

分类号: TP317.4

密 级: 公开

学校代码: 10363

学 号: 2220130103



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于多模式特征组合的无参考
图像质量评价方法研究

论 文 作 者 何欢

指 导 教 师 江本赤 教授 史晨阳 讲师

学 科 (专 业) 人工智能

研 究 方 向 计算机视觉

论文提交日期: 2025 年 5 月 20 日

分类号：TP317.4

单位代码：10363

密 级：公 开

学 号：2220130103

基于多模式特征组合的无参考图像质量评价方法研究

**Research on No-Reference Image Quality Assessment
Method Based on Multi-Mode Feature Combination**

学生姓名：_____何欢_____

指导教师：江本赤 教授 史晨阳 讲师

专 业：_____人工智能_____

研究方向：_____计算机视觉_____

论文答辩日期：_____2025 年 5 月 9 日_____

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：何欢

日期：2025 年 5 月 20 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：何欢

指导教师签名：江萍

日期：2025年5月20日

日期：2025年5月20日

基于多模式特征组合的无参考图像质量评价方法研究

摘 要

随着数字成像技术的发展与应用，图像质量评价（Image Quality Assessment, IQA）已成为图像处理领域的重要研究课题。依赖主观实验的传统评价方式存在效率低下、操作性差等局限性，而客观图像质量评价技术则具备效率高、适应性强等优势。现有方法未充分利用图像的色差及信息熵等特征，因此开展基于多模式特征组合的图像质量评价相关研究具有重要意义。本文聚焦无参考客观图像质量评价问题，通过对图像中有价值的特征信息进行挖掘及组合，结合机器学习方法，建立量化评价模型，以实现对图像质量的准确客观评价。论文的主要工作如下：

（1）为充分利用模糊图像的彩色信息，提高无参考图像质量评价的准确性，研究了一种基于感兴趣区域的无参考图像质量评价算法。首先，分析图像像素灰度值和颜色值分布的不确定性，以及两个相邻像素之间的差异性，在此基础上进行失真图像的颜色空间转换；然后提取并组合三通道的信息熵和色差信息特征，将图像中信息量较大的区域作为感兴趣区域；最后针对选定区域，深入分析色差特征，并对其进行统计学分析，从而获得最终的图像质量评价模型。实验验证了所提方法较高的预测精度，改善了对不同模糊失真图像评价的准确性。

（2）针对传统统计学方法在无参考图像质量评价中的局限性，研究了一种基于粒子群优化算法（PSO）与广义回归神经网络

(GRNN) 相结合的无参考图像质量评价框架 (PSO-GRNN)。首先, 通过提取模糊图像的显著性、色差和梯度特征关键信息并获取关键特征分数; 然后依据特征分数和主观分数标签将数据库划分为训练集和测试集, 用于模型的训练与验证; 再采用粒子群优化算法对 GRNN 的网络参数进行优化, 基于 PSO-GRNN 优化模型, 通过输入特征分数预测图像质量评分。在多失真和自然失真图像数据库上对所提模型进行了验证, 实验结果表明, 基于 PSO-GRNN 模型的图像质量评价准确度高于传统方法。

关键词: 图像质量评价, 感兴趣区域, 信息熵, 色差, 特征组合, PSO-GRNN

RESEARCH ON NO-REFERENCE IMAGE QUALITY ASSESSMENT METHOD BASED ON MULTI-MODE FEATURE COMBINATION

ABSTRACT

With the development and application of digital imaging technology, image quality assessment (IQA) has become an important research topic in the field of image processing. The traditional evaluation method relying on subjective experiments has the limitations of low efficiency and poor operability, while the objective image quality evaluation technology has the advantages of high efficiency and strong adaptability. The existing methods do not make full use of the characteristics of image color difference and information entropy. Therefore, it is of great significance to carry out research on image quality evaluation based on multi-mode feature combination. This paper focuses on the problem of no-reference objective image quality evaluation. By mining and combining valuable feature information of images, combined with machine learning methods, a quantitative evaluation model is established to achieve accurate and objective evaluation of image quality. The main work of this paper is as follows:

- (1) In order to make full use of the color information of blurred images

and improve the accuracy of no-reference image quality assessment, a no-reference image quality assessment algorithm based on region of interest is studied. Firstly, the uncertainty of the gray value and color value distribution of the image pixels and the difference between two adjacent pixels are analyzed. On this basis, the color space conversion of the distorted image is carried out. Then, the information entropy and color difference information features of the three channels are extracted and combined, and the region with large amount of information in the image is taken as the region of interest. Finally, for the selected area, the color difference characteristics are analyzed in depth, and the statistical analysis is carried out to obtain the final image quality evaluation model. Experiments verify that the proposed method has higher prediction accuracy and improves the accuracy of evaluation of different blurred distorted images.

(2) Aiming at the limitations of traditional statistical methods in no-reference image quality assessment, a no-reference image quality assessment framework (PSO-GRNN) based on particle swarm optimization (PSO) and generalized regression neural network (GRNN) is studied. Firstly, the key information of saliency, color difference and gradient features of blurred images is extracted and the key feature scores are obtained. Then, the database is divided into training set and test set according to the feature score and subjective score label for training and

verification of the model. Then, the particle swarm optimization algorithm is used to optimize the network parameters of GRNN. Based on the PSO-GRNN optimization model, the image quality score is predicted by inputting the feature score. The proposed model is verified on the multi-distortion and natural distortion image database. The experimental results show that the accuracy of image quality evaluation based on PSO-GRNN model is higher than that of traditional methods.

KEY WORDS: image quality evaluation, area of interest, information entropy, color difference, feature combination, PSO-GRNN

目录

第 1 章 绪论.....	1
1.1 选题背景及意义.....	1
1.2 图像质量评价研究现状.....	2
1.2.1 主观质量评价	3
1.2.2 客观质量评价	4
1.3 本文研究内容与结构安排.....	8
1.3.1 主要研究内容	8
1.3.2 结构安排	9
第 2 章 图像质量评价理论基础.....	11
2.1 人眼视觉系统.....	11
2.1.1 人眼生理结构	11
2.1.2 人眼视觉感知特性	12
2.2 图像信息表征.....	13
2.2.1 信息熵	14
2.2.2 梯度	14
2.2.3 色差	15
2.2.4 显著性	16
2.3 图像质量评价常用数据库.....	17
2.3.1 多失真图像数据库	17
2.3.2 自然失真图像数据库.....	19
2.4 算法性能验证指标.....	21
2.5 本章小结	23
第 3 章 基于感兴趣区域的无参考图像质量评价	24
3.1 特征提取与组合.....	24
3.2 基于感兴趣区域的评价模型构建.....	26
3.2.1 图像感兴趣区域选择.....	26
3.2.2 基于感兴趣区域的质量评价模型.....	29
3.3 实验验证	30
3.3.1 颜色空间测试结果分析.....	32
3.3.2 AOI 尺寸选择的效果分析.....	32
3.3.3 实验结果与分析	33
3.3.4 CLIVE 数据库测试结果.....	36
3.4 本章小结	37
第 4 章 基于 PSO-GRNN 算法的图像质量评价.....	38
4.1 算法模型构建.....	38
4.1.1 广义回归神经网络 (GRNN)	38
4.1.2 PSO 算法原理	39
4.1.3 PSO-GRNN 图像质量评价模型优化.....	40
4.2 特征图提取	41
4.2.1 梯度特征提取	41
4.2.2 色差特征提取	42
4.2.3 视觉显著性特征提取.....	42

4.3 模型性能对比实验.....	43
4.3.1 特征提取与组合	43
4.3.2 算法优化效果验证	45
4.4 实验结果与分析.....	46
4.4.1 多场景验证分析	46
4.4.2 人工失真数据库实验结果.....	48
4.4.3 自然失真模糊图像质量评价算法性能分析.....	49
4.5 本章小结	51
第 5 章 总结与展望.....	52
5.1 工作总结	52
5.2 不足之处与展望.....	52
参考文献.....	54
攻读学位期间参与的科研项目与学术成果.....	60
致谢	61

插图清单

图 1-1 清晰图像和失真图像质量对比	1
图 1-2 图像质量评价方法分类	3
图 1-3 基于学习的方法的一般流程	7
图 2-1 人眼生理结构图	12
图 2-2 测试图像及其视觉显著区域	13
图 2-3 TID2008 数据库参考图像	18
图 2-4 BID 数据库中四张模糊图	20
图 3-1 不同模糊程度色差图, (a)到(d)模糊程度递增	25
图 3-2 原始图像以及三个不同区域图像	27
图 3-3 选择感兴趣区域的流程	28
图 3-4 原始图像和感兴趣区域的直方图	28
图 3-5 不同 α 值在 CSIQ 数据库测试的 PLCC 和 SROCC 值变化趋势	30
图 3-6 基于感兴趣区域的图像质量评价整体流程	31
图 3-7 不同模糊程度失真典型图像及其主观分数和所提方法计算结果	31
图 3-8 所提算法计算结果与四个人工失真数据库中主观评分的拟合曲线	36
图 4-1 GRNN 结构图	38
图 4-2 PSO-GRNN 算法的总体框架	40
图 4-3 相同内容的模糊图像及其对应的梯度伪色图	41
图 4-4 相同内容的图像对应的色差伪色图	42
图 4-5 相同内容的图像及其对应的显著性图伪色图	42
图 4-6 提取图像特征值的流程	43
图 4-7 CID2013 不同场景的箱线图	48
图 4-8 BID、CID2013 和 CLIVE 数据库的散点图和拟合曲线	51

插表清单

表 2-1 多失真数据库信息说明	17
表 2-2 自然失真数据库信息说明	20
表 3-1 多失真数据库模糊图像信息说明	32
表 3-2 不同颜色空间测试结果	32
表 3-3 TID2008 数据库不同尺寸选择感兴趣区域测试结果	33
表 3-4 所提算法在四个人工失真数据库上的测试结果.....	33
表 3-5 所提算法和一些经典 NR-IQA 算法的比较.....	34
表 3-6 所提算法与近几年 NR-IQA 算法的比较.....	35
表 3-7 所提算法在 CLIVE 数据库上的计算结果.....	37
表 4-1 不同的特征组合的性能	44
表 4-2 PSO-GRNN 使用不同的自适应优化方法在 BID 数据库上的性能.....	45
表 4-3 CID2013 数据库不同场景介绍.....	46
表 4-4 在 CID2013 上的不同场景下的测试结果.....	47
表 4-5 PSO-GRNN 与近五年 NR-IQA 算法在人工失真数据库上的比较.....	49
表 4-6 PSO-GRNN 与其他 NR-IQA 算法在自然失真数据库上的比较.....	50

第 1 章 绪论

1.1 选题背景及意义

随着数字媒体技术的迅猛发展，相关产业得到了显著推动，数字图像呈现出爆炸式增长态势。在信息获取与传播等过程中，数字图像扮演着极为重要的角色，成为电视、互联网、虚拟现实系统等众多领域承载的关键内容^[1]。图像已经成为当今社会不可或缺的信息载体，在社会需求不断提升以及物质极大丰富的背景下，图像质量对用户体验和信息传递的影响力日益凸显。用户对图像的要求已从传统的视觉美观性向信息清晰度转变，对高质量图像的需求更加强烈。



图 1-1 清晰图像和失真图像质量对比

然而，在图像的采集、压缩、传输、处理及显示等众多的环节中^[2]，由于各种失真，图像不可避免会受到影响，导致其质量下降，图 1-1 展示了图像数据库中一幅清晰图像和失真图像。不同类型的图像失真情况，在降低用户视觉体验效果的同时都可能造成图像中重要信息的丢失。例如，在对图像进行处理时，压缩、去噪、去模糊、图像增强等操作可能会引起图像颜色信息、纹理信息和结构信息等方面的失真，不仅如此，还有许多复杂情况也会导致图像质量的下降，其中模糊就是常见的一种失真，并且该失真不仅会影响人眼对图像的视觉感知，还会导致图像重要信息缺失。因此，设计准确可靠的图像质量评价（Image Quality Assessment, IQA）方法，精准评估数字图像的模糊程度，已成为当下图像处理领域备受关注的研究方向，图像质量评估算法是通过模拟人眼视觉机制，代替人眼去进行图像质量评估，算法计算出来的分数其有效性是通过与数据库里每个测试图像的主观评分拟合进行验证。图像质量评价方法的应用，极大地提升了人们在

生产和生活中的便利性。

IQA 技术通过对图像的视觉效果、信息完整性和结构保真度进行量化分析，为图像采集、传输、压缩等环节提供了客观的评估依据。IQA 技术广泛应用于医疗诊断、遥感监测、工业检测等领域^[3]，已成为保障图像数据可靠性和可用性的关键技术。在医学影像领域，IQA 技术可用于病灶识别的辅助诊断，通过对医学图像的质量评估，提高病灶检测的准确性，从而为临床诊断提供更可靠的依据。在卫星遥感领域，IQA 技术能够对地物分类进行优化，通过对卫星图像质量的评估，确保地物信息的准确提取，为环境监测和资源管理提供支持。在视频监控领域，IQA 技术可实现对监控视频质量的实时优化，提升监控系统的可靠性和有效性。在图像处理领域，IQA 技术能够优化智能手机的拍照算法，提升图像质量，从而改善用户体验。随着人工智能和计算机视觉技术的快速发展，IQA 技术正从单一指标分析向多维度、智能化评估转变，这一转变不仅提升了图像质量评价的精准度，还为各行业的数字化转型提供了有力的技术支撑。

由于人类视觉系统十分复杂，导致人们对视觉认知存在局限，目前关于客观质量评价算法的研究还有很大挑战，模糊是图像失真中一种常见类型，会导致图像重要信息的丢失。因此，针对模糊图像客观评价问题，研究与人类视觉感知一致的客观方法成为了一个重要课题，具有重要的学术价值，并且对多媒体领域具有重要意义。

1.2 图像质量评价研究现状

目前，按照评价主体的不同，图像质量评价方法可分为主观质量评价和客观质量评价两种^[4]。而在客观质量评价方法中，根据所需参考图像信息量的多少，客观图像质量评价方法分为三类：全参考图像质量评价（Full-Reference Image Quality Assessment, FR-IQA）、半参考图像质量评价（Reduced-Reference Image Quality Assessment, RR-IQA）和无参考图像质量评价（No-Reference Image Quality Assessment, NR-IQA）^[5]，具体分类如图 1-2 所示。主观质量评价是以人类视觉感知为基础，通过观察人员的评分完成图像质量的表征，其结果被认为最能反应图像实际视觉效果。但由于主观评价实验耗费大量时间和人力，导致其应用性较

差，在具有较高实时性要求的图像处理系统中难以发挥作用。在实际应用中，客观质量评价方法时间效率高，稳定性好，具有更为广泛的时间价值和应用前景。

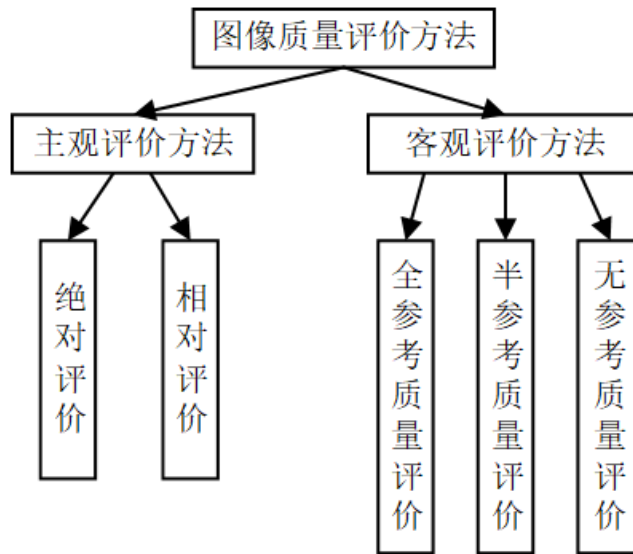


图 1-2 图像质量评价方法分类

1.2.1 主观质量评价

主观图像质量评价方法的核心是通过观察人员对图像进行打分，从而反映人类视觉对图像质量的真实感知。为确保评估结果的科学性和一致性，首先要建立明确统一的评分标准，然后在一定的观察条件下，依据评分标准组织观察人员根据主观判断对图像质量进行打分，得到该图像质量的主观测试分数。国际电信联盟（International Telecommunication Union, ITU）^[6]制定颁发多个标准用于指导主观图像质量评价的实施，其中 ITU-R BT.500.11 标准是针对平面图像质量测试的规范，详细规定了实验测试的环境、呈现准则、评分尺度等关键要素。主观 IQA 实验一般采用 5 分制评分测度得到评分，图像质量极佳为 5 分，低劣为 1 分，最终图像的分数主要通过平均主观得分（Mean Opinion Score, MOS）或平均主观分差（Differential Mean Opinion Score, DMOS）^[5]来呈现。得到并收集观察者的评分后，进行归一化处理得到图像的 MOS，直接对图像的整体质量进行评估；而 DMOS 则是通过分析观察者对参考图像（无失真）和失真图像评分的差异，然后进行归一化处理后得到，用其量化图像质量的相对变化。

目前，主观图像质量评价是基于人类视觉系统对图像进行直接视觉感知，最能反映人对图像的真实感受，从而评判出图像的质量，在图像质量评价中具有重

要作用。但即使在标准的实验环境下，受到实验动机、环境、心理、认知等多种因素的影响，不同观察者对图像质量的主观感受也难以统一，导致主观评分结果稳定性和可移植性较差。此外，主观实验费时费力，在实际图像处理系统的应用中有较大的局限性。

1.2.2 客观质量评价

客观图像质量评价方法分为三类：全参考评价、半参考评价和无参考评价。

(1) 全参考评价

全参考图像质量评价算法是一种基于现有参考图像信息，通过比较待测失真图像与参考图像之间的差异，分析待测图像的失真程度，并据此得到待测图像质量评估分数的方法。Wang 等人^[7]基于结构相似度（Structural Similarity, SSIM），通过比较参考图像与失真图像之间的局部亮度信息、局部对比度信息和局部结构信息，得到图像局部相似度得分，然后使用闵可夫斯基模型对局部得分进行综合处理，从而得到图像的整体质量分数。Pei 等人^[8]通过计算多个 DoG 频带特征的非线性组合用于模拟人眼视觉系统，并采用随机森林回归方法训练特征，最终得到图像质量评分。Gao 等人^[9]使用深度神经网络测试失真参考图像间的局部特征相似性，再汇集局部相似分数从而获取全局图像质量分数。江等人^[10]利用色貌指标提取图像结构信息并得到相位一致性量值；然后，利用均方根法获取对比度相似图，通过相应颜色通道得到色度相似图；最后，结合三部分特征，采用标准差法进行求和池化，得到图像最终评分。

该类方法充分利用参考图像的全部信息进行图像质量评价，所以在这三类方法中，其预测性能最高，但在实际应用具有一定的局限性。

(2) 半参考评价

半参考图像质量评价算法的关键在于仅借助参考图像的部分信息对失真图像进行质量评估。这种方法通常采用更具表征能力的全局特征，一般适用于无线传输等特定场景下的图像质量评价问题。Wu 等人^[11]基于视觉显著性失真（Attended Visual Content Degradation, AVCD），首先提取图像视觉显著性和梯度上两方面的结构直方图，然后通过比较原始图像和失真图像在上述两种直方图上

的信息差异，最终得到质量评价分数。Hu 等人^[12]提出了一种新的基于多尺度特征增强和聚合的半参考图像去模糊质量评估方法。利用多功能视觉转换器 Transformer 生成预测模糊-去模糊图像（BDI）对的多尺度特征和差异信息隐式地增强初始深度特征，然后将不同尺度的增强特征聚合后映射到 BDI 对的质量差异上。

由于仅依赖图像的部分信息，RR-IQA 方法相较于 FR-IQA 方法在数据处理量上显著减少，因而具有更高的计算效率和处理速度并且能得到与主观评价一致性较高的评价分数。然而，在缺失参考图像的情况下，RR-IQA 方法的适用性受到限制，无法独立评估失真图像的质量。在实际应用中，原始无失真参考图像的获取通常较困难，甚至根本不存在无失真图像，因此进一步探索复杂度高但适用性强的无参考图像质量评价算法，对满足实际应用中图像质量评价的需求有重要意义。

（3）无参考评价

无参考图像质量评价算法也称盲参考评价算法，是一种无需依赖任何参考图像的信息即可完成图像质量评价任务的方法。无参考评价算法仅仅基于失真图像自身的特征和统计信息，通过分析图像的内在属性^[14]（如清晰度、梯度和失真程度等）进行图像质量的评估，将其量化为一个数值化的质量分数。相较于 FR-IQA 和 RR-IQA，无参考图像质量评价在计算复杂度、参数需求以及成本控制方面更具显著优势，尤其适用于无法获取参考图像的实际应用场景。

针对模糊图像的质量评价方法研究，无参考图像质量评价方法可分为三类：基于空域与频域的方法、基于学习的方法以及基于视觉感知的方法。

① 基于空域与频域的方法

空间域评价函数利用图像在空间域中的相关特征识别模糊图像和清晰图像，而图像的频率^[13]特性是衡量图像中灰度变化强度的重要指标，通常是通过傅里叶变换（Fourier Transform, FT）、小波变换（Wavelet Transform, WT）等方法进行提取。Zhan 等人^[14]利用最大梯度和梯度的可变性，以与主观评分高度一致的方式预测模糊图像的质量。Li 等人^[15]通过将梯度图像分割成等大小的块并计算正交矩表征图像的形状，从而获得锐度分数。Zhang 等人^[16]通过计算像素八个方向

的灰度值评估边缘锐度。Liu 等人^[17]利用基于激活机制的 Canny 边缘检测算法获取图像边缘的位置和方向，建立边缘边宽直方图，通过加权平均边宽得到锐度评价度量。Chen 等人^[18]使用 Prewitt 算子和 Gabor 滤波器计算图像的平均梯度，以预测织物表面图像的局部锐度值。Bahrami 等人^[19]通过计算图像像素的最大局部变异 (Maximum Local Variation, MLV) 及其分布的标准差作为测量锐度的度量标准。Gu 等人^[20]基于自回归 (Autoregressive, AR) 模型参数，通过计算局部估计的 AR 系数的能量和对比度差异，从而得到图像清晰度分数。沈等人^[21]利用灰度图像显著性分类提出 SOA 方法，对模糊图像进行质量评价。Zhang 等人^[22]利用卡通纹理分解 (Cartoon-Texture Decomposition, CTD) 的特性，将图像分为边缘突出的卡通部分和有噪声的纹理部分，然后分别估计模糊度水平和噪声水平从而预测结果。De 等人^[23]使用基于离散余弦变换 ((Discrete Cosine Transform, DCT) 系数的特征生成图像清晰度评估模型。Hu 等人^[24]利用 Swin 转换器 (Swin Transformer, ST) 提取图像特征，然后设计了一个即插即用的全局自注意力模块 (Global Self-Attention Block, GSAB) 同时在模型中引入 Transformer 模块，最后，通过双分支结构推导出图像质量分数的预测。Bae 等人^[25]采用结构对比指数 (Structural Comparison Index, SCI) 和 DCT 块表征人眼视觉系统 (Human Visual System, HVS) 对局部图像特征和各种失真类型的视觉质量感知。Baig 等人^[26]基于离散傅里叶变化 (Discrete Fourier Transform, DFT)，首先计算图像的导数，然后计算基于块的 DFT，并在块级别对每个交流系数的比率进行平均，然后将二者结合，从而估计整体图像感知质量。Gvozden 等人^[27]首先计算图像的均方根用于提取图像的局部对比度信息，并利用小波变换中的对角小波系数进行排序和加权处理，最终得到图像的质量评价分数。刘等人^[28]利用灰度图像梯度获取人眼视觉敏感区域和非人眼视觉敏感区域提出 Is 模型，对图像进行质量评价。

基于空间域的方法具有计算量小的优点，但大多依赖于图像的细节特征，容易受到噪声的影响，故该类方法预测性能比较低；基于频域的方法需要将图像信息从空间域转换到光谱域，在一定程度上增加了计算的复杂度。

② 基于学习的方法

与传统方法相比，基于学习的方法的优势在于能够从大量数据中自动提取图

像特征的多层表示，进而获取丰富的图像特征信息，实现质量分数的映射，从而提高评价结果的准确性。基于学习的方法的一般流程如图 1-3 所示。

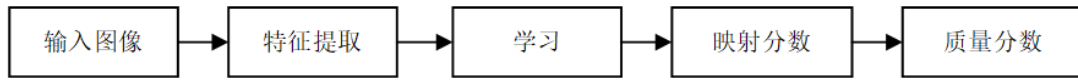


图 1-3 基于学习的方法的一般流程

早期基于学习的无参考图像清晰度评价方法主要是基于机器学习的支持向量机（Support Vector Machine, SVM）模型，包括支持向量回归（Support Vector Regression, SVR）和支持向量聚类（Support Vector Clustering, SVC）方法。Chen 等人^[29]利用最大局部变化图从多分辨率图像中提取结构特征，同时提取多尺度 Log-Gabor 滤波图像的梯度信息作为结构特征的结构，并与熵特征相结合；然后，将提取的特征进行融合形成质量感知特征向量，最后通过 SVR 映射到质量评分中。Liu 等人^[30]利用方向选择机制的各向异性和梯度方向效应对视觉的影响，提取结构信息，然后进一步利用 Toggle 算子提取边缘信息作为局部模式的权值，最后通过 SVR 训练具有优化特征和主观评分的预测模型。近年来，基于深度学习的图像质量评价方法逐渐成为研究热点，此类方法通过构建深度学习模型，并直接从图像数据中学习并生成图像评价分数。Zhu 等人^[31]基于优化卷积神经网络结构，自动提取独特的图像质量特征，并通过归一化和丢包策略完成预测评价得分。Li 等人^[32]采用一种分块处理的策略，将整个图像分割成多个小块，利用深度卷积神经网络（Deep Convolutional Neural Network, DCNN）提取各个块中高级特征，然后，聚合来自不同块的特征信息并输入最小二乘回归模型，从而获得锐度评估值。Zhang 等人^[33]提出了一种盲图像质量评价的深度双线性（BIQA）模型，该模型适用于包含综合失真和真实失真的图像，能够有效预测图像质量值。Ji 等人^[34]通过对 HSV 空间中的色调图像进行引导滤波，提取图像颜色中的失真信息，然后对传统残差块进行增强形成从图像中捕获多尺度特征的特征提取模块，最后双分支残差网络（Dual-Branch Residual Network, DBRIQA）对图像进行质量评价。

基于机器学习的方法能够自动学习得到图像特征的多层表示，从而对图像质量进行评估，但此类方法通常对大数据集的依赖性较强，并且评价的结果很大程度上依赖特征的提取。而基于深度学习的图像质量评价方法通过神经网络的自动

特征提取能力，为图像质量评价领域带来了新的突破和发展。

③ 基于视觉感知的方法

近年来，研究者致力于将 HVS 的感知特性融入图像质量评价算法，以期提升算法的性能，此类方法通过模拟人类视觉通道的完整视觉感知过程，强调了不同的视觉信息对人眼的不同贡献度。Jia 等人^[35]将图像视觉显著性相似性和图像对比度相似性两个特征分量进行结合，提出了 CVSS (Contrast and Visual Saliency Similarity) 算法得到图像质量分数。Liu 等人^[36]利用遮掩图像的结构信息，设计了 GSM (Gradient Similarity Metric) 算法模型，从而对失真图像进行质量评分。闫等人^[37]考虑到人眼视觉特性，融合图像区域灰度信息对比度与 SSIM 算法，将参考图像与失真图像的对比度结构相似度值进行加权归一，由此提出一种改进的基于区域对比度和结构相似度的图像质量评价方法。然而，HVS 本身是一个高度复杂的非线性系统，而图像质量评价的应用场景又极为广泛。因此，许多算法尝试基于人眼视觉特性提取特征，并结合深度学习等先进方法对图像质量进行评价。如何将人类视觉系统的特性深度融入图像质量评价算法，进而提出既具备高精度又兼具高效率的图像质量评价算法，仍是当前该领域亟待解决的研究难点。

综上所述，无参考图像质量评价方法研究主要是基于频域和空域，基于学习以及基于视觉感知三类。然而，目前仍然存在一些问题，多数方法对视觉感知特性的利用有所不足，图像特征的组合较为单一。为了更好地评估模糊图像的质量情况，在人眼视觉系统的基础上，关注视觉感知特性，提取图像合适的特征并组合，建立模型进行质量评价，进一步提高算法的预测精度，具有重要的研究意义。

1.3 本文研究内容与结构安排

1.3.1 主要研究内容

人类视觉系统是人类处理图像信息的基础，而图像特征信息对于图像质量评价存在优劣之分。因此，如何在人类视觉感知基础上选择合适的图像特征并将其进行组合，进而将其应用到图像质量评价算法中，并结合现代图像处理技术对算法的需求，提出高精度和高效率的图像质量评价算法模型是研究的关键。根据前

面对研究现状的分析总结,分别针对基于空间域和频域的方法和基于学习的方法设计了两种模型。

本文具体研究内容如下:

(1) 基于感兴趣区域的无参考图像质量评价方法。首先提取图像色差和信息熵特征信息,并进行组合,从而确定整幅彩色图像的兴趣区域(Area of Interest, AOI);然后对上述局部 AOI 图像深入分析色差指标,分别进行绝对与相对色差变化计算,获得质量评价分数。为了验证该方法的性能,对四个常用图像数据库中的模糊图像进行对比实验测试,分析预测精度和泛化性,结果表明该方法表现良好。该算法将信息熵从灰度图像扩展到彩色图像,并结合色差指标,弥补了对图像中颜色信息利用的不足。

(2) 基于 PSO-GRNN 算法的模糊图像清晰度评价。无参考图像质量评价算法的研究框架主要分为两步,即图像特征的提取以及特征池化。这里研究一种基于机器学习的不同光照成像条件下图像清晰度客观评价方法。该方法选择视觉显著性、色差和梯度特征信息进行组合,从图像中提取这三个方面的相关特征信息作为不同光照成像条件下模糊图像评价的特征值;然后,建立一个基于粒子群优化的广义回归神经网络(PSO-GRNN),对上述提取的特征值进行训练,并确定最终的模糊图像评价结果。分别在多失真数据库和自然失真数据库上进行验证,实验结果表明,所提方法整体的预测准确度表现优良。尤其针对不同成像条件下的模糊失真图像,该质量评价算法具有良好的性能。

1.3.2 结构安排

论文主要以人眼视觉感知为基础,对相关的彩色图像特征进行选择 and 组合,结合机器学习内容从而设计算法。论文的整体结构安排如下。

第一章 绪论。本章首先介绍了图像质量评价研究的背景和意义,然后阐述图像质量评价的研究现状,最后介绍了本文的具体研究内容。

第二章 图像质量评价理论基础。本章首先从人眼视觉系统出发,描述人类视觉系统的感知特性及其在图像质量评价算法中的应用,然后进一步介绍相关图像特征,以及评估比较图像质量算法所要用到的相关数据库,为本文后续的工作

奠定了理论基础。

第三章 基于感兴趣区域的无参考图像质量评价方法。首先提取色差和信息熵特征信息并进行组合，确定整幅彩色图像的感兴趣区域（AOI）；然后对上述局部 AOI 图像分别进行绝对与相对色差变化计算，获得质量评价分数。为了验证该方法的性能，对 TID2008 等四个常用数据库中的模糊失真图像进行对比实验测试，结果表明，该模型在预测精度和泛化性两方面表现良好。

第四章 基于 PSO-GRNN 算法的模糊图像清晰度评价。这里研究一种基于机器学习的模糊图像清晰度客观评价方法。该方法选择视觉显著性、色差和梯度信息作为图像特征，从图像中提取这三个方面的相关特征信息作为不同光照成像条件下模糊图像评价的特征值；然后，建立一个基于粒子群优化的广义回归神经网络（PSO-GRNN），对上述提取的特征值进行训练，并确定最终的模糊图像评价结果。分别在多失真数据库和自然失真数据库上进行验证，实验结果表明，所提方法在预测准确度有优良的表现。尤其针对不同成像条件下的模糊失真图像，该质量评价算法具有良好的性能。

第五章 总结与展望。对论文的研究内容进行简要阐述，思考现有研究不足之处，然后对无参考图像质量评价领域的研究前景做出展望。

第 2 章 图像质量评价理论基础

图像作为一种实用且常用的信息载体，为人类生活提供了大量的信息。图像的呈现是为了服务于人眼的，而图像质量评价的最终目标是替代人类视觉系统（Human Vision System, HVS），对图像质量做出准确判断。因此，深入研究人类视觉系统在认知和读取外界信息时所表现出的特性，对图像质量评价有着深远的意义。本章将从人类视觉系统的相关知识、彩色图像相关特征提取以及用于图像质量评价算法评估的数据集和评价指标四个方面展开介绍，这些内容将为后续研究工作提供理论基础。

2.1 人眼视觉系统

人眼视觉系统是人体中一种复杂的感官系统，外界光线通过大脑皮层的神经通信，被转化为意识中的图像信息。利用这一系统，人类能够对不同环境和场景中的视觉信息进行自适应地感知，从而精准识别和区分颜色、亮度、纹理等关键信息。本节将简要介绍人眼视觉系统，并深入探讨可能对图像质量产生影响的相关视觉特征。

2.1.1 人眼生理结构

人眼视觉系统是一个复杂且精密的感知系统，主要由眼球、视觉神经以及大脑的视觉皮层组成，其核心功能是接收并处理来自外界的光信号，从而让我们能够感知周围环境并获取相关信息。眼球作为视觉系统的核心器官，包含角膜、瞳孔、虹膜、晶状体、视网膜等重要结构^[38]，结构如图 2-1 所示。视网膜上分布着杆状细胞和锥状细胞两类感光细胞，前者主要负责低光照条件下的视觉，后者主要负责颜色和细节的感知。

当光线依次经过眼睛的各个结构时，会发生折射与聚焦，由此精准投射到视网膜上的感光细胞，接着感光细胞将获取的光信号转化为电信号，然后信息通过视神经被传输到大脑。视神经在传输过程中会对双眼接收到的信息进行初步整合，随后传送到大脑的视觉皮层。位于枕叶的视觉皮层，是负责处理视觉信息的

关键区域。视觉信息首先在大脑的初级视觉皮层进行初步加工,包括图像的边缘、颜色和运动等基本特征的提取和分析。然后,这些基础信息会被传递到更高层次的视觉区域,进行更为复杂的分析,如物体识别、面部辨识和空间定位等。

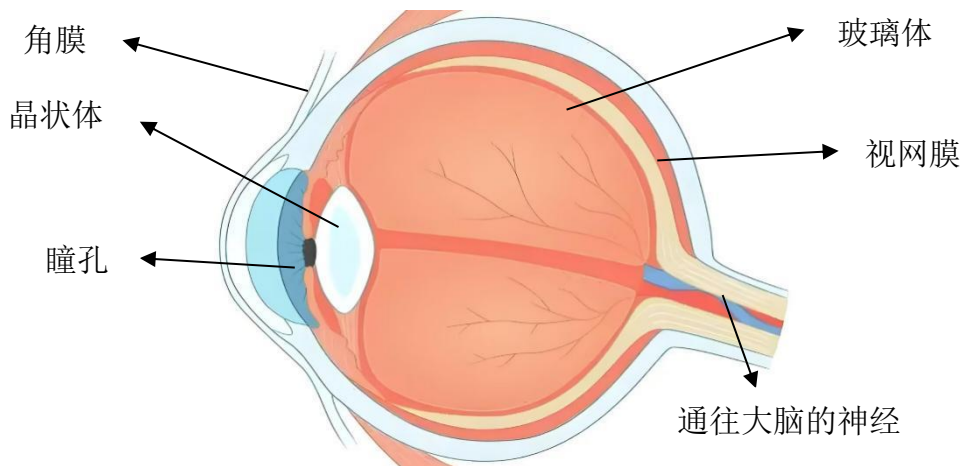


图 2-1 人眼生理结构图

2.1.2 人眼视觉感知特性

视觉感知特性,就是人类视觉系统感知图像信息时所表现出的一系列特性,这一过程涉及高度复杂的信息处理机制。人类视觉系统的视觉感知特性涵盖多个学科领域,其中包括视觉心理学、视觉生理学以及物理学等^[39]。日常生活当中,对同一物体进行观察,人眼的感知效果由于时间、角度、光照等因素也存在差异。在差异产生过程中,人类视觉感知特性起到决定性作用,其中包括对比敏感度、视觉注意力机制等。

(1) 对比敏感度

视觉感知中的对比敏感度 (Contrast Sensitivity) 是衡量人眼在不同亮度差异下分辨细节能力的核心指标^[40]。它反应了视觉系统检测物体与其背景之间细微对比度变化的能力,而非单纯依赖高对比度下的清晰度。例如,在低能见度条件下辨别远处的物体轮廓时,对比敏感度发挥关键作用。人眼的对比敏感度通过对比敏感度函数 (CSF) 量化,计算公式如式 2-1,该函数揭示了视觉系统对不同空间频率 (即图像中明暗变化的密集程度) 的响应特性。具体而言,在中等空间频率 (如日常物体的边缘和纹理) 下敏感度最高,而对极粗 (低频率) 或极细 (高频率) 的细节敏感度较低。

$$CSF(f) = A \times e^{-\left(\frac{\ln(f/f_0)}{2\sigma}\right)^2} \quad (2-1)$$

式中： A 是与最大对比敏感度相关的常数，决定了函数的峰值高度，反映了人眼在最佳状态下的对比敏感度； f_0 是对比敏感度达到峰值时对应的空间频率，即人眼最敏感的空间频率， σ 是与函数宽度相关的参数，决定了对比敏感度随空间频率变化的快慢。

(2) 视觉注意力机制

视觉注意力机制^[39]是大脑的一种认知功能，通过对大脑有限的注意力资源进行选择分配，优先加工处理感兴趣视觉信息，而忽略弱化其他不重要的信息。在日常生活中，视觉显著性区域检测（Visual Saliency Detection）是视觉注意力机制的运行表现。具体而言，人眼会自动识别并聚焦图像中最引人注目（有重要信息）的区域，也被称为显著性区域。如图 2-2 所示，当人眼观察一幅人物图片时，视觉注意力机制会引导观察者优先关注图片中心的人物，而对背景信息则仅作短暂扫视，不会过多停留。这种选择性关注的特性是视觉注意力机制的核心功能，它优化了视觉信息的处理效率，使大脑能够高效地提取和分析关键信息。



图 2-2 测试图像及其视觉显著区域

2.2 图像信息表征

在图像质量评价研究中，可以利用的图像特征有许多，有空间域的特征、频域率的特征，如梯度、对比度等，在此基础上图像质量评价方法也有不同分类。

为了衡量图像的失真情况，更好地对图像进行质量评价，主要使用信息论中的信息熵特征、图像空间域的色差和梯度特征以及图像频域率的显著性特征进行后续算法的构建。

2.2.1 信息熵

信息熵是信息论中一个用于量化随机变量的不确定性的概念。在图像处理领域，图像每个像素的灰度值或颜色值可以看作随机变量。因此，图像信息熵可被定义为对图像中像素灰度值或颜色值分布不确定性的度量。由于图像信息的丰富程度与像素值分布的不确定性密切相关，不确定性越高，意味着图像包含的信息越丰富，所以图像信息熵能够有效反映一幅图像包含信息的丰富程度。传统图像信息熵的计算如公式 2-2 所示，对象通常是灰度图像^[41]：

$$E = -\sum_{l=1}^{L-1} P_l \log_2 P_l \quad (2-2)$$

式中： L 表示最大灰度， P_l 为图中灰度级 l 的概率，而

$$P_l = \frac{n_l}{M \times N} \quad (2-3)$$

式中： n_l 表示灰度为 l 的像素个数， $M \times N$ 表示区域大小。

2.2.2 梯度

梯度是图像空间域的一个重要特征，常被用于评估模糊失真图像的质量^[14-19]。对于一幅失真图像而言，梯度能够有效表征图像像素变化的幅度，具体而言，在图像的边缘区域，灰度值变化较大，对应的梯度值也较高；而在图像中比较平滑（清晰）的区域，其灰度值变化较小，相应梯度值也较低。目前，计算图像梯度的方法有许多，常见的有 Prewitt 算子、Canny 算子、Sobel 算子和 Roberts 算子。Prewitt 算子通过计算特定区域内像素的灰度值所产生的差值实现边缘检测，适用于简单场景下的边缘提取。Canny 算子是一种基于多尺度分析的边缘检测算法，能够有效抑制噪声并提取出清晰的边缘，是目前应用较为广泛的边缘检测方法之一。Sobel 算子则结合了高斯平滑和微分导数，主要用于检测图像中的水平和垂直边缘，其在边缘定位精度上具有一定优势。Roberts 算子通过对图像相邻

像素进行交叉差分运算计算梯度，具有计算量小、计算速度快的特点，适用于对实时性要求较高的场景。其中 Roberts 算子定义像素点在 (i, j) 处的梯度值为 $G(i, j)$ ，其计算公式如 2-4 所示。

$$G(i, j) = \sqrt{[I(i, j) - I(i, j+1)]^2 + [I(i, j) - I(i+1, j)]^2} \quad (2-4)$$

通过对不同模糊程度的图像进行梯度计算，可以得到相应的梯度图，进而识别其结构信息的变化。研究表明，模糊图像的模糊程度与结构信息的变化呈正相关关系，即模糊程度越高，图像的结构信息丢失越严重，梯度图中的边缘特征也越不明显。因此，梯度图能够有效反映图像的模糊程度，通过计算梯度可以更准确地量化图像的模糊程度，为图像质量评估提供重要依据。

2.2.3 色差

色差是颜色科学研究中提出的一个概念，用于表示颜色色度之间的差异。国际照明委员会（The International Commission on Illumination, CIE）负责包括色差在内的照明相关技术标准的国际协调。CIE 技术委员会制定 CIEDE2000 色差方程，开始色差公式常用于色调匹配中的颜色再现。近年来，相关研究者已将色差指标应用于图像质量评价领域。如有研究者在 CMYK 颜色空间^[42]中使用图像的全局平均色差值对图像的质量进行评估，除此之外，有研究者对色差指标进行了扩展，在 CIELAB 颜色空间构建了彩色图像差异（Color Image Difference, CID）和改进的彩色图像差分（improved Color Image Difference, iCID）方法评估图像的质量。

在现有研究中，像素的色差指标表示两个像素在特定颜色空间的各通道之间的差异，能够衡量图像模糊失真程度。在 RGB 模糊图像中，两个相邻像素之间的差可以表示色差，进行相应颜色空间转换后，得到对应颜色空间的图像 X ，其每个像素包含 3 个颜色分量：亮度 S_1^* 、颜色通道 S_2^* 和 S_3^* 通道，其色差指数 (ΔE) 通过式 2-5 计算。

$$(\Delta E) = \sqrt{(S_{11}^* - S_{12}^*)^2 + (S_{21}^* - S_{22}^*)^2 + (S_{31}^* - S_{32}^*)^2} \quad (2-5)$$

式中： S_{11}^* 和 S_{12}^* 是 X 两个相邻像素的亮度通道； S_{21}^* 和 S_{22}^* 是 X 两个相邻像素的颜色通道； S_{31}^* 和 S_{32}^* 是 X 两个相邻像素的另一个颜色通道。

2.2.4 显著性

人类能够轻松识别图像中不同区域的重要性，并将注意力有选择地集中在相对重要的部分，这一过程主要依赖于人眼视觉感知中的视觉注意力机制。

在图像显著性检测领域，常规方法有基于频域的频谱残差（Spectral Residual, SR）算法，对一幅图像 $f(x, y)$ ，首先进行傅里叶变换 $F(u, v)$ ，接着计算幅度谱的对数 $L(u, v)$ ，然后计算频谱残差 $R(u, v)$ ，然后通过式 2-6 计算得到显著图 $S(x, y)$ ，其中 $IFFT$ 为逆傅里叶变换。

$$S(x, y) = IFFT\left(e^{R(u, v)}\right) \quad (2-6)$$

除此之外，有研究人员提出了一种显著区域检测方法，即SDSP^[43]算法，通过结合三个简单的先验进行显著性区域的检测。给定一个图像 $f(x)$ ， $x \in \Omega$ 其中 Ω 表示RGB颜色空间中的图像空间域，首先， $f(x)$ 需要转换到CIELab空间，产生三个通道 f_L 、 f_a 和 f_b ，然后，将显著性 $S_F(x)$ 定义为：

$$S_F(x) = \left((f_L \times g)^2 + (f_a \times g)^2 + (f_b \times g)^2 \right)^{1/2}(x) \quad (2-7)$$

每个像素 x 都可以映射到颜色平面 f_{an} 、 f_{bn} 中的一个点，点 x 的颜色显著性 $S_C(x)$ 定义为：

$$S_C(x) = 1 - \exp\left(-\frac{f_{an}^2(x) + f_{bn}^2(x)}{\sigma_C^2}\right) \quad (2-8)$$

设 c 为图像 $f(x)$ 的中心。然后，“位置先验”下 x 点的“位置显著性”可以表示为一个高斯映射：

$$S_D(x) = \exp\left(-\frac{\|x - c\|_2^2}{\sigma_D^2}\right) \quad (2-9)$$

给定的一幅图像 $f(x)$ ，利用这三个简单先验分别计算出三个显著性映射 $S_F(x)$ 、 $S_C(x)$ 和 $S_D(x)$ 。其计算公式为：

$$SDSP = S_F(x) \cdot S_D(x) \cdot S_C(x) \quad (2-10)$$

2.3 图像质量评价常用数据库

为进一步推动图像质量评价课题研究发展，按照统一标准，国际上多个研究单位建立了多个包含参考图像和失真图像的图像质量评价数据库。这些数据库可根据图像失真类型的差异分为多失真数据库与自然失真数据库。

2.3.1 多失真图像数据库

多失真图像数据库中包含多种不同失真类型和程度的图像，通过压缩、添加噪声、模糊处理等人工操作对图像生成失真，利用该数据库中的失真图像进行实验测试，能够评估提出的图像质量评价算法的性能表现。表 2-1 展示了 TID2008 等四个常用多失真数据库的基本信息。

表 2-1 多失真数据库信息说明

数据库	参考图像	图像数量	主观分数指标	图像尺寸
TID2008	25	1700	MOS	512×384
TID2013	25	3000	MOS	512×384
CSIQ	30	866	DMOS	512×512
LIVE	29	799	DMOS	——

(1) TID2008 图像数据库

TID2008 (Tampere Image Database 2008) 是由芬兰坦佩雷大学 (Tampere University of Technology) [44] 开发的一个标准化图像质量评估数据集，被广泛用于图像质量研究中相关算法性能测试。该数据集包含 25 张参考图像，如图 2-3 所示，其涵盖自然景观、人脸、建筑以及合成图像等多种类型，以及 1700 张测试图像，每张参考图像都经过多种类型的失真处理 (包括 17 种不同的失真类型，每种失真类型都有 4 种不同的强度级别)，模拟了图像在现实世界中遇到的失真

情况。这些失真类型包括模糊、噪声、压缩、对比度、亮度等，为图像质量评估提供了丰富且多样化的测试样本。通过这些多样化的失真类型和强度级别，TID2008 为研究人员提供了一个全面的测试平台，有效推动图像质量评估技术的发展，并为学术界和工业界提供了有价值的参考数据。



图 2-3 TID2008 数据库参考图像

(2) TID2013 图像数据库

TID2013 (Tampere Image Database 2013) 是一个标准化 IQA 数据集^[45]，旨在支持和推动图像质量评估领域的研究。作为 TID2008 数据库的升级版，TID2013 在多个方面进行了改进和扩展，提供了更为广泛和多样化的测试样本，为图像质量评估算法的开发、比较和验证提供了更加可靠的基准。TID2013 数据集涵盖了 25 张参考图像。相较于 TID2008，该数据库一个显著特点是它包含了更为丰富的失真类型和强度级别。在 TID2008 所包含的 17 种失真类型的基础上，TID2013 进一步扩充了失真类型的范围，考虑了更多真实世界中频繁出现的失真因素。每种失真程度均包含了 120 张图像，这一设计使得数据集可以更精准地映射出各种不同的图像质量退化情形，极大地增强了其在实际应用中的代表性与普适性。此外，TID2013 还提供了更为丰富的失真强度级别，失真程度可以从细微的改变延伸至严重的失真，这使得研究人员能够更加精细地评估算法在不同失真程度下的性能表现，进而为图像质量评估领域的研究提供了更为精准、更具深度的实验依

据与分析基础。

（3）CSIQ 图像数据库

CSIQ (Challenge Set on Image Quality) 是一个常用的人工失真图像数据库, 该数据库涵盖了 30 张高质量的参考图像, 5 种不同类型和程度的失真, 总共包括 1162 张图像。这些图像涵盖多种自然场景类别, 包括人脸、风景、建筑、动物、静物等。CSIQ 数据库图像有 6 种失真类型, 分别为 JPEG 人工伪影、高斯模糊、色彩偏移、加性白噪声和运动模糊, 针对每种失真类型, 设计不同等级的失真程度, 更加具体地模拟图像失真情况。CSIQ 数据库中每幅图像都有丰富的主观实验评分数据, 由多位观察者对图像质量进行评分后汇总而成, 最终以 DMOS 的形式呈现。这些主观评价数据为图像质量评估算法的性能测试和比较提供了重要依据。

（4）LIVE 图像数据库

LIVE 图像质量评价数据库 (LIVE Image Quality Assessment Database) 是一个广泛应用于 IQA 研究中的标准数据集^[46], 用于评估不同图像和视频处理算法与人类视觉感知的一致性。LIVE 数据库共有包括 29 幅高质量的参考图像在内的 799 张图像, 其中失真类型有高斯模糊、JPEG 压缩、色彩量化、加性白噪声和运动模糊 5 种。这些图像涵盖了多种自然场景, 如人脸、动物、风景、建筑等, 能够有效模拟现实应用中的图像失真。此外, LIVE 数据库有充足的主观实验数据, 除此之外, 多名观察者对图像质量进行评分后, 汇总得到 DMOS, 为图像质量评估算法的性能测试和比较提供了重要支持。

2.3.2 自然失真图像数据库

自然失真图像数据库是指从真实场景中的失真图像中专门收集和整理的数据集, 涵盖了模糊、压缩、噪声等多种失真类型, 常用于验证和评估 IQA 算法的性能。由于真实世界的不同场景、曝光条件以及各种拍摄设备, 数据库中图像失真情况更加复杂且具有挑战性, 能够有效模拟现实应用中的图像质量退化。这类数据库的主要目的是为 IQA 算法的性能验证和评估提供基准支持, 帮助研究人员更好地评估算法针对真实失真图像的表现。近年来, 随着机器学习以及深度学

习技术的发展,自然失真图像数据库也随之进一步发展,其规模更大且质量更高,如 BID 数据库、CID2013 数据库和 CLIVE 数据库,简要介绍见表 2-2。

表 2-2 自然失真数据库信息说明

数据库	失真图像数	主观分数指标	图像尺寸
BID	586	MOS	——
CID2013	474	MOS	1600×1200
CLIVE	1162	MOS	500×500

(1) BID 数据库

BID 数据库^[47]包含 586 张图像,分辨率从 1280×960 到 2272×1704。数据库中的图像根据模糊种类分为 5 类,清晰图像(无模糊)、失焦模糊图像、简单运动模糊图像、复杂运动模糊图像和其他图像。在主观评价实验中,约有 180 名受试者对数据库中的 586 张图像进行了评分,每张图像至少获得 10 个评分。测试图像的最终分数是基于分配给它的所有主观评分的平均值计算得出。图 2-4 是选取 BID 数据库中的四张模糊图像。



图 2-4 BID 数据库中四张模糊图

(2) CID2013 数据库

CID2013 (Chinese Image Database 2013) 数据库^[48]是用于图像质量评估研究的自然失真数据集,尤其在计算机视觉和图像处理领域中被广泛用于图像质量评估。CID2013 数据库由 6 个图像集组成,共包含 474 张图像,涵盖 8 种不同场景,这些图像是由 79 种不同的成像设备(包括手机、数码相机、数码单反相机

等)拍摄得到。在主观评分方面,采用混合绝对类别评级-配对比较的方法,共有多位观察者参与评分(其中67%为女性)。评分结构不仅提供了平均意见得分值(MOS),还针对图像的清晰度、颗粒度、亮度和色彩饱和度进行了单独评级。此外,还有少部分观察者的单独量表重新校准数据,以便对六个图像集的数据进行合并分析。

(3) CLIVE 数据库

CLIVE (Camera Model Image Quality Evaluation) 数据库^[49]是一个用于 IQA 研究的标准数据集,旨在系统性地评估图像质量评价方法在实际相机图像中的表现。此数据库包含 1162 张来自不同类型的相机设备拍摄的自然图像,适用于各种失真类型的质量评估,包括噪声、模糊、压缩等。该数据库主要用于评估与相机设备相关的图像质量问题,尤其是相机图像中普遍存在的失真情况。CLIVE 数据库的质量评分基于主观评分(MOS),每张图像的 MOS 分数通常是通过所有参与者的评分进行平均计算得出的。为了确保评分结果的可靠性和有效性,数据库在数据处理过程中采用了去除极端评分和异常值的方法,从而为图像质量评估算法的开发和验证提供了高质量的基准数据。

2.4 算法性能验证指标

客观图像质量评价算法旨在实现对图像质量的自动化、精准化评估,以此替代复杂且繁琐的主观实验评价方法。对于一个给定的测试图像集,通过深入分析其客观评价分数值 x 与主观分数值 y 之间的关系,进而有效衡量客观评价算法的性能。视频质量专家组 (Video Quality Experts Group, VQEG) 建议从预测准确性、预测单调性和预测一致性三个方面说明图像质量评价方法的性能。主要有以下五种性能指标:

(1) 皮尔逊线性相关系数 (Pearson Linear Correlation Coefficient, PLCC), 描述算法评价客观分数与主观分数的相关性,说明算法的预测准确性,定义为:

$$PLCC = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x} \right) \left(\frac{y_i - \bar{y}}{\sigma_y} \right) \quad (2-11)$$

式中： n 表示测试图像的数量， x_i 表示第 i 幅图像的主观分数， y_i 表示第 i 幅图像的客观分数， $\bar{x}$ 和 $\bar{y}$ 分别为主观分数 $\{x_1, x_2, \dots, x_n\}$ 和客观分数 $\{y_1, y_2, \dots, y_n\}$ 的均值； σ_x 和 σ_y 分别其标准差。

(2) 斯皮卡尔蛮排序相关系数 (Spearman Rank ordered Correlation Coefficient, SROCC)，说明算法预测的单调性，定义为：

$$SROCC = 1 - \left[6 \sum_{i=1}^n \frac{(r_{xi} - r_{yi})^2}{n(n^2 - 1)} \right] \quad (2-12)$$

式中： r_{xi} 和 r_{yi} 分别为 x_i 和 y_i 在其分数序列进行排序后的位置。

(3) 肯德尔排序相关系数 (Kendall Rank ordered Correlation Coefficient, KROCC)，说明算法预测的单调性，定义为：

$$KROCC = \frac{n_c - n_d}{0.5(n-1)n} \quad (2-13)$$

式中： n_c 表示排序后的主观分数和客观分数间的同序对 (Concordant Pairs) 的数量，而 n_d 则是异序对 (Discordant Pairs) 的数量。

(4) 均方根误差 (Root Mean Squared Error, RMSE) 比较算法评价客观分数与主观分数的绝对误差，说明算法预测的一致性，定义为：

$$RMSE = \left[\left(\frac{1}{n} \right) \sum_{i=1}^n (x_i - y_i)^2 \right]^{1/2} \quad (2-14)$$

(5) 平均绝对误差 (Mean Absolute Error, MAE)，定义为：

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (2-15)$$

但由于图像客观分数值和主观分数值往往存在不一样的值域范围和分布特性，因此在进行性能比较之前，需要将客观分数值映射到统一到同一尺度。对于 PLCC 和 RMSE 指标建议先用非线性方程进行统一尺度的转换，其中涵盖 4 参数和 5 参数的逻辑回归方程^[50]，其公式如下，本研究选择了后者进行转换，以确保后续性能比较的科学性与准确性。

$$p(x) = \frac{\beta_1 - \beta_2}{1 + \exp\left(\frac{x - \beta_3}{\beta_4}\right)} - \beta_2 \quad (2-16)$$

$$p(x) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + \exp[\beta_2(x - \beta_3)]} \right] + \beta_4 x + \beta_5 \quad (2-17)$$

式中： $\beta_1, \dots, \beta_n$ 表示参数拟合； x 代表客观 IQA 分数； $p(x)$ 表示逻辑回归后分数。

在客观图像质量评价算法的性能评估中，PLCC、SROCC 和 KROCC 的值越高（接近于 1），而 RMSE 和 MAE 的值越低（接近于 0），则表明该算法的预测性能更为优异。其中 PLCC 和 RMSE 两者都需要先进行非线性拟合，两个指标评价的效果具有一定的相似性；而 SROCC 和 KROCC 都是说明主观分数和算法评价分数的顺序相关的系数，也具有一定的相似性，而 MAE 指标在进行验证的时候，仅仅比较两个值的绝对误差，无法有效展示算法计算结果与测试图像主观分数的关系，特别是当数据库的主观分数值和模型计算得到的分数值偏差会比较大时，如果直接使用 MAE 效果更差。鉴于此，本文在分析客观图像质量评价算法的预测效果时，重点关注 PLCC、SROCC 和 RMSE 三个指标的数值表现，以此作为评判所提算法模型性能优劣的主要依据。

2.5 本章小结

本章主要介绍了人眼的生理结构，重点介绍了视觉感知特性中视觉注意力机制，其次简要的介绍了信息熵、梯度、色差以及显著性图像特征信息，然后具体介绍了评判算法模型性能所要用的数据库以及相关的评价指标。本文将利用相关图像特征对图像质量评价算法模型进行深入研究。

第3章 基于感兴趣区域的无参考图像质量评价

为提高图像质量评价模型的准确性,考虑到传统方法仅仅利用灰度图像相关特征进行图像质量评价,该方法难以全面反映图像质量的复杂性,导致评价结果存在偏差。在第一章研究现状中了解到图像质量评价有基于空间域的方法与基于视觉感知的方法。鉴于此,研究一种基于感兴趣区域的无参考图像质量评价方法。本章首先针对图像信息熵和色差特征的提取过程进行了系统性分析,并探讨了二者组合的合理性与优势。基于人眼视觉感知特性,精准定位失真图像中的感兴趣区域,以该区域作为图像质量评价的核心对象,替代传统方法中对整幅图像的全面评价。在此基础上进行色差的扩展计算,从而得到最终的评价分数,由此构建出无参考图像质量评价计算模型。最后在公共图像数据集上进行了系统性实验,根据实验结果,将本章算法与现有多多种图像质量评价算法进行性能比较分析,本章所提方法在预测准确性上有一定优势。

3.1 特征提取与组合

为更好利用图像彩色信息,从模糊图像的色差特征出发,提取了图像其他多种特征并与色差组合进行尝试实验,最终选择信息熵指标和色差进行组合,共同用于无参考图像质量评价研究中。

(1) 多通道色差

由前文可知,色差已被用于图像质量评价研究中,下面给出了一幅图像在某个像素点的色差指数 ΔE 的计算公式^[51],由于彩色图像具有多通道,因此具有 S 通道的水平和垂直方向的局部色差算子可以定义为:

$$\Delta E_h(i, j) = \sqrt{\sum_{n=1}^S (S_n^*(i, j) - S_n^*(i, j+1))^2} \quad (3-1)$$

$$\Delta E_v(i, j) = \sqrt{\sum_{n=1}^S (S_n^*(i, j) - S_n^*(i+1, j))^2} \quad (3-2)$$

式中: $S_n^*(i, j)$ 表示每个颜色通道在像素位置的强度, S 表示颜色通道数。

然后将两个方向的局部色差算子进行结合,得到像素局部色差这个新的指标,表示为式 3-3:

$$\Delta E_l(i, j) = (\Delta E_h(i, j) + \Delta E_v(i, j)) / 2 \quad (3-3)$$

公式 3-3 的计算过程类似于图像梯度特征提取，但是其含义不同于一般的梯度计算。色差是一个颜色科学指标，使用 3 个通道信息表示两种颜色之间的差异。图 3-1 是模糊程度不同的彩色图像的色差图，从色差图可以看到图像模糊程度不同，其对应的色差图也有明显的区别，故色差指标能够用于表征模糊图像的失真程度。

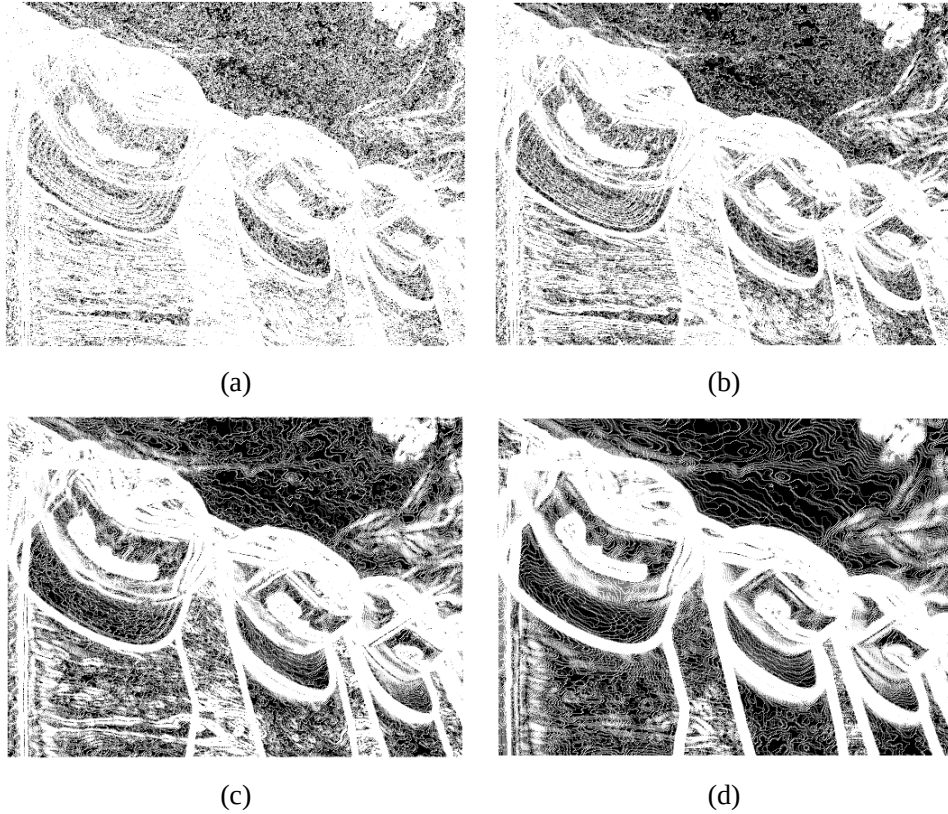


图 3-1 不同模糊程度色差图，(a)到(d)模糊程度递增

(2) 彩色信息熵

彩色图像相较于灰度图像，单个像素包含的信息多于灰度图像。这里考虑到仅仅利用灰度图像的信息熵可能对彩色图像的信息利用不够，故而将传统信息熵的计算过程扩展应用到彩色图像，对彩色图像各通道分别进行信息熵计算，然后选取各通道信息熵最大值作为该图像整体的信息熵结果，其计算公式如下所示：

$$E_i = -\sum_{l=0}^{L-1} P_{il} \log_2 P_{il} \quad (3-4)$$

$$P_{il} = \frac{n_{il}}{M \times N} \quad (3-5)$$

$$E_s = \max(E_i) \quad (3-6)$$

式中： $i=1,2,3$ ， P_{il} 为图像第 i 个通道的灰度级 l 的概率， n_{il} 表示第 i 个通道中灰度为 l 的像素个数， $M \times N$ 表示区域大小。

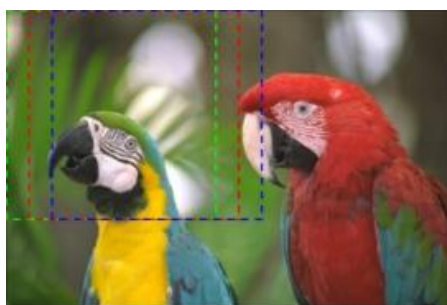
已知色差指标能够表示图像模糊失真程度，但对于一幅图像而言，包含最大色差的区域可能有多个，仅根据色差信息选取出的感兴趣区域并不唯一，这时考虑到不同区域对应的信息熵的值不一样，故本研究在色差基础上引入信息熵进行特征组合从而选取图像的感兴趣区域。

3.2 基于感兴趣区域的评价模型构建

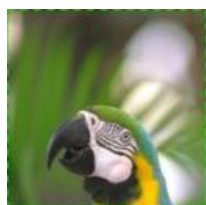
3.2.1 图像感兴趣区域选择

提取出彩色图像的信息熵和色差信息后，进一步利用这些信息对失真图像进行质量评价是关键步骤之一。对于一幅图像而言，其包含重要信息的区域即感兴趣区域，该区域的选择方法各不相同。在 IQA 研究中，可以通过合理选取感兴趣区域，获得更加符合 HVS 的 IQA 客观分数。目前，常规算法是以图像中心位置作为感兴趣区域，然而由于图像内容的差异性，图像的重点区域可能并不在图像的中心位置。因此，设计能够选择出包含相对重要信息的区域的算法显得尤为重要。已知图像的模糊度可以通过像素的强度变化表示，故像素强度变化的大小可以作为模糊度的准确指标。文献[16]已提出一种基于目标像素相对于其 8 邻域的最大灰度变化。除此之外，还有许多算法中都用到了局部最大值的概念。由此，可以将最大值这一概念应用到信息熵和色差这两个指标上。在选择感兴趣区域时，可以利用图像最大色差和最大信息熵指标进行选择。

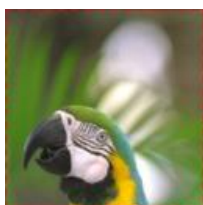
为了确保整幅图像都被充分考虑，采用重叠区域划分方法进行循环处理。具体而言，通过将图像划分为多个重叠的子区域，并逐一分析这些子区域的最大色差和信息熵，从而确定感兴趣区域。如图 3-2 所示，是 TID2008 数据库中一幅模糊图像及其三个位置不同的重叠区域图像。



(a) 整体图像



(b) 绿色区域



(c) 红色区域



(d) 蓝色区域

图 3-2 原始图像以及三个不同区域图像

本文提出的感兴趣区域选择方法的流程如图 3-3 所示：

第一步：图像裁剪框初始化。设定特定尺寸的图像裁剪框，用于从原始 RGB 图像中提取局部图像块。裁剪框的移动顺序为从左到右、从上到下，依次覆盖整个失真图像，确保图像的每个区域均被完整遍历。

第二步：颜色空间转换。将裁剪得到的每个局部 RGB 图像转换至预设的目标颜色空间，以获得该颜色空间下的三通道局部图像信息。这一转换过程旨在通过特定颜色空间的特性，更好地提取与图像质量相关的特征信息。

第三步：色差计算与最大值筛选。基于转换后的三通道局部图像，利用色差计算公式生成每块局部图像的色差图。通过分析色差图，确定包含最大色差值的局部图像块。色差图的生成与分析是识别图像中视觉显著区域的关键步骤，最大色差值所对应的局部图像块通常包含丰富的视觉信息，是图像质量评价的重要参考区域。

第四步：多通道信息熵计算与感兴趣区域确定。针对包含最大色差值的局部图像块，进一步计算其多通道信息熵。通过对比各局部图像块的多通道信息熵，确定具有最大信息熵值的局部图像块，该图像块即为模糊失真图像的感兴趣区域。

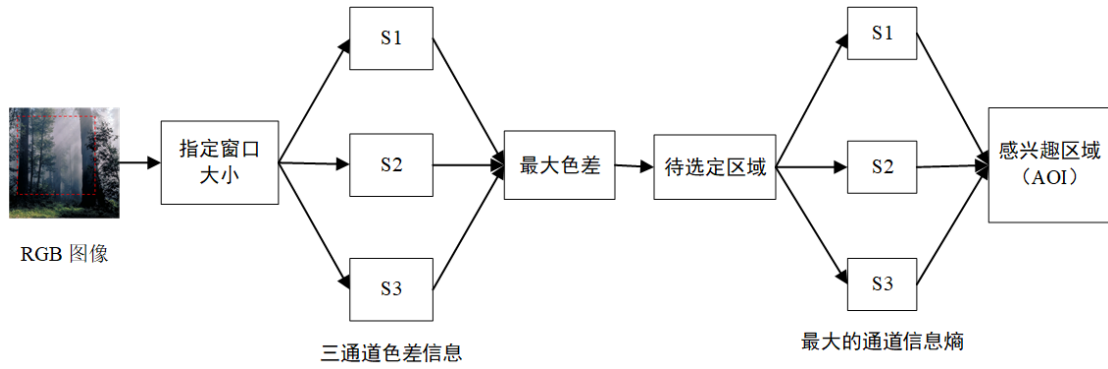


图 3-3 选择感兴趣区域的流程

为进一步验证所提感兴趣区域选择方法的可靠性，选择与上述相同的图像作为测试对象，通过计算待测失真图像及其对应的 AOI 的直方图，如图 3-4 所示，分析两者相似性。这一结果表明，AOI 能够有效保留待测失真图像的大部分信息，且其信息分布与原始图像具有高度相关性。直方图作为图像信息分布的重要统计特征，能够反映图像中像素值的分布情况。在本研究中，AOI 的直方图与原始失真图像的直方图高度相似，说明所选择的 AOI 在信息保留方面具有显著优势。这种相似性不仅验证了 AOI 选择方法的有效性，还表明选择的 AOI 可以作为图像质量评估的替代对象。

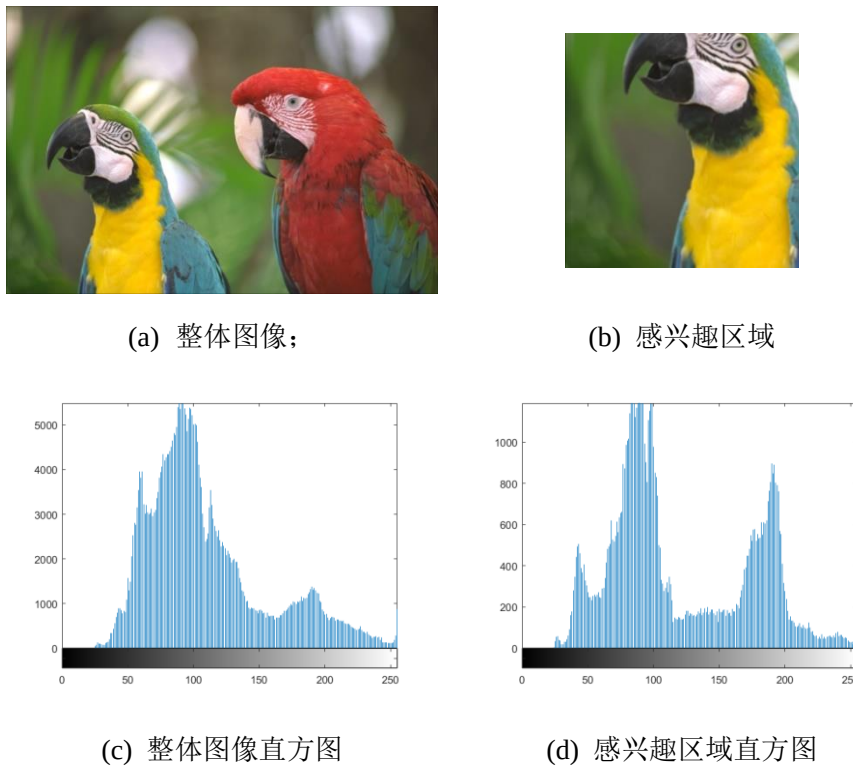


图 3-4 原始图像和感兴趣区域的直方图

3.2.2 基于感兴趣区域的质量评价模型

对于输入的 RGB 彩色图像，为了更好地利用彩色图像的亮度等相关信息，从而接近人眼视觉（HVS）的颜色感知，考虑将原始的 RGB 图像转换到更符合人类视觉的颜色空间，这一转换过程参考了常见的颜色空间转换方法。为了进一步处理已提取的 AOI，这里参考文献[51]中的相关操作。针对高斯模糊失真下的无参考图像质量评价，本研究提出的模型通过两个局部分量预测图像清晰度，即局部绝对色差变化 $LCDV_{\alpha}$ 和局部相对色差变化 $LCDV_r$ ，其中绝对色差变化是局部区域中最大色差和最小色差的差异，而相对色差变化则是基于最大色差、最小色差以及平均色差的综合计算结果。绝对色差变化可由式 3-7 定义：

$$LCDV_{\alpha} = \max(\Delta E_l) - \min(\Delta E_l) \quad (3-7)$$

式中：分量 $\max(\Delta E_l)$ 用于计算模糊图像的最清晰位置，而 $\min(\Delta E_l)$ 计算最模糊位置。其中 ΔE_l 表示给定 AOI 图像的局部色差图。在本节中，使用局部感兴趣图像色差的极大值作为所提算法一部分。除此之外，所提算法还考虑了模糊图像内容的相对变化。图像区域越清晰，邻域像素的相对变化越大；图像区域越模糊，邻域像素的相对变化越小，这表明像素间颜色变化的比例不同，整体清晰度评估也会产生一定影响。因此，本研究定义相对色差变化为：

$$LCDV_r = \frac{\max(\Delta E_l) - \min(\Delta E_l)}{\text{average}(\Delta E_l)} \quad (3-8)$$

最后通过简单的池化，得到所提算法的最终公式：

$$S = LCDV_{\alpha}^{\alpha} \times LCDV_r^{1-\alpha} \quad (3-9)$$

式中： α 为权重因子，用于确定两个部分的权重。

为了探究参数 α 值对图像质量评价最终结果的影响，本研究以 CSIQ 数据库中的模糊失真图像为实验对象。实验中，参数值在 $[0.05, 0.95]$ 区间内以 0.05 为步长进行取值，实验结果如图 3-5 所示。图中纵坐标轴表示 IQA 研究中常用的性能评价指标值，即 PLCC 和 SROCC。从实验结果中可以观察到，参数值的最优区间为 $[0.55, 0.75]$ 。在此区间内，PLCC 和 SROCC 值均表现出较高的相关性，

表明 IQA 模型的预测结果与人眼主观评价之间的一致性较好。因此，在后续模型计算中，参数值被设定为 0.65，以确保模型在图像质量评价中具有较高的准确性和可靠性。

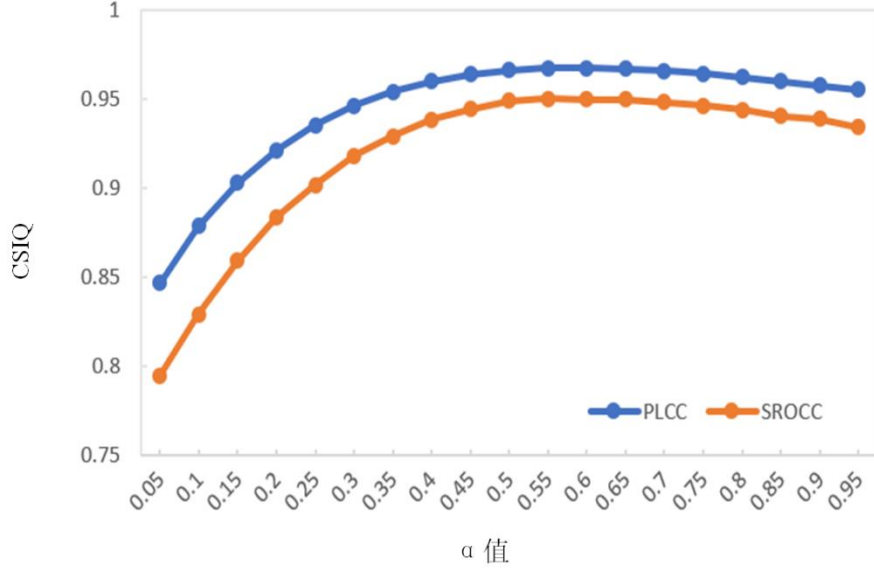


图 3-5 不同 α 值在 CSIQ 数据库测试的 PLCC 和 SROCC 值变化趋势

基于上述内容，本节所提评价算法的整体流程如图 3-6 所示。首先，将输入的局部 AOI 的 RGB 彩色图像转换至特定颜色空间，通过该转换过程，获得该颜色空间下的色差图，为后续的特征提取提供基础；然后基于生成的色差图，分别计算局部绝对色差和局部相对色差；最后，将计算得到的绝对色差图和相对色差图进行池化操作，通过聚合色差特征信息，最终生成图像质量评价分数。

3.3 实验验证

图 3-7 展示了从 CSIQ 数据库中选取的不同模糊程度失真图像的主观评分 (DMOS) 以及所提算法计算得到的客观质量分数 score。结果分析表明，随着图像模糊程度的加强，图像 DMOS 值逐渐增大，而由算法计算得到的分数逐渐减小，两者之间具有一定的负相关性，说明所提算法可有效区分测试图像质量变化。进一步说明所提算法能够有效反映图像质量的主观感知变化，具有较高的准确性和可靠性，能够为图像质量评价提供有力支持。

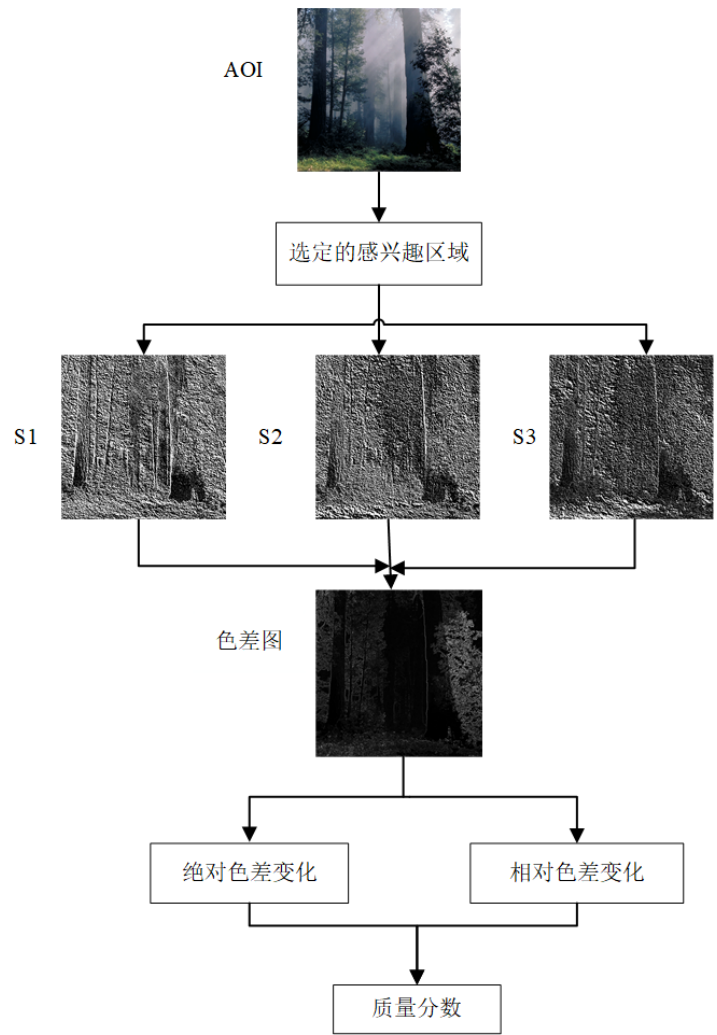


图 3-6 基于感兴趣区域的图像质量评价整体流程

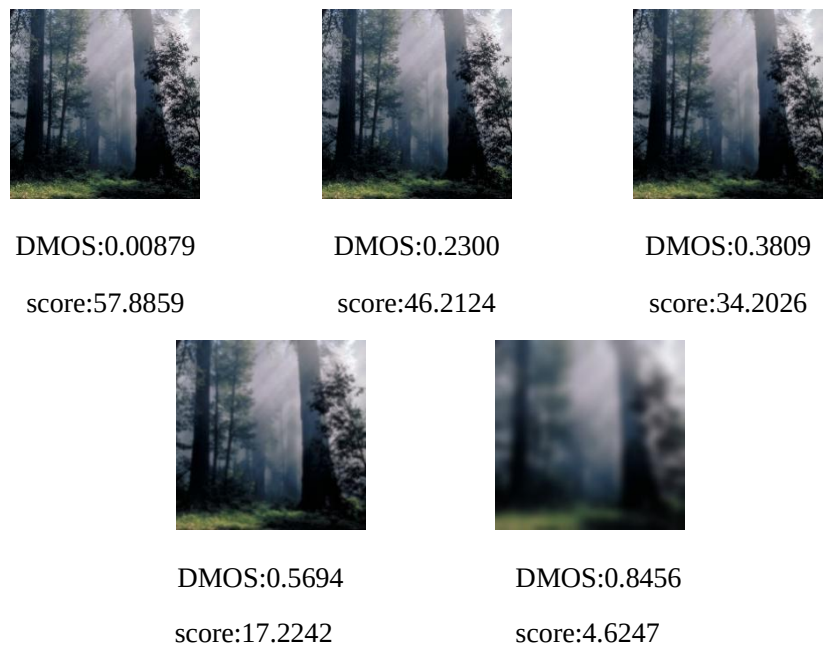


图 3-7 不同模糊程度失真典型图像及其主观分数和所提方法计算结果

为进一步验证所提模型的性能,本研究选取了四个公共数据库中的模糊失真图像进行实验,具体包括 TID2008、TID2013、CSIQ 和 LIVE 数据库。表 3-1 介绍了这四个数据库模糊图像的基本信息。这些数据库在 IQA 研究被广泛使用,具有较高的权威性和代表性。

表 3-1 多失真数据库模糊图像信息说明

数据库	参考图像数	模糊图像数	主观分数指标	图像尺寸
CSIQ	30	150	DMOS	512×512
LIVE	29	145	DMOS	——
TID2008	25	100	MOS	512×384
TID2013	25	125	MOS	512×384

3.3.1 颜色空间测试结果分析

在本研究中,为了确定所提算法的最佳颜色空间,本节针对模糊失真图像开展了不同颜色空间的对比测试。本次测试选取了 RGB、LAB、YIQ、LMN 四种颜色空间,并以 TID2008 数据库中的模糊失真图像作为测试样本,测试结果如表 3-2 所示,其中最佳结果数据已用加粗字体标出。在测试过程中,将图像全域默认作为感兴趣区域(AOI)进行分析。通过对测试结果的分析可知,YIQ 颜色空间在测得的 PLCC 和 SROCC 值均处于领先地位,同时其 RMSE 值最小。基于以上分析,本研究选择 YIQ 颜色空间作为所提算法的图像空间处理基础。在选定颜色空间后,所有待测图像均将被转换至 YIQ 颜色空间,进而开展相关计算,以获取所提算法的最终测试结果。

表 3-2 不同颜色空间测试结果

TID2008	PLCC	SROCC	RMSE
RGB	0.8040	0.9331	0.6978
YIQ	0.9363	0.9434	0.4123
LAB	0.9118	0.9177	0.4818
LMN	0.7382	0.7286	0.7917

3.3.2 AOI 尺寸选择的效果分析

在本节研究中,为了探究 AOI 尺寸对所提方法最终评价结果的影响,本研

究选取了不同尺寸的 AOI 图像作为测试对象。测试图像选自 TID2008 数据库中模糊失真图像，具体所选尺寸及对应的测试结果如表 3-3 所示，其中最佳结果数据已用加粗字体标出。

表 3-3 TID2008 数据库不同尺寸选择感兴趣区域测试结果

数据库	AOI 尺寸	384× 384	360× 360	352× 352	320× 320	288× 288	256× 256	224× 224	192× 192
TID2008	PLCC	0.9386	0.9385	0.9386	0.9385	0.9320	0.9252	0.9246	0.9213
	SROCC	0.9427	0.9440	0.9447	0.9458	0.9402	0.9334	0.9305	0.9260
	RMSE	0.4047	0.4053	0.4050	0.4052	0.4254	0.4453	0.4471	0.4564

由于 TID2008 数据库中模糊图像的尺寸为 512×384 ，本研究在进行感兴趣区域（AOI）尺寸测试时，基于常规思路将实验窗口形状设定为正方形，并将像素尺寸范围设定为 192×192 至 384×384 ，尺寸变化步长为 8×8 。通过对不同窗口大小的性能测试结果分析（如表 3-3 所示），发现性能指标在不同窗口尺寸下存在一定波动。其中，当窗口大小为 320×320 时，整体性能表现最佳。因此，针对 TID2008 数据库，本研究确定采用 320×320 的窗口大小进行后续实验。为进一步验证该方法的普适性，进一步对 LIVE、CSIQ 以及 TID2013 数据库也进行了类似的实验，测试不同尺寸实验窗口对性能的影响。根据实验结果，确定 LIVE、CSIQ 和 TID2013 数据库的最佳窗口大小分别为 288×288 、 512×512 和 384×384 。基于上述确定的参数设置，本文在四个数据库上进行了实验，并获得了所提算法的最终测试结果，如表 3-4 所示。

表 3-4 所提算法在四个人工失真数据库上的测试结果

数据库	PLCC	SROCC	RMSE
LIVE	0.9520	0.9437	4.8117
CSIQ	0.9661	0.9486	0.0739
TID2008	0.9385	0.9458	0.4052
TID2013	0.9502	0.9562	0.3890

3.3.3 实验结果与分析

在本节中，首先对所提出的方法与若干经典方法进行了对比分析，结果如表 3-5 所示。所涉及的经典方法包括 JNB^[52]、CPBD^[53]、SPRIASH^[27]、LPC_SI^[54]、

BIBLE^[55]、MLV^[19]、ARISM^[20]等。表 3-5 中列出了四个数据库的加权平均值（Weighted Average, W.A.）和直接平均值（Direct Average, D.A.）从而更好地对整体性能进行综合评估，其中根据数据库包含的模糊失真图像数量确定权重得到加权平均值。进一步为更清晰呈现各方法的性能差异，三个评价指标的前三名结果以粗体形式突出显示。从表 3-5 的结果可以看出，所提出的方法在四个数据库上的整体表现均较为出色，在 TID2008、TID2013 和 CSIQ 三个数据库上，该方法得到 PLCC 等三个评价指标的最佳排名。而对于 LIVE 数据库，其性能仅略低于前三名算法。此外，对于加权平均值和直接平均值，该方法三个评价指标均表现优异。其中，PLCC 和 SROCC 的直接平均值分别为 0.9517 和 0.9486，加权平均值则分别为 0.9530 和 0.9485，这些结果表明所提出的方法在整体性能上具有显著优势。

表 3-5 所提算法和一些经典 NR-IQA 算法的比较

数据库		JNB	CPBD	SPRIASH	LPC_SI	BIBLE	MLV	ARISM	所提算法
CSIQ	SROCC	0.7624	0.8847	0.9141	0.9071	0.9131	0.9247	0.9255	0.9486
	PLCC	0.8061	0.8818	0.9380	0.9159	0.9404	0.9489	0.9456	0.9661
	RMSE	0.1719	0.1370	0.1007	0.1166	0.0988	0.0917	0.0945	0.0739
LIVE	SROCC	0.7872	0.9182	0.9593	0.9389	0.9607	0.9316	0.9511	0.9437
	PLCC	0.8161	0.8955	0.9595	0.9181	0.9622	0.9429	0.9560	0.9520
	RMSE	10.824	8.3361	5.2750	7.4246	5.1022	6.2386	5.4939	4.8117
TID2008	SROCC	0.6667	0.8412	0.8869	0.8561	0.8915	0.8548	0.8548	0.9458
	PLCC	0.6931	0.8236	0.8891	0.8574	0.8929	0.8583	0.8428	0.9385
	RMSE	0.8634	0.6974	0.5483	0.6165	0.5393	0.6147	0.6447	0.4052
TID2013	SROCC	0.6902	0.8518	0.8927	0.8888	0.8988	0.8787	0.8980	0.9562
	PLCC	0.7114	0.8553	0.9004	0.8918	0.9051	0.8815	0.8953	0.9502
	RMSE	0.8914	0.6573	0.5519	0.5740	0.5392	0.5990	0.5651	0.3890
D.A.	SROCC	0.7266	0.8740	0.9133	0.8977	0.9160	0.8975	0.9074	0.9486
	PLCC	0.7567	0.8641	0.9218	0.8958	0.9252	0.9079	0.9099	0.9517
W.A.	SROCC	0.7336	0.8778	0.9163	0.9018	0.9188	0.9021	0.9124	0.9485
	PLCC	0.7644	0.8681	0.9256	0.8995	0.9286	0.9136	0.9166	0.9530

为进一步验证所提方法的有效性，将其与近年来其他研究者提出的方法进行了对比分析，涉及的方法包括文献[56]、HVS-MaxPol^[57]，Is^[28]，SOA^[21]，GCDV^[58]，

以及文献[59]的方法等。在本研究中,表 3-6 中仅列出了各方法的 PLCC 和 SROCC 值,并且把最优的两个结果以加粗形式突出显示。由于部分方法的源代码未开源,故其中部分数据参考了文献中的实验结果数据,因此有些算法只有部分数据库上的实验结果。从表 3-6 的结果来看,所列方法均未能在四个数据库中同时取得最佳结果。然而,深入分析数据发现,所提方法在四个数据库的性能比较中,相较于文献提出的方法,仍具有一定的优势。特别是在 TID2008 和 TID2013 数据库中,所提方法在三个评价指标上均处于领先地位。在另外两个数据库中,所提方法与最优方法之间的差距也相对较小。因此,可以认为所提方法相较于主要参考文献所提算法,整体性能仍具有一定优势。

表 3-6 所提算法与近几年 NR-IQA 算法的比较

数据库	评价指标	[56]	HVS- MaxPol	Is	SOA	GCDV	[51]	[59]	所提算法
CSIQ	SROCC	0.9500	0.9208	0.9152	0.9358	0.9528	0.9487	0.9270	0.9486
	PLCC	0.9666	0.9381	0.9145	0.9421	0.9683	0.9662	0.9290	0.9661
LIVE	SROCC	0.9448	0.9596	—	0.9622	0.9550	0.9407	0.9630	0.9437
	PLCC	0.9494	0.9569	—	0.9667	0.9566	0.9489	0.9540	0.9520
TID2008	SROCC	0.9419	0.8507	0.9552	—	0.9177	0.9403	0.9030	0.9458
	PLCC	0.9370	0.8533	0.9152	—	0.9118	0.9335	0.8960	0.9385
TID2013	SROCC	0.9605	0.8746	0.9499	0.9185	0.9403	0.9521	0.9330	0.9562
	PLCC	0.9540	0.8774	0.9365	0.9130	0.9368	0.9457	0.9240	0.9502
W.A.	SROCC	0.9495	0.9070	—	—	0.9415	0.9457	0.9339	0.9485
	PLCC	0.9531	0.9064	—	—	0.9466	0.9502	0.9284	0.9530
D.A.	SROCC	0.9493	0.9014	—	—	0.9415	0.9455	0.9315	0.9486
	PLCC	0.9518	0.9064	—	—	0.9434	0.9486	0.9258	0.9517

综合表 3-5 和表 3-6 的结果分析,本研究提出的方法通过结合色差和局部信息熵两个特征确定图像的感兴趣区域(AOI),并在该区域内完成色差信息的计算,从而对图像的质量进行有效评估。和部分经典方法进行比较,所提方法在其中三个数据库上的预测指标值排名前二,均能表现出显著优势。此外,在与近年来提出的方法的对比中,尽管所提方法在预测精度上与新方法存在一定差距,但差距较小,仍能够满足模糊图像质量评价的实际需求。经过分析,原因可能在于,模型构建过程中更多地关注图像局部信息,但全局视觉关注区域对图像质量感知

有一定影响,没有平衡好这两方面,从而导致所提方法在预测精度上的表现不佳。未来的研究可以进一步探索如何更好地结合局部与全局信息,以进一步提升算法的性能。

图 3-8 展示了所提方法与四个人工失真数据库中模糊图像主观评分的散点图及拟合曲线。通过图中拟合曲线,可以看出该方法预测的质量评分与主观得分具有良好的一致性,表明所提方法能够有效反映图像质量的主观感知。所有测试数据均在MATLAB R2016b环境下运行得到,确保了实验结果的可靠性和可重复性。

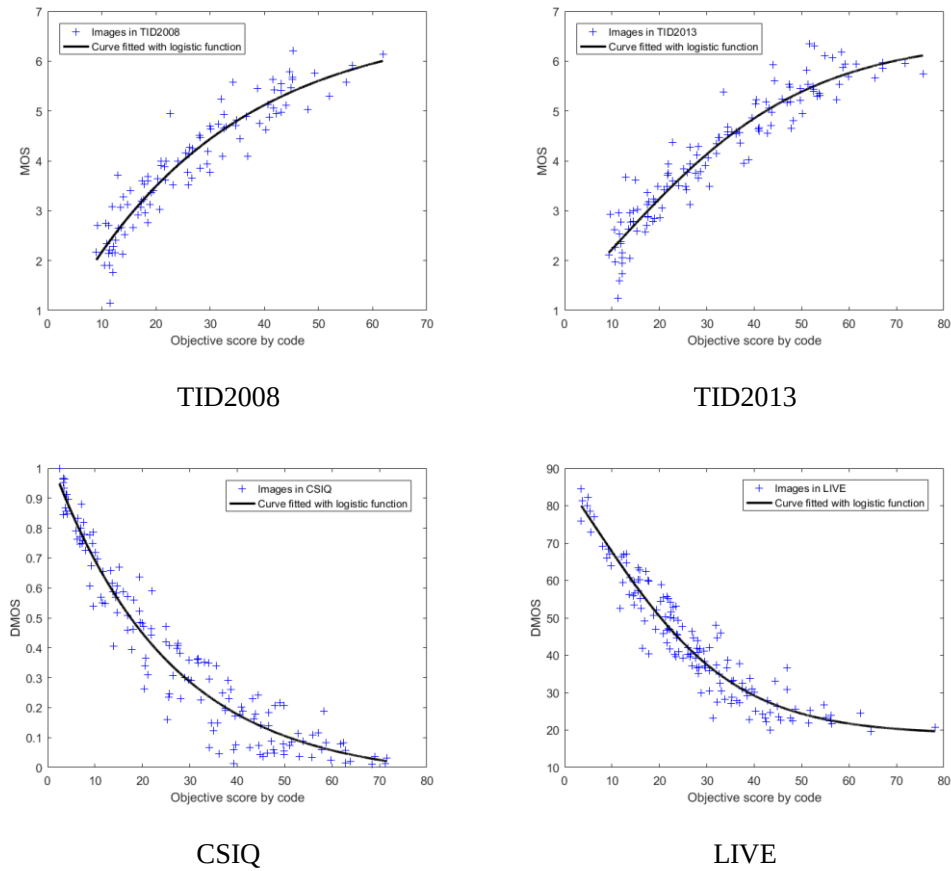


图 3-8 所提算法计算结果与四个人工失真数据库中主观评分的拟合曲线

3.3.4 CLIVE 数据库测试结果

上述方法在 TID2008 等人工失真数据库上测试,相关结果说明所提方法能够对人工模糊失真图像的质量评价具有较好性能。由于人工合成模糊失真具有一定规律,传统建模方法可以较好解决该问题。而真实场景的模糊失真为随机产生,

传统方法需要进一步验证。为了说明所提方法在真实场景失真数据库测试效果，选择 CLIVE 数据库作为测试对象。该数据库中模糊图像大小为 500×500 ，所提方法的 AOI 尺寸参照表 3-3 的最优测试结果对应的结果，选择 320×320 进行测试后续测试，相关结果如表 3-7 所示。得到的结果为 SROCC 值为 0.0320，PLCC 值为 0.0977，RMSE 值为 20.1993。从这三个数据可以看出，所提方法在在真实场景失真数据库测试时效果非常差。从结果可以发现基于传统方法构建的算法，无法很好解决真实场景的图像质量评估问题。当下，机器学习和深度学习的应用越来越广泛，可以尝试利用基于学习的方法解决该问题。

表 3-7 所提算法在 CLIVE 数据库上的计算结果

数据库	PLCC	SROCC	RMSE
CLIVE	0.0320	0.0977	20.1993

3.4 本章小结

在本章中，研究了一种基于感兴趣区域的无参考图像质量评价方法。该方法首先通过计算彩色图像的最大色差和多通道最大信息熵两项指标，确定待测图像的感兴趣区域（AOI）。随后，对选定的 AOI 图像进行局部绝对和相对色差计算，最终基于这些计算结果生成图像的质量评价分数。为了验证所提方法的优越性，本研究选取了近年来被广泛引用的以及近年提出的若干 NR-IQA 方法进行对比分析，并在四个大型图像质量评估数据库上对其拟合效果进行了验证。结果表明，所提方法的 SROCC 和 PLCC 加权平均值分别为 0.9485 和 0.9530，SROCC 和 PLCC 直接平均值分别为 0.9486 和 0.9517。结果表明，所提方法在预测精度和泛化能力方面均表现出色，能够有效评估人工失真模糊图像的质量。但是该方法在真实场景失真数据库测试时效果非常差，需要进一步提出能更好解决该问题的算法。

第4章 基于 PSO-GRNN 算法的图像质量评价

第三章研究了一种基于感兴趣区域的无参考图像质量评价方法,能够有效评估人工失真模糊图像的质量,但该方法在真实场景失真数据库测试时效果非常差。进一步针对真实模糊图像,为了获得精确的模糊图像质量评价,本章研究一种基于机器学习的图像质量客观评价方法。该方法选择并提取图像视觉显著性、色差和梯度特征,从图像中提取这三个方面的相关特征信息作为模糊图像评价的特征值;然后,建立一个基于粒子群优化的广义回归神经网络(PSO-GRNN),对上述提取的特征值进行训练,并得到最终的模糊图像评价结果。

4.1 算法模型构建

通过前期对图像质量评价研究现状的调研,发现神经网络已被广泛应用于图像的质量评价研究,相关网络结构也有很多种,其中比较典型的网络就是 GRNN。本章后续将以 GRNN 网络为基础,构建继续机器学习的图像质量评价方法。

4.1.1 广义回归神经网络 (GRNN)

GRNN 是一个具有动态网络结构^[60]的功能强大的回归工具,它是一种基于非参数估计的非线性回归的径向基函数神经网络。GRNN 的结构如图 4-1 所示,包括条件样本的输入层、网络模式层、求和层以及输出层共四层结构。

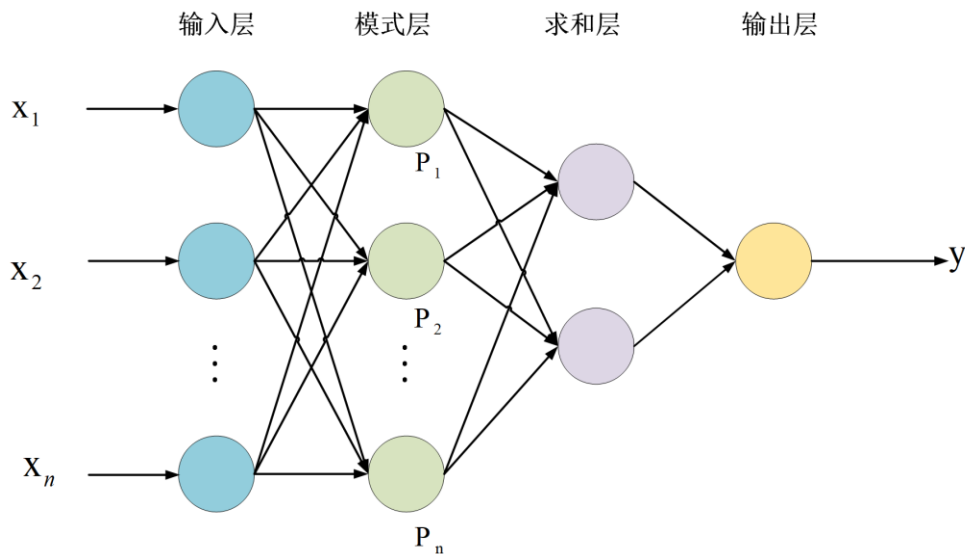


图 4-1 GRNN 结构图

其中输入层的神经元数目等于学习样本中输入向量的维数，然后直接将输入变量传递给模式层；模式层的神经元数目等于学习样本的数目 n 。每个神经元对应一个样本，其传递函数为高斯函数，用于计算输入向量与样本之间的距离；求和层包含两种类型的神经元，用于计算所有训练样本的响应输出；输出层的神经元数目等于学习样本中输出向量的维数 k ，每个神经元将求和层的输出相除，输出最终的预测结果。其中对应的网络输入和网络输出分别为：

$$Y(X) = \frac{\sum_{i=1}^n Y_i \exp(-D_i^2 / 2\sigma^2)}{\sum_{i=1}^n \exp(-D_i^2 / 2\sigma^2)} \quad (4-1)$$

式中， n 为样本中观测值的个数， X_i 和 Y_i 是样本的值， $D_i^2 = (X - X_i)^T (X - X_i)$ ， σ 为传递参数。 σ 的值越大，函数近似就越平滑。

4.1.2 PSO 算法原理

针对 GRNN 网络的参数 σ ，考虑选用粒子群优化算法（Particle Swarm Optimization, PSO）进行参数的优化。其中 PSO 算法的计算步骤如图 4-2 虚线框所示。具体的实现步骤如下：

（1）初始化总体：随机初始化每个粒子的位置(P_i)和速度(v_i)，最大迭代次数等。

（2）根据适应度函数计算粒子的适应度值，保存每个粒子的最佳位置 i ，并保存个体的最佳适应度值 $pbest_i$ 和总体的全局最佳位置 $gbest_i$ 。

（3）更新粒子的速度和位置，公式如下。

$$v_i^{t+1} = \omega \cdot v_i^t + c_1 r_1 (pbest_i - P_i^t) + c_2 r_2 (gbest_i - P_i^t) \quad (4-2)$$

$$P_i^{t+1} = P_i^t + v_i^t \quad (4-3)$$

其中 c_1 、 c_2 是学习因子，也称为加速度常数。 r_1 、 r_2 为 $[0,1]$ 范围内的均匀随机数。 ω 为惯性权重。 t 是迭代次数。首先计算每个粒子新的适应度值，然后比较该粒子新的值和历史上的最佳适应度值，若新的适应度值更优，则将更新为

该粒子的最佳位置，反之则保留历史数值；接着，比较每个粒子的最优位置对应的适应度值与整个种群的最优适应度值。若某个粒子的适应度值更优，则对种群的全局最佳位置和适应度值进行更新。

(4)确定搜索结果是否满足停止条件(达到最大迭代次数或满足精度要求)。如果满足停止条件，则输出最优值。否则，继续执行第二步并继续运行，直到满足停止条件。

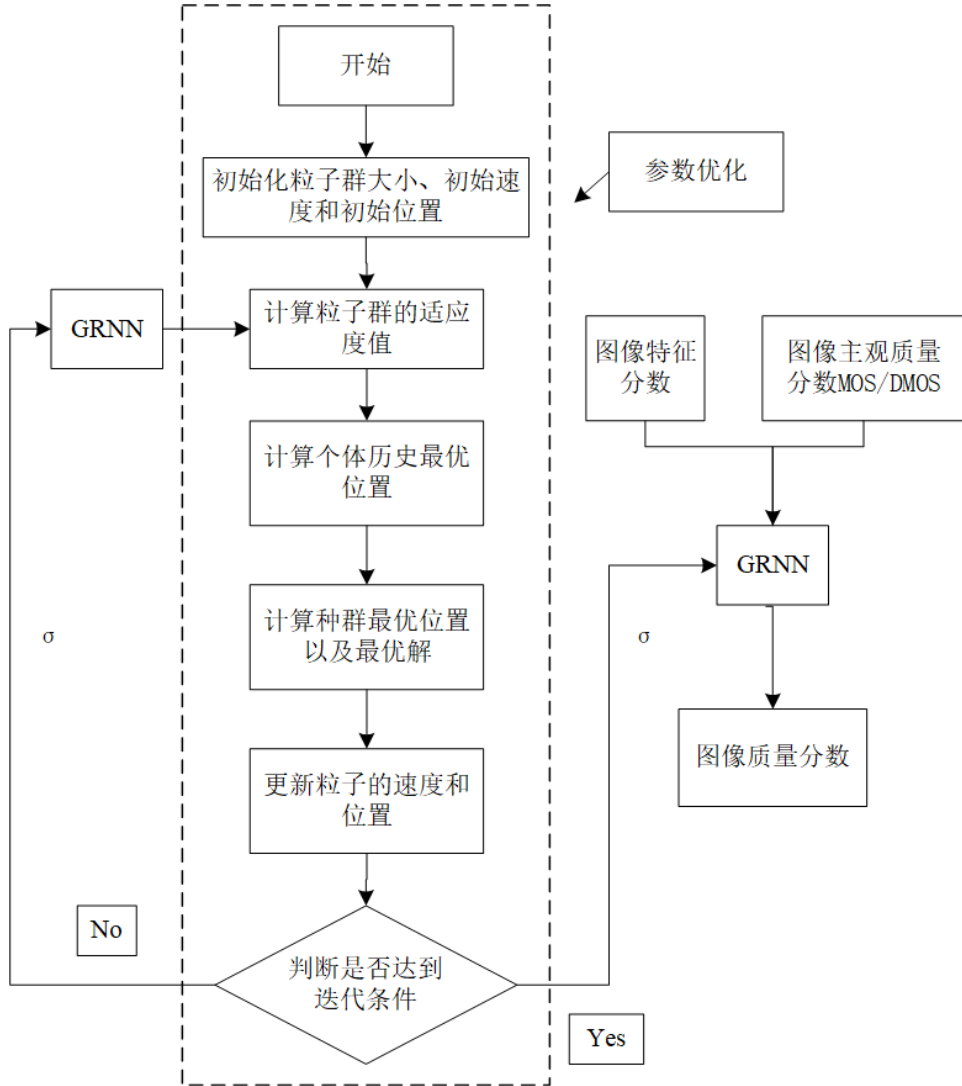


图 4-2 PSO-GRNN 算法的总体框架

4.1.3 PSO-GRNN 图像质量评价模型优化

GRNN 具有较强的非线性映射能力和灵活的网络结构，文献[61]中提到了 PSO-GRNN 预测模型。由此受到启发，针对模糊图像质量评价设计基于 PSO-GRNN 的 IQA 方法模型如图 4-2 所示。

具体而言，该方法首先对彩色模糊图像进行显著性、色差和梯度处理，提取图像的特征图。随后，从图像数据库中随机抽取 80% 的图像作为训练集，将其特征值及对应的主观质量评分（如 MOS 或 DMOS）输入到 GRNN 中进行训练。在此基础上，利用 PSO 算法对 GRNN 的平滑因子等关键参数进行优化，最终，使用经过优化训练的 PSO-GRNN 模型对数据库中剩余 20% 的图像进行质量评价。

在本研究中，PSO-GRNN 模型输入层中的神经元数为 n ，与待测失真图像特征值的数量一致。

4.2 特征图提取

由于 PSO-GRNN 模型输入层需要进行若干个值的输入，根据论文 2.2 节介绍，选择图像的梯度、色差以及视觉显著性特征作为网络的输入。

4.2.1 梯度特征提取

IQA 研究表明，灰度图像梯度是一种常用的特征。图像变化的幅度可以通过梯度进行说明，对于边缘部分，其梯度值发生明显变化；而平滑部分，梯度值变化幅度较小。图 4-3 显示了相同内容（a）和（b）的模糊图像及其对应的色差伪色图（c）和（d），可以看出不同模糊图像对应红色的矩形区域色差伪色图有明显差异。

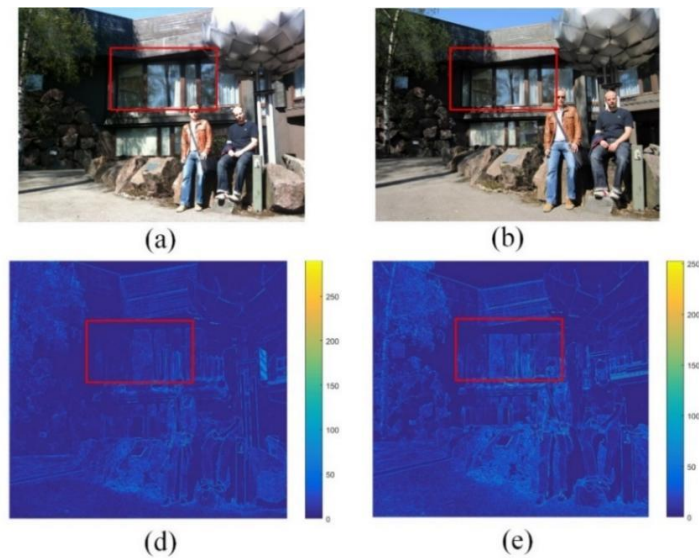


图 4-3 相同内容的模糊图像及其对应的梯度伪色图

4.2.2 色差特征提取

图 4-4 显示了在不同光照条件下，具有相同内容的图像对应的色差伪色图。经过比较可以看到不同光照下图像的色差图有明显的差异，红色矩形位置更能说明差异性。

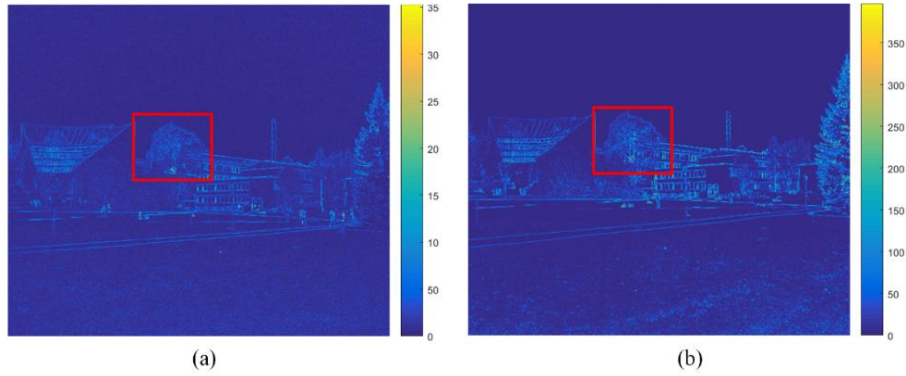


图 4-4 相同内容的图像对应的色差伪色图

4.2.3 视觉显著性特征提取

图像的视觉显著性能够有效反映局部区域的“突出”特征，这一特性已被广泛整合到图像质量评估的研究中。对于彩色模糊图像而言，场景中的重要区域因模糊而减少，导致图像的视觉显著性发生显著变化。这种变化与图像质量的感知密切相关，因此视觉显著性可以作为评估图像质量的重要依据。本研究提出的算法基于 SDSP 方法^[43]，通过视觉显著性提取图像中的重要区域。如图 4-5 所示，展示了在不同光照条件下相同内容的模糊和失真图像及其对应的视觉显著性图。从图中可以看出，该图能够准确地提取出图像中的重要区域。图中的 (c) 和 (d) 分别展示了 (a) 和 (b) 对应的显著性图，清晰地反映了特定光照条件下图像的模糊水平，且与人类视觉感知保持一致。

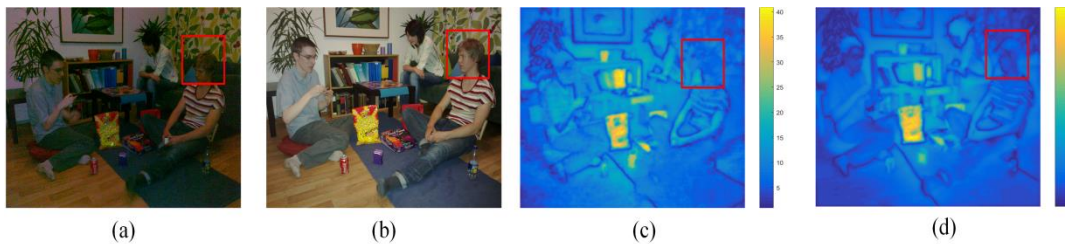


图 4-5 相同内容的图像及其对应的显著性图伪色图

4.3 模型性能对比实验

4.3.1 特征提取与组合

由于 GRNN 网络需要进行输入，那么选择哪些图像特征输入网络进行训练成为关键问题。最初尝试提取单个特征（如梯度）进行测试，但测试结果表明，利用单一特征测试的 PLCC 和 SROCC 值均较低，无法有效反映失真图像质量。因此，转变思路探索其他特征组合的可能性。经过多组不同特征组合的实验验证，最终确定采用视觉显著性、色差和梯度信息作为图像质量评价的特征。具体而言，对于模糊图像，首先提取其显著性图、色差图和梯度图。随后，针对这三个特征分别计算最大值（Max）、相对变化（RC）和方差（Var），从而得到 9 个特征值，如图 4-6 所示。这些特征值共同构成了图像的特征向量，并被输入到 PSO-GRNN 模型中进行质量评价。这一特征选择策略充分利用了视觉显著性、色差和梯度信息在图像质量感知中的重要作用。

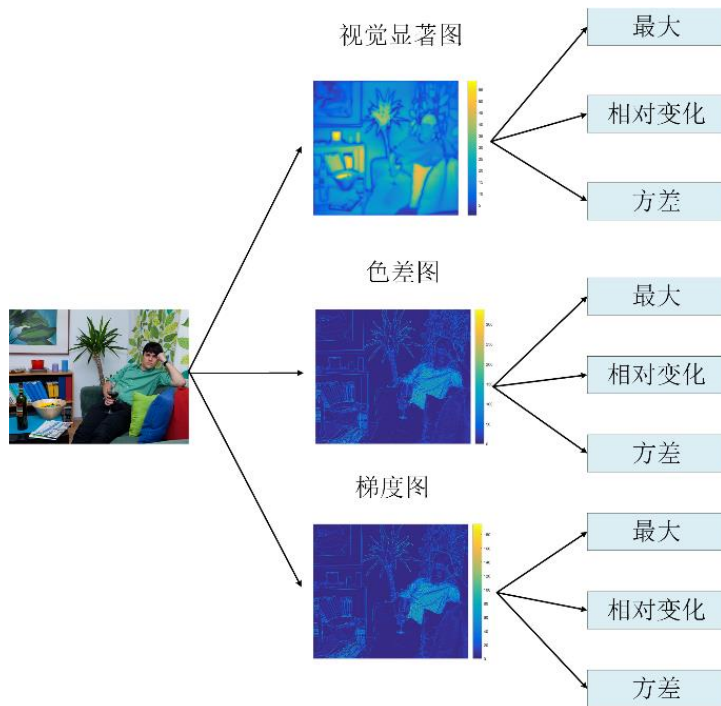


图 4-6 提取图像特征值的流程

对所得到的梯度、色差、视觉显著性图，其 Max 、 RC 和 Var 特征值的计算公式为：

$$Max = \max(M(i, j)) \quad (4-4)$$

$$RC = \frac{\max(M(i, j)) - \min(M(i, j))}{\text{mean}(M(i, j))} \quad (4-5)$$

$$Var = \frac{\left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right]}{n} \quad (4-6)$$

式中, $x_1, x_2, \dots, x_n$ 表示每个特征图的像素值, $\bar{x}$ 表示一个特征图的平均像素值, n 表示像素总数。

表 4-1 不同的特征组合的性能

数据库	特征组合	PLCC	SROCC
BID	Max	0.828	0.825
	RC	0.802	0.788
	Var	0.852	0.850
	Max + RC	0.846	0.844
	Max + Var	0.886	0.877
	RC + Var	0.877	0.869
	Max + RC + Var	0.890	0.885
CID2013	Max	0.902	0.892
	RC	0.843	0.828
	Var	0.917	0.909
	Max + RC	0.902	0.888
	Max + Var	0.925	0.915
	RC + Var	0.922	0.916
	Max + RC + Var	0.924	0.913
CLIVE	Max	0.834	0.821
	RC	0.785	0.769
	Var	0.851	0.858
	Max + RC	0.866	0.857
	Max + Var	0.871	0.863
	RC + Var	0.861	0.855
	Max + RC + Var	0.873	0.867

在 BID、CID2013 及 CLIVE 三个数据库上, 测试了 3 个特征值、6 个特征值以及 9 个特征值共 7 种不同的特征值组合, 结果如表 4-1 所示, 其中最佳结果用

加粗字体标出。从表 4-1 中结果可以看出，三个数据库上所有特征组合的 PLCC 和 SROCC 值均接近或高于 0.8，表明所提方法在不同特征组合下均能获得较为理想的性能。进一步分析表明，将视觉显著性、色差和梯度信息相结合，并分别计算其最大值、相对变化和方差所得到的 9 个特征值组合，在图像质量评价测试中表现最为出色。此外，通过对比不同特征组合的实验数据可以发现，方差在特征组合中的贡献略大于最大值和相对变化。这表明方差在反映图像质量特征的差异性和复杂性方面可能更具优势，能够为模型提供更丰富的信息，从而在一定程度上提升模型对图像质量评价的精度。

4.3.2 算法优化效果验证

为了实现 GRNN 的最佳性能，确定理想的变量值 σ 是关键。在 IQA 的研究中，通常采用控制变量的方法确定参数值。然而，该方法选择了自适应优化方法，以获得更好的性能。在表中，在 BID 数据库上测试了三种比较成熟的自适应优化方法，即果蝇优化算法（Fruit fly Optimization Algorithm, FOA）、萤火虫算法（Firefly Algorithm, FA）和粒子群优化算法（PSO）。这三种自适应优化算法都受到自然界中动物的某些行为启发，FOA 源于果蝇的觅食行为，主要利用果蝇的嗅觉和视觉特性；FA 算法受萤火虫发光吸引行为启发，以亮度和吸引力为核心机制；而 PSO 受鸟群觅食启发，强调粒子间的信息共享和协作。此外，还测试了一种不采用自适应优化方法的 GRNN，共四种进行测试对比，结果如表 4-2 所示，最佳结果在表中以粗体突出显示。根据相应的评价标准值(SROCC 和 PLCC)，PSO 的性能优于其他方法。因此，本研究选择了 PSO-GRNN 作为所提出的 IQA 方法的主要框架。

表 4-2 PSO-GRNN 使用不同的自适应优化方法在 BID 数据库上的性能

数据库	评价标准	GRNN	FOA-GRNN	FA-GRNN	PSO-GRNN
BID	SROCC	0.880	0.880	0.876	0.885
	PLCC	0.885	0.887	0.883	0.890

4.4 实验结果与分析

4.4.1 多场景验证分析

在三种自然失真数据库中，其中 CID2013 数据库由 6 个数据集（I-VI）组成，每个数据集包含 6 个不同的场景。其具体场景分类如表 4-3 所示：

表 4-3 CID2013 数据库不同场景介绍

组别	拍摄环境 亮度	拍摄距离	场景描述	示例图像	图像集
1	2	0.5	Close-up in dark lighting conditions		I-VI
2	100	1.5	Close-up in typical indoor lighting conditions		I-VI
3	10	4.0	Small group in dim lighting conditions		I-VI
4	1000	1.5	Studio image		I-IV
5	>3400	3.0	Small group in cloudy bright to sunny lighting conditions		I-V
6	>3400	>50	Close-up in high-dynamic range lighting conditions		I-VI
7	>3400	3.0	Small group in cloudy bright to sunny lighting conditions (~3x optical or digital zoom)		VI
8	>3400(outdoors) and<100(in doors)	1.5	Close-up in high-dynamic range lighting conditions		V, VI

对 CID2013 数据库中不同组的图像进行测试，以验证所提方法对不同光照

条件下相同内容的图像的评价效果，实验结果见表 4-4。

表 4-4 在 CID2013 上的不同场景下的测试结果

场景	SROCC	PLCC	情景	SROCC	PLCC
IS_I_C01	0.789	0.801	IS_I_C02	0.878	0.934
IS_II_C01	0.709	0.860	IS_II_C02	0.781	0.745
IS_III_C01	0.979	0.991	IS_III_C02	0.772	0.741
IS_IV_C01	0.908	0.890	IS_IV_C02	0.955	0.997
IS_V_C01	0.902	0.890	IS_V_C02	0.968	0.990
IS_VI_C01	0.866	0.938	IS_VI_C02	0.928	0.937
IS_I_C03	0.989	0.991	IS_I_C04	0.964	0.956
IS_II_C03	0.841	0.934	IS_II_C04	0.968	0.948
IS_III_C03	0.985	0.990	IS_III_C04	0.972	0.971
IS_IV_C03	0.960	0.960	IS_IV_C04	0.908	0.968
IS_V_C03	0.993	0.999	IS_VI_C07	0.849	0.967
IS_VI_C03	0.966	0.991	IS_V_C08	0.965	0.984
IS_I_C05	0.884	0.930	IS_I_C06	0.964	0.984
IS_II_C05	0.877	0.836	IS_II_C06	0.775	0.733
IS_III_C05	0.844	0.871	IS_III_C06	0.791	0.807
IS_IV_C05	0.977	0.996	IS_IV_C06	0.952	0.969
IS_V_C05	0.975	0.961	IS_V_C06	0.765	0.833
IS_VI_C08	0.804	0.968	IS_VI_C06	0.743	0.983

共检验 36 组，经分析，50%（20 组）的拟合结果在 0.90 以上，75%（28 组）的拟合结果在 0.80 以上。上述检验的最低 SROCC 和 PLCC 结果分别为 0.709 和 0.732。在这部分实验中，各组的测试图像的内容相同，但光照条件存在差异。从试验数据中可以得出结论，所提出的方法在不同的光照条件下，对 IQA 具有良好的性能和稳定性。

根据表 4-4 的结果，图 4-7 显示了不同场景下的 SROCC 和 PLCC 的箱线图。箱线图是用于显示一组数据分散情况的统计图，可以通过这种图直观的探索数据特征。从图 4-7 可以看出，该方法在场景 3 和场景 4 中表现出更好的性能，这些场景都是主题光照度在 10 到 1000 勒克斯之间的室内场景。此外，该方法在场景 3 和场景 4 中也表现出了稳定的性能。对于其他场景，所提出的方法也产生了类

似的性能，这进一步证明了该方法在不同光照条件下的普适性和可靠性。

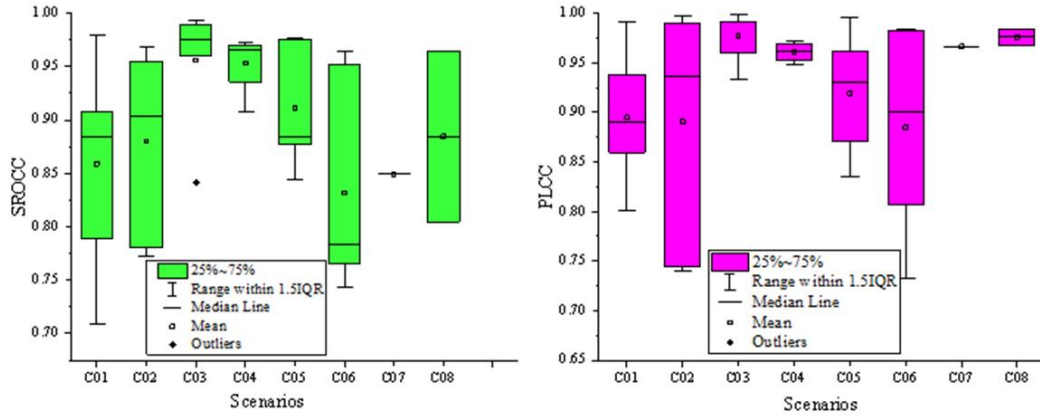


图 4-7 CID2013 不同场景的箱线图

4.4.2 人工失真数据库实验结果

在本节中，首先针对人工失真模糊图像，对 TID2008、TID2013 和 CSIQ 和 LIVE 数据库的模糊图像提取 9 个相应特征值，然后进行测试，以获得每个图像的预测分数。接着将预测的得分与图像的主观得分进行线性拟合，得到相应的 PLCC、SROCC 和 RMSE 值，以评估预测结果与主观评分的一致性。为了确保结果的稳定性和可靠性，对每个数据库进行了 20 次独立的检验，并以这 20 次检验结果的平均值作为该数据库的最终拟合结果，并与近五年所提出的一些算法进行比较，其中所提算法 1 是前面第三章提出的模型方法，所提算法 2 是基于 PSO-GRNN 方法。相关结果如表 4-5 所示，其中较好的两个结果粗体表示。

从表 4-5 中可以看出，该模型针对人工失真模糊图像的预测效果较好，能够有效反映图像质量的主观感知。然而，图像模糊失真不仅限于人为生成的情况，自然失真图像的复杂性更高，其失真类型和程度往往更加多样化。因此，进一步思考该模型在自然失真图像上的表现具有重要的研究意义。

表 4-5 PSO-GRNN 与近五年 NR-IQA 算法在人工失真数据库上的比较

数据库	评价指标	[56]	HVS-MaxPol	Is	SOA	GCDV	[51]	[59]	所提算法 1	所提算法 2
CSIQ	SROCC	0.9500	0.9208	0.9152	0.9358	0.9528	0.9487	0.9270	0.9486	0.9734
	PLCC	0.9666	0.9381	0.9145	0.9421	0.9683	0.9662	0.9290	0.9661	0.9819
LIVE	SROCC	0.9448	0.9596	—	0.9622	0.9550	0.9407	0.9630	0.9437	0.9719
	PLCC	0.9494	0.9569	—	0.9667	0.9566	0.9489	0.9540	0.9520	0.9778
TID2008	SROCC	0.9419	0.8507	0.9552	—	0.9177	0.9403	0.9030	0.9458	0.9621
	PLCC	0.9370	0.8533	0.9152	—	0.9118	0.9335	0.8960	0.9385	0.9622
TID2013	SROCC	0.9605	0.8746	0.9499	0.9185	0.9403	0.9521	0.9330	0.9562	0.9713
	PLCC	0.9540	0.8774	0.9365	0.9130	0.9368	0.9457	0.9240	0.9502	0.9690

4.4.3 自然失真模糊图像质量评价算法性能分析

在本节中，针对 BID、CID2013 和 CLIVE 数据库中的真实模糊图像，提取了 9 个相应的特征值，并计算每个图像的预测分数。随后，将预测分数与图像的主观评分进行线性拟合，以获得相应的 PLCC 和 SROCC 值。为了确保结果的可靠性，仍对每个数据库进行了 20 次独立检验，并以这 20 次检验结果的平均值作为整个数据库的最终拟合结果。相关结果如表 4-6 所示。大多数研究基本上测试都是基于 TID2008 等传统四个数据库的验证，采用真实失真数据库进行验证测试的比较少，另外，部分方法没有源代码没有开源，无法对真实数据库进行计算测试。所以选择了表 4-6 中的方法进行对比。表 4-6 中列出了所提方法与其他一些方法的 PLCC 和 SROCC 值对比，其中最佳结果以黑体字突出显示。对比方法涵盖了空间域、频域、机器学习和深度学习等多个领域。其中 RISE^[62]，SR^[63]，Yu's CNN^[64]，DIQA^[65]，SSEQ^[66]，SFA^[67]，DB-CNN^[68]，DIVINE^[69]和 NIQE^[70]是基于学习的方法，而 MLV^[19]，BIBLE^[55]，BISHARP^[27]和 GCDV^[51]是基于空间域和频率域的方法。

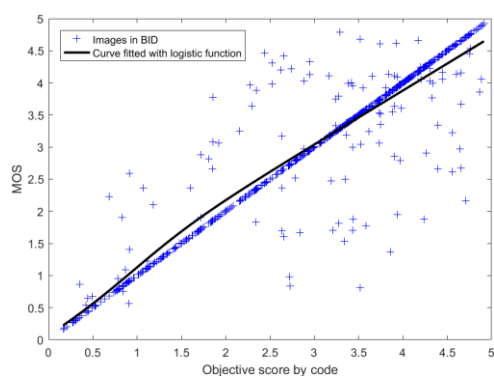
从实验结果可以看出，所提算法在 BID、CID2013 和 CLIVE 数据库上的 PLCC 和 SROCC 值均高于 0.85，表明该方法在评估真实模糊图像质量方面具有较高的准确性和稳定性。总体而言，基于学习的算法在处理真实模糊和人工失真图像的质量评估任务时，相较于传统的空间域和频域方法，展现出显著的优越性。

特别是基于 PSO-GRNN 的方法，在多个数据库上的表现优于其他基于网络结构的方法，如 Yu's CNN、DIQA、SFA 和 DB-CNN。这一结果表明，全连接的 GRNN 在处理 IQA 问题时具有显著优势。

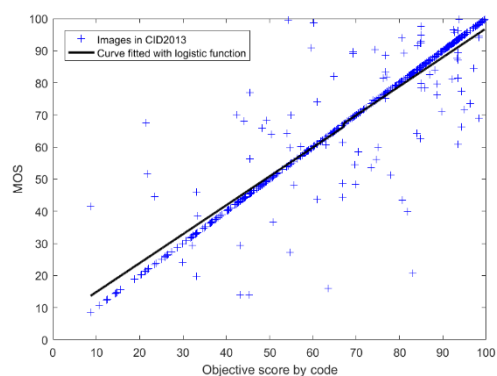
在 BID、CID2013 和 CLIVE 数据库上，针对所提方法共进行了 20 次独立实验，每次实验均获得了 PLCC 和 SROCC 值。最终，以这 20 次实验结果的平均值作为该方法的最终 PLCC 值和 SROCC 值，相关结果如表 4-1 所示。为了进一步验证方法的可靠性和泛化能力，对这些数据库进行了随机抽样检验。所提方法的预测结果与数据库中的主观评分之间的散点图如图 4-8 所示。从图 4-8 中可以看出，所提方法在评估真实模糊图像质量方面表现不错，其回归曲线与主观观测值之间的相关性较高，表明该方法能够反映图像质量的主观感知。

表 4-6 PSO-GRNN 与其他 NR-IQA 算法在自然失真数据库上的比较

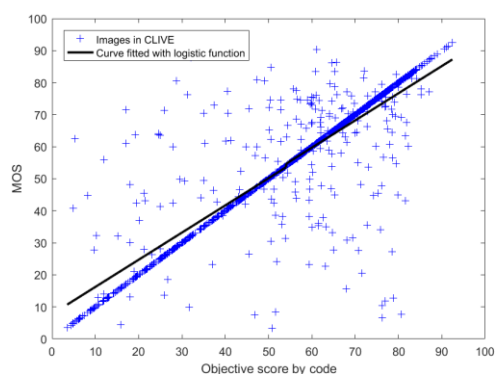
数据库	BID		CID2013		CLIVE	
评价指标	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
BISHARP	0.356	0.307	0.678	0.681	-	-
BIBLE	0.392	0.361	0.698	0.687	0.515	0.427
MLV	0.375	0.317	0.689	0.621	0.400	0.339
GCDV	0.338	0.294	0.681	0.596	0.405	0.334
RISE	0.602	0.584	0.793	0.769	0.555	0.515
SR	0.415	0.467	0.621	0.634	-	-
Yu's CNN	0.560	0.557	0.715	0.704	0.501	0.502
DIQA	0.506	0.492	0.720	0.708	0.704	0.703
SSEQ	0.604	0.581	0.689	0.676	-	-
SFA	0.840	0.826	-	-	0.833	0.812
DB-CNN	0.471	0.464	0.686	0.672	0.869	0.851
DIVINE	0.506	0.489	0.499	0.477	0.558	0.509
NIQE	0.471	0.469	0.693	0.633	0.478	0.421
所提方法	0.890	0.885	0.924	0.913	0.873	0.867



BID



CID2013



CLIVE

图 4-8 BID、CID2013 和 CLIVE 数据库的散点图和拟合曲线

4.5 本章小结

这里研究了一种基于机器学习的无参考图像清晰度客观评价方法。该方法选择视觉显著性、色差和梯度信息作为图像特征，从图像中提取这三个方面的相关特征信息作为不同光照成像条件下模糊失真图像评价的特征值。然后，建立一个基于粒子群优化的广义回归神经网络（PSO-GRNN），对上述提取的特征值进行训练，并得到最终的模糊图像评价结果。分别在多失真数据库和自然失真数据库上进行验证，实验结果表明，所提方法在预测准确度上有优良的表现。尤其针对不同成像条件下的模糊失真图像，该质量评价算法具有良好的性能。

第 5 章 总结与展望

5.1 工作总结

信息技术的不断发展，图像在人类生活中也发挥着愈来愈重要的作用。然而图像从采集到传输再到编码等过程中，不可避免的遭受到各种各样的失真，因此提高无参考图像质量评价方法的性能有重要研究意义。本文主要研究针对模糊图像的无参考图像质量评价方法。考虑人眼注意力机制以及不同图像特征组合，通过理论分析和实验测试，研究了两种有效的 IQA 方法。

(1) 信息熵是表征图像的常用特征，而色差指标能够表征彩色图像的信息，利用这两个特征，结合人眼视觉注意力机制，研究了基于感兴趣区域的无参考图像质量评价方法。首先构建最大局部色差和多通道信息熵两个指标，确定整幅彩色图像的兴趣区域；然后对上述局部 AOI 图像分别进行绝对与相对色差变化计算，获得质量评价分数。为了验证该方法的性能，对四个常用图像质量数据库中的模糊失真图像进行对比实验测试，分析其预测精度和泛化性，该算法表现良好。

(2) 研究了一种基于机器学习的图像清晰度客观评价方法。该方法选择视觉显著性、色差和梯度信息作为图像特征，从图像中提取这三个方面的相关特征信息作为不同光照成像条件下模糊图像评价的特征值。然后，建立一个基于粒子群优化的广义回归神经网络 (PSO-GRNN)，输入上述提取的特征值进行训练，得到最终的模糊图像评价结果。在多失真数据库和自然失真数据库上进行实验验证，结果表明，所提方法在预测准确度有优良的表现。特别是针对不同成像条件下的模糊失真图像，该质量评价算法具有良好的性能。

5.2 不足之处与展望

图像质量评价是图像处理领域中重要的一个方面，具有广阔的研究意义和应用前景。本文提出的两种算法模型虽然经过实验验证预测效果优良，但是仍有提升的空间。

(1) 对于基于感兴趣区域的无参考图像质量评价方法，利用色差和信息熵

两个指标选取感兴趣区域，可以对这两个特征进行改进优化，除此之外，本文利用方形框选取感兴趣区域，未来可以探索提取图像中心圆形区域作为 AOI，可能更加符合人眼视觉感知。

(2) 深度学习在图像处理领域取得了重大突破，对于本文基于 PSO-GRNN 算法的无参考图像质量评价，需要先提取图像相关特征，这导致了算法的计算效率不高，故而有优化的空间。除此之外，针对框架中优化算法，考虑能否在 PSO 等成熟优化算法的基础上，参考近几年的优化算法做进一步的研究。目前对基于深度学习的图像质量评价方法的研究日益增多，后续研究中进一步深入探索构建基于深度学习的模型。

对于今后的进一步研究，有必要深入探究人类视觉系统机理，将视觉注意力机制、视觉认知模型等融入无参考图像质量评价研究中，并结合人工智能、深度学习等新技术和新方法使评价结果与人类主观感受更一致，进而为图像增强等应用提供更精准指导，满足不断增长的应用需求。

参考文献

- [1] 韩翰, 卓力, 张菁, 等. 基于深度学习的无参考图像质量评价综述[J]. 测控技术, 2022, 41(04): 1-10.
- [2] 卢鹏, 刘楷贇, 邹国良, 等. 基于多特征融合和卷积神经网络的无参考图像质量评价[J]. 液晶与显示, 2022, 37(01): 66-76.
- [3] 陈健, 李诗云, 林丽, 等. 模糊失真图像无参考质量评价综述[J]. 自动化学报, 2022, 48(3): 689-711.
- [4] Gu K, Zhou J, Qiao J F, et al. No-reference quality assessment of screen content pictures[J]. IEEE Transactions on Image Processing, 2017, 26(8): 4005-4018.
- [5] Saad M A, Bovik A C, Charrier C. Blind prediction of natural video quality[J]. IEEE Transactions on Image Processing, 2014, 23(3): 1352-1365.
- [6] 高敏娟, 党宏社, 魏立力, 等. 全参考图像质量评价回顾与展望[J]. 电子学报, 2021, 49(11): 2261-2272.
- [7] Wang Z, 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.
- [8] Pei S C, Chen L H. Image quality assessment using human visual DOG model fused with random forest[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3282-3292.
- [9] Gao F, Wang Y, Li P, et al. Deepsim: Deep similarity for image quality assessment[J]. Neurocomputing, 2017, 257: 104-114.
- [10] 江本赤, 卞仕磊, 史晨阳, 等. 基于色貌尺度相位一致性的全参考图像质量评价[J]. 光学精密工程, 2023, 31 (10): 1509-1521.
- [11] Wu J, Liu Y, Li L, et al. Attended visual content degradation based reduced reference image quality assessment[J]. IEEE Access, 2018, 6: 12493-12504.
- [12] Hu B, Wang S, Gao X, et al. Reduced-reference image deblurring quality assessment based on multi-scale feature enhancement and aggregation[J]. Neurocomputing, 2023, 547: 126378.
- [13] Anju M, Mohan J. Deep image compression with lifting scheme: Wavelet transform domain based on high-frequency subband prediction[J]. International Journal of Intelligent Systems, 2022, 37(3): 2163-2187.

- [14] Zhan Y, Zhang R. No-reference image sharpness assessment based on maximum gradient and variability of gradients[J]. IEEE Transactions on Multimedia, 2017, 20(7): 1796-1808.
- [15] Li L, Lin W, Wang X, et al. No-reference image blur assessment based on discrete orthogonal moments[J]. IEEE Transactions on Cybernetics, 2015, 46(1): 39-50.
- [16] Zhang K, Huang D, Zhang B, et al. Improving texture analysis performance in biometrics by adjusting image sharpness[J]. Pattern Recognition, 2017, 66: 16-25.
- [17] Liu Z, Hong H, Gan Z, et al. An improved method for evaluating image sharpness based on edge information[J]. Applied Sciences, 2022, 12(13): 6712.
- [18] Chen M, Yu L, Zhi C, et al. Improved faster R-CNN for fabric defect detection based on Gabor filter with Genetic Algorithm optimization[J]. Computers in Industry, 2022, 134: 103551.
- [19] Bahrami K, Kot A C. A fast approach for no-reference image sharpness assessment based on maximum local variation[J]. IEEE Signal Processing Letters, 2014, 21(6): 751-755.
- [20] Gu K, Zhai G, Lin W, et al. No-reference image sharpness assessment in autoregressive parameter space[J]. IEEE Transactions on Image Processing, 2015, 24(10): 3218-3231.
- [21] 沈飞鹏, 朱彤, 张赫男, 等. 基于显著性目标分类的无参考模糊图像质量评价[J]. 激光与光电子学进展, 2021, 58 (22): 237-247.
- [22] Zhang F, Roysam B. Blind quality metric for multidistortion images based on cartoon and texture decomposition[J]. IEEE Signal Processing Letters, 2016, 23(9): 1265-1269.
- [23] De K, Masilamani V. No-reference image sharpness measure using discrete cosine transform statistics and multivariate adaptive regression splines for robotic applications[J]. Procedia Computer Science, 2018, 133: 268-275.
- [24] Hu Z, Yang G, Du Z, et al. No-reference image quality assessment based on global awareness[J]. PloS one, 2024, 19 (10): e0310206.
- [25] Bae S H, Kim M. A novel image quality assessment with globally and locally consilient visual quality perception[J]. IEEE Transactions on Image Processing. 2016, 25(5): 2392-2406.
- [26] Baig M A, Moinuddin A A, Khan E, et al. DFT-based no-reference quality assessment of blurred images[J]. Multimedia Tools and Applications, 2022, 81(6): 7895-7916.
- [27] Gvozden G, Grgic S, Grgic M. Blind image sharpness assessment based on local contrast map

- p>statistics[J]. Journal of Visual Communication and Image Representation, 2018, 50 145-158.
- [28] 刘玉玺, 郑秋梅, 高绍姝, 等. 基于人眼视觉敏感特性的图像清晰度评价[J]. 光学技术, 2020, 46(03): 354-360.
- [29] Chen J, Li S, Lin L, et al. No-reference blurred image quality assessment method based on structure of structure features[J]. Signal Processing: Image Communication, 2023, 118: 117008.
- [30] Liu L, Gong J, Huang H, et al. Blind image blur metric based on orientation-aware local patterns[J]. Signal Processing. Image Communication, 2020, 80: 115654.
- [31] Zhu M, Ge D. Image quality assessment based on deep learning with FPGA implementation[J]. Signal Processing. Image Communication, 2020, 83: 115780.
- [32] Li D, Jiang T, Jiang M. Exploiting high-level semantics for no-reference image quality assessment of realistic blur images[C]//Proceedings of the 25th ACM international conference on Multimedia. 2017: 378-386.
- [33] Zhang W, Ma K, Yan J, et al. Blind image quality assessment using a deep bilinear convolutional neural network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(1): 36-47.
- [34] Ji P, Liu C, Chen H. No-reference image quality assessment via a dual-branch residual network[J]. IET Image Processing, 2024, 18 (7): 1719-1732.
- [35] Jia H Z, Zhang L, Wang T H. Contrast and visual saliency similarity-induced index for assessing image quality[J]. IEEE Access, 2018: 65885-65893.
- [36] Liu A, Lin W, Narwaria M. Image quality assessment based on gradient similarity[J]. IEEE Transactions on Image Processing, 2012: 1500-1512.
- [37] 闫乐乐, 李辉, 邱聚能, 等. 基于区域对比度和 SSIM 的图像质量评价方法[J]. 应用光学, 2015, 36(1): 58-63.
- [38] 姚军财. 基于人眼对比度敏感视觉特性的图像质量评价方法[J]. 液晶与显示, 2011, 26(3): 390-396.
- [39] Mozhaeva A I, Vlasuyk I V, Potashnikov A M, et al. The method and devices for research the parameters of the human visual system to video quality assessment[C]//Systems of Signals

- Generating and Processing in the Field of on Board Communications. IEEE, 2021: 1-5.
- [40] 冯丹青, 陈亮. 基于多特征信息融合的彩色图像复杂度研究[J]. 通信技术, 2019, 52(09): 2136-2142.
- [41] Jeon, J, Yoon, Kim D, et al. Fully digital auto-focusing system with automatic focusing region selection and point spread function estimation[J]. IEEE Transactions on Consumer Electronics, 2010, 56(3): 1204-1210.
- [42] Niu Y, Zhang H, Guo W, et al. Image quality assessment for color correction based on color contrast similarity and color value difference[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 28(4): 849-862.
- [43] Zhang L, Shen Y, Li H. VSI: A visual saliency-induced index for perceptual image quality assessment[J]. IEEE Transactions on Image Processing, 2014, 23(10): 4270-4281.
- [44] Ponomarenko N, Lukin V, Zelensky A, et al. TID2008-a database for evaluation of full-reference visual quality assessment metrics[J]. Advances of Modern Radioelectronics, 2009, 10(4): 30-45.
- [45] Ponomarenko N, Ieremeiev O, Lukin V, et al. Color image database TID2013: Peculiarities and preliminary results[C]//European workshop on visual information processing (EUVIP). IEEE, 2013: 106-111.
- [46] Sheikh H. LIVE image quality assessment database release 2[J]. <http://live.ece.utexas.edu/research/quality>, 2005.
- [47] Ciancio A, Da Silva E A B, Said 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.
- [48] Virtanen T, Nuutinen M, Vaahteranoksa M, et al. CID2013: A database for evaluating no-reference image quality assessment algorithms[J]. IEEE Transactions on Image Processing, 2015, 24(1): 390-402.
- [49] Ghadiyaram D, Bovik A C. Massive online crowdsourced study of subjective and objective picture quality[J]. IEEE Transactions on Image Processing, 2015, 25(1): 372-387.
- [50] Lee D, Plataniotis K N. Towards a novel perceptual color difference metric using circular

- p>processing of hue components[C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014: 166-170.
- [51] Shi C, Lin Y. No reference image sharpness assessment based on global color difference variation[J]. Chinese Journal of Electronics, 2024, 33(1): 293-302.
- [52] 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.
- [53] Narvekar N D, Karam L J. A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)[J]. IEEE Transactions on Image Processing, 2011, 20(9): 2678-2683.
- [54] Hassen R, Wang Z, Salama M A. Image sharpness assessment based on local phase coherence[J]. IEEE Transactions on Image Processing, 2013, 22(7): 2798-2810.
- [55] Li L, Lin W, Wang X , et al. No-Reference image blur assessment based on discrete orthogonal moments[J]. IEEE Transactions on Cybernetics, 2016, 46(1): 39-50.
- [56] Zhan Y, Zhang R. No-reference image sharpness assessment based on maximum gradient and variability of gradients[J]. IEEE Transactions on Multimedia, 2017, 20(7): 1796-1808.
- [57] Hosseini M S, Zhang Y, Plataniotis K N. Encoding visual sensitivity by maxpol convolution filters for image sharpness assessment[J]. IEEE Transactions on Image Processing, 2019, 28(9): 4510-4525.
- [58] 史晨阳, 卞仕磊, 江本赤, 等. 基于局部色貌的无参考彩色模糊图像质量评价方法[J]. 光学技术, 2023, 49 (05): 615-622.
- [59] Baig M A, Moinuddin A A, Khan E. A simple spatial domain method for quality evaluation of blurred images[J]. Multimedia Systems, 2024, 30(1): 28.
- [60] Li C, Bovik A C., Wu X. Blind image quality assessment using a general regression neural network[J]. IEEE Transactions on Neural Networks, 2011, 22(5): 793-799.
- [61] Zhao M, Ji S, Wei Z. Risk prediction and risk factor analysis of urban logistics to public security based on PSO-GRNN algorithm[J]. Plos one, 2020, 15(10): e0238443.
- [62] Li L, Xia W, Lin W, et al. No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features[J]. IEEE Transactions on Multimedia, 2017, 19(5):

1030-1040.

- [63] Lu Q, Zhou W, Li H. A no-reference Image sharpness metric based on structural information using sparse representation[J]. Information Sciences, 2016, 369: 334-346.
- [64] Yu S, Wu S, Wang L, et al. A shallow convolutional neural network for blind image sharpness assessment[J]. Plos One, 2017, 12(5): e0176632.
- [65] Kim J, Nguyen A D, Lee S. Deep CNN-based blind image quality predictor[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 30(1): 11-24.
- [66] Liu L, Liu B, Huang H, et al. No-reference image quality assessment based on spatial and spectral entropies[J]. Signal Processing: Image Communication, 2014, 29(8): 856-863.
- [67] Li D, Jiang T, Lin W, et al. Which has better visual quality: the clear blue sky or a blurry animal?[J]. IEEE Transactions on Multimedia, 2019, 21(5): 1221-1234.
- [68] Zhang W, Ma K., Yan J, et al. Blind image quality assessment using a deep bilinear convolutional neural network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(1): 36-47.
- [69] Moorthy A K, Bovik A C. Blind image quality assessment: From natural scene statistics to perceptual quality[J]. IEEE Transactions on Image Processing, 2011, 20(12): 3350-3364.
- [70] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3), 209–212.

攻读学位期间参与的科研项目与学术成果

一、论文发表情况

- [1] He H, Jiang, B; Shi, C, et al. An Objective Evaluation Method for Image Sharpness Under Different Illumination Imaging Conditions. PHOTONICS. 2024. 11. (11) :1032. (SCI 收录)
- [2] 刘懿平,何欢,彭丰,等. 基于改进 YOLOv5 算法的密集动态目标检测方法[J]. 安徽科技学院学报,2024,38(02):79-86.

二、专利申请情况

- [1] 江本赤, 何欢, 刘懿平等. 一种自动化激光切割件取出装置: CN217965399U[P]. 2022-12-06.

三、参与项目情况

- [1] 安徽省教育厅重点项目, 2022AH050983, 机器视觉辅助驾驶场景下的成像效果评价与提升研究, 2022.09-2024.08.
- [2] 安徽省车载显示集成系统工程研究中心开放基金项目, VDIS2023C01, 车载机器视觉成像退化机制建模及其在图像质量评价中的应用研究, 2023.12-2025.11.

致谢

时间如白驹过隙稍纵即逝，转眼间三年研究生学习生涯接近尾声。这三年，虽然只是整个人生中短暂的一段时间，但这段时间内经历的种种都是宝贵的，不久我即将毕业，临近大论文即将完成之际，在这里我由衷地感谢我的导师江本赤教授和史晨阳老师。

江老师在学术上学识渊博，对待学生认真负责，科研严谨且经验丰富，有着极强的论文写作水平，给了我很多指导。尤其是在小论文的发表中，开始进展一直不顺利，我比较着急也很烦躁，但是江老师和史老师及时给予指导，并细致地告诉我论文写作过程中的注意事项；不仅如此，在老师们的身上我也学习到了有些书本上没有的道理。

也感谢我的同学和朋友，研究生期间，在学习道路上，大家共同进步、相互鼓励。最后，我要感恩我的家人，尤其是我的爸爸妈妈，他们年纪已大，但对我的决定都是大力支持，物质和精神上他们都是我坚强的后盾。

我相信人生没有白走的路，我走的每一步都算数。努力是会得到回报的！

尚德敏学 唯实惟新

