 71(8): 479–491. doi: [10.1097/00006324-199408000-00001](https://doi.org/10.1097/00006324-199408000-00001).
- [17] FEZZA S A, CHETOUANI A, and LARABI M C. Using distortion and asymmetry determination for blind stereoscopic image quality assessment strategy[J]. *Journal of Visual Communication and Image Representation*, 2017, 49: 115–128. doi: [10.1016/j.jvcir.2017.08.009](https://doi.org/10.1016/j.jvcir.2017.08.009).
- [18] WANG Zhou, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612. doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [19] LAGO-FERNÁNDEZ L F and DECO G. A model of binocular rivalry based on competition in IT[J]. *Neurocomputing*, 2002, 44/46: 503–507. doi: [10.1016/S0925-2312\(02\)00408-3](https://doi.org/10.1016/S0925-2312(02)00408-3).
- [20] SHEIKH H R and BOVIK A C. Image information and visual quality[J]. *IEEE Transactions on Image Processing*, 2006, 15(2): 430–444. doi: [10.1109/TIP.2005.859378](https://doi.org/10.1109/TIP.2005.859378).
- [21] GEISLER W S. Visual perception and the statistical properties of natural scenes[J]. *Annual Review of Psychology*, 2008, 59: 167–192. doi: [10.1146/annurev.psych.58.110405.085632](https://doi.org/10.1146/annurev.psych.58.110405.085632).
- [22] WANG Jiheng, ZENG Kai, and WANG Zhou. Quality prediction of asymmetrically distorted stereoscopic images from single views[C]. 2014 IEEE International Conference on Multimedia and Expo, Chengdu, China, 2014: 1–6.
- [23] KHAN S and CHANNAPPAYYA S S. Estimating depth-salient edges and its application to stereoscopic image quality assessment[J]. *IEEE Transactions on Image Processing*, 2018, 27(12): 5892–5903. doi: [10.1109/TIP.2018.2860279](https://doi.org/10.1109/TIP.2018.2860279).
- [24] JIANG Gangyi, XU Haiyong, YU Mei, *et al.* Stereoscopic image quality assessment by learning non-negative matrix factorization-based color visual characteristics and considering binocular interactions[J]. *Journal of Visual Communication and Image Representation*, 2017, 46: 269–279. doi: [10.1016/j.jvcir.2017.04.010](https://doi.org/10.1016/j.jvcir.2017.04.010).
- [25] MA Jian, AN Ping, and SHEN Liquan. Reduced-reference stereoscopic image quality assessment using natural scene statistics and structural degradation[J]. *IEEE Access*, 2018, 6: 2768–2780. doi: [10.1109/ACCESS.2017.2785282](https://doi.org/10.1109/ACCESS.2017.2785282).
- [26] KARIMI M, SOLTANIAN N, SAMAVI S, *et al.* Blind stereo image quality assessment inspired by brain sensory-motor fusion[J]. *Digital Signal Processing*, 2019, 91: 91–104. doi: [10.1016/j.dsp.2019.03.004](https://doi.org/10.1016/j.dsp.2019.03.004).
- [27] YANG Jiache, SIM K, LU Wen, *et al.* Predicting stereoscopic image quality via stacked auto-encoders based on stereopsis formation[J]. *IEEE Transactions on Multimedia*, 2019, 21(7): 1750–1761. doi: [10.1109/TMM.2018.2889562](https://doi.org/10.1109/TMM.2018.2889562).
- [28] MITTAL A, MOORTHY A K, and BOVIK A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695–4708. doi: [10.1109/TIP.2012.2214050](https://doi.org/10.1109/TIP.2012.2214050).
- [29] LIN Yancong, YANG Jiachen, LU Wen, *et al.* Quality index for stereoscopic images by jointly evaluating cyclopean amplitude and cyclopean phase[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2017, 11(1): 89–101. doi: [10.1109/JSTSP.2016.2632422](https://doi.org/10.1109/JSTSP.2016.2632422).
- [30] ZHANG Wei, QU Chenfei, MA Lin, *et al.* Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network[J]. *Pattern Recognition*, 2016, 59: 176–187. doi: [10.1016/j.patcog.2016.01.034](https://doi.org/10.1016/j.patcog.2016.01.034).
- 陈 勇: 男, 1963年生, 博士, 教授, 主要从事图像处理。
- 金曼莉: 女, 1997年生, 硕士生, 主要从事无参考图像质量评价。
- 朱凯欣: 女, 1994年生, 硕士, 主要从事立体图像质量评价。
- 刘焕淋: 女, 1970年生, 博士生导师, 教授, 主要从事信号处理等方面的研究。
- 陈 东: 男, 1996年生, 硕士, 主要从事无参考图像质量评价与图像增强。