

分类号：TP317.4

密 级：公开

学校代码：10363

学 号：2210120110



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于人眼视觉特性的图像
 质量评价算法研究

论 文 作 者 卞仕磊

校 内 导 师 江本赤 副教授

校 外 导 师 崔 岩 高级工程师

专业学位类别 机械

研究方向（领域） 机器视觉

论文提交日期： 2024 年 4 月 16 日

分类号：TP317.4

单位代码：10363

密 级：公开

学 号：2210120110

基于人眼视觉特性的图像质量评价算法研究

Research on Image Quality Assessment Based on
Human Visual Characteristics

学 生 姓 名： 卞仕磊

校 内 导 师： 江本赤（副教授）

校 外 导 师： 崔岩（工程师）

专 业 学 位 类 别： 机械

研究方向（领域）： 机器视觉

论 文 答 辩 日 期：2024 年 5 月 17 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：

日期：2024 年 5 月 31 日

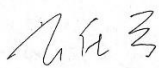
安徽工程大学学位论文版权使用授权书

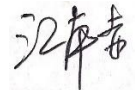
学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

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： 

指导教师签名： 

日期：2024年5月31日

日期：2024年5月31日

基于人眼视觉特性的图像质量评价算法研究

摘 要

随着信息技术的发展，图像作为一种常用的视觉信息载体，在人类生产生活中受到广泛应用。然而在采集、传输和存储等图像处理环节中，由于各种因素影响，会造成图像信号的损坏甚至丢失，导致图像视觉体验下降，并对后续的图像应用产生不良影响。因此，量化图像质量即研究图像质量评价算法，对于图像处理技术及应用具有着重要意义。

图像质量评价算法的主要作用是代替人眼完成图像质量的量化任务并做出评分。基于此，论文拟在现有图像质量评价算法基础上，探索符合人眼视觉感知结果的改进方法，并引入机器学习方法进一步提升主观和客观评分的一致性。论文的主要内容如下：

(1) 在分析现实场景中颜色明度和色度共变性的基础上，研究了一种基于色貌尺度相位一致性的全参考图像质量评价算法。首先采用 CIELAB 色域中的 Vividness 图来同步描述颜色的亮度和色度变化，为更精确地检测彩色图像的相位信息变化量，引入了色彩尺度相位一致性指标。其次，使用均方根法计算图像对比度差异值，以弥补相位对比度变化。然后，在 CIELAB 色空间中的色度通道上计算图像色度相似性。最后，采用标准差法将相位一致性、对比度和色度三种特征值进行求和池化，得到全参考图像质量评价计算模型。实验结果表明，与常用算法相比，本方法具有更高的预测精度，并与人眼视觉感知结果保持较高的一致性。

(2) 为充分利用颜色空间中颜色分量上的结构信息, 研究了一种基于多通道视觉信息融合的全参考图像质量评价算法。首先构造了高斯拉普拉斯(Laplacian of Gaussian, LoG)滤波器, 使用 LoG 滤波器分解图像对应的三个颜色通道上的频率信息, 融合滤波后的三个分量, 得到多通道 LoG 滤波算子, 通过该算子提取的图像频段相较于传统方法更加符合人眼感知结果。然后引入图像对比度相似度来弥补滤波后图像对比度的变化。最后, 计算图像色度相似度, 并将多通道 LoG 滤波、对比度以及色度三种图像特征值融合, 得到全参考图像质量评价计算模型。实验结果表明, 该算法无论在总体性能比较还是单个失真性能比较中都有较好的表现。

(3) 针对图像特征在池化过程中造成的差异性弱化问题, 研究了一种基于机器学习的改进池化策略。该方法通过粒子群优化的广义神经网络 (Particle Swarm Optimization General Regression Neural Network, PSO-GRNN) 学习不同图像特征的相似分数在空间中的最佳非线性组合, 实现了对传统池化方法性能提升。首先将传统图像特征相似分数作为图像标签来构建数据库, 弥补了标签样本不足的问题。然后将数据库中的所有标签按照一定比例随机分为训练集和测试集。最后将特征相似分数输入到已经训练完成的 PSO-GRNN 网络中, 获得最终的图像质量评分。实验将 PSO-GRNN 池应用于五个经典图像质量评价算法, 与原算法的结果相比, 本方法在预测精度和泛化性能两个方面都表现出较高的可靠性。

关键词: 图像质量评价, 人眼视觉系统, 多通道 LoG 滤波, 相位一致

性，PSO-GRNN。

RESEARCH ON IMAGE QUALITY ASSESSMENT BASED ON HUMAN VISUAL CHARACTERISTICS

ABSTRACT

With the development of information technology, image, as one of the most commonly used visual information carriers, has been widely used in human life. However, in various image processing links such as acquisition, transmission and storage, the image signal will be damaged or even lost due to various factors, resulting in the decline of image visual experience and adverse effects on subsequent image applications. Therefore, quantifying image quality, that is, studying image quality assessment algorithm, is of great significance to image processing technology and application.

The main function of image quality assessment algorithm is to replace human eyes to complete the quantification task of image quality and make a score. Based on this, this paper intends to explore an improvement method that conforms to human visual perception results on the basis of existing image quality assessment algorithms, and introduce machine learning methods to further improve the consistency of subjective and objective scores. The main contents of the paper are as follows:

(1) On the basis of analyzing the co variability of color brightness and chromaticity in real scenes, a full-reference image quality assessment

algorithm based on color appearance scale phase consistency is studied. Firstly, the CIELAB Vividness map in the color domain is used to describe color brightness and chroma changes synchronously, and a color scale phase consistency index is introduced to detect phase information changes of color images more accurately. Secondly, the root-mean-square method is used to calculate the image contrast difference to compensate for the phase contrast change. Then, the image chromaticity similarity is calculated on the chromaticity channel in the CIELAB color space. Finally, the standard deviation method is used to sum and pool the three characteristic values of phase consistency, contrast and chrominance to get the full reference image quality assessment model. The experimental results show that the proposed method has higher prediction accuracy than the common algorithms, and maintains a high consistency with human visual perception results.

(2) To fully utilize the structural information of color components in color space, a comprehensive reference image quality assessment algorithm based on multi-channel visual information fusion was studied. Firstly, a Laplacian of Gaussian (LoG) filter is constructed. The frequency information on the three color channels corresponding to the image is decomposed by the LoG filter, and the three filtered components are fused to obtain the multi-channel LoG filter operator. Compared with the traditional method, the image frequency band extracted by this operator is

more consistent with the human eye perception results. Then the image contrast similarity is introduced to compensate for the change of image contrast after filtering. Finally, the image chromaticity similarity is calculated, and the three image feature values of multi-channel LoG filtering, contrast and chromaticity are fused to obtain a full-reference image quality assessment model. The experimental results show that the proposed algorithm has good performance both in the overall performance comparison and in the single distortion performance comparison.

(3) A machine learning based improved pooling strategy was studied to address the issue of weakening differences caused by image features during the pooling process. This method learns the optimal nonlinear combination of similarity scores of different image features in space through Particle Swarm Optimization General Regression Neural Network (PSO-GRNN), achieving performance improvement over traditional pooling methods. Firstly, traditional image feature similarity scores are used as image labels to construct the database, which makes up for the problem of insufficient label samples. Then randomly divide all labels in the database into training and testing sets according to a certain proportion. Finally, the feature similarity score is input into the trained PSO-GRNN grid to obtain the final image quality score. The experiment applied the PSO-GRNN pool to five classic image quality evaluation algorithms. Compared with the results of the original algorithm, this method showed

high reliability in both prediction accuracy and generalization performance.

KEY WORDS: image quality assessment, human vision system, multi-channel LoG filtering, phase consistency, PSO-GRNN

目 录

第 1 章 绪 论.....	1
1.1 课题研究背景及意义.....	1
1.1.1 失真的产生.....	1
1.1.2 图像质量评价的应用.....	2
1.2 图像质量评价的研究现状.....	3
1.2.1 研究对象.....	4
1.2.2 方法分类.....	4
1.2.3 研究方向.....	5
1.2.4 分析验证.....	11
1.3 研究内容与结构安排.....	14
1.3.1 研究内容.....	14
1.3.2 结构安排.....	15
第 2 章 人类视觉系统及其特性.....	17
2.1 人类视觉系统的生理结构.....	17
2.1.1 人眼与视觉中枢神经系统.....	17
2.2 人类视觉系统的感知特性.....	19
2.2.1 对比敏感度.....	19
2.2.2 视觉掩盖.....	20
2.2.4 最小可觉差和视觉多通道分解.....	21
2.3 本章小结.....	22
第 3 章 基于色貌尺度相位一致性的全参考图像质量评价.....	23
3.1 色貌尺度相位一致性相似图.....	23
3.1.1 色貌算子.....	23
3.1.2 获取色貌尺度相位一致性相似图.....	24
3.2 对比度和色度相似图.....	26
3.4 基于色貌尺度相位一致性的全参考图像质量评价算法.....	28
3.5 实验结果与分析.....	30
3.5.1 数据库和模型评价标准.....	30
3.5.2 算法参数优化.....	31
3.5.3 算法性能比较.....	32
3.6 本章小结.....	38
第 4 章 基于多通道融合的全参考图像质量评价.....	39
4.1 多通道信息融合.....	39
4.1.1 LoG 的定义.....	39
4.1.2 多通道 LoG 滤波相似性.....	40
4.2 对比度和色度相似图.....	42
4.4 基于多通道融合的全参考图像质量评价算法.....	43
4.5 实验结果与分析.....	44
4.5.1 数据库和模型评价标准.....	44
4.5.2 算法参数优化.....	45
4.5.3 算法性能比较.....	46

4.6 本章小结.....	52
第 5 章 基于 PSO-GRNN 池化策略的全参考图像质量评价.....	53
5.1 PSO-GRNN 池化方法.....	53
5.1.1 GRNN 网络结构	53
5.1.2 PSO 方法原理	54
5.1.2 基于 PSO 优化的 GRNN 网络.....	55
5.1.3 图像预处理和 PSO-GRNN 池化策略.....	56
5.2 实验结果与分析.....	59
5.2.1 数据库和模型评价标准.....	59
5.2.2 算法参数设置.....	59
5.2.3 算法性能比较.....	60
5.3 本章小结.....	64
第 6 章 总结与展望.....	65
6.1 工作总结.....	65
6.2 未来展望.....	66
参考文献.....	67
攻读学位期间发表的学术论文目录.....	73
致谢.....	74

第1章 绪论

本章首先从图像质量评价的研究背景和意义出发,阐述了图像失真的来源以及图像质量评价技术在数字通信中的应用。接着分别从研究对象、研究方向、研究进展和分析验证等四个部分,对图像质量评价领域的研究现状进行详细介绍。最后给出了文章的研究内容和组织结构。

1.1 课题研究背景及意义

视觉是人类获取外界信息的主要感官之一,而图像作为视觉最常见的受体,在人类生活中扮演着重要角色。随着手机和摄像机等拍摄工具的发展,人类获取图像的方式变得更加便捷。面对质量不一的海量图像,人类对图像质量的追求愈加强烈。图像质量评价(Image Quality Assessment, IQA)算法是图像处理环节中最常用的参数调整工具,其为图像质量的稳定传输提供了可靠的客观依据^[1]。因此,研究出高精度和高效率的图像质量评价算法是满足人类对图像质量需求的重要途径。

1.1.1 失真的产生

图像失真又称为畸变,指信号在传输过程中与原有信号相比有产生了偏差即信号损失。图像失真一般发生在图像处理环节中,包括图像的采集、水印、压缩、传输、存储、认证和重建等^[2]。在上述环节中,图像的信息无法保证无损传播。例如图像在采集阶段会受到人为(摄者技术)和环境因素(深海水压)的影响,导致收集到的图像模糊不清;图像在传输时,由于受到网络带宽和设备性能的限制,产生的图像信号带有噪声。图像在压缩时,为了释放存储空间,一般采用有损的压缩方式,从而造成了图像信息的丢失。上述图像信息的损失不仅造成了较差的观感还会使得图像失去可用性。如在艺术领域,丢失信息的艺术照片会大大降低鉴赏价值。在科技领域,失真图像会直接影响到实验结果的准确性,从而减缓实验进度。图 1-1 中给出了图像在不同压缩水平下的失真实例。图 1-1 中,图

1-1 (a) 是原始未失真图像，图 1-1 (b)、图 1-1 (c) 和图 1-1 (d) 是失真等级依次为 2、3 和 4 的 JPEG 压缩失真图像（失真等级越大，失真程度越高）。如图 1.1 所示，图像的质量随着压缩程度的增加而逐渐降低，当压缩达到一定时，图像不仅失去了部分信息甚至没有了参考价值。因此，研究出高精度的图像质量评价算法，量化图像质量，对图像处理和应用具有重要意义。



(a) 原始图像



(b) 1 级 JPEG 压缩



(c) 2 级 JPEG 压缩



(d) 3 级 JPEG 压缩

图 1-1 不同失真等级的 JPEG 压缩失真图像

1.1.2 图像质量评价的应用

图像质量评价在多个领域中都有着广泛的应用价值，如动态监控、智能筛选、参数调整和性能评判等。通过对不同领域中的图像质量进行评估和优化，可以提高图像的应用效果、满足用户需求和提高产品竞争力并保障安全与隐私。下面详细介绍图像质量评价在不同领域中的应用。

图像质量评价主要有以下四点应用：第一，图像质量的动态监控和反馈。例如，在日常的监控摄像中，摄像头随着画面的变动，监控系统会自适应的改变焦

距和曝光等参数,依据实时图像质量的变化为摄像系统调整参数提供依据。第二,海量图像的智能筛选。例如,通过图像质量评价算法实时量化图像质量,将海量图像中不符合质量要求的图像剔除。第三,优化图像处理系统的参数。例如,对一幅图像采用不同参数设置的同一种图像处理方法,通过比较处理后的图像质量来分析不同参数值的优劣,选出最优结果。第四,图像处理方法的性能指标。例如,对相同的一组图像进行不同类型的压缩处理,通过比较压缩后的图像质量来筛选更加合适的压缩方法。

1.2 图像质量评价的研究现状

本小节将分别从研究对象、方法分类、研究方向和分析验证等多方面介绍 IQA 的研究现状。图 1-2 给出了 IQA 研究现状框架图。其中, IQA 分类为主观评估和客观评估。客观评估分为客观算法和评价标准。

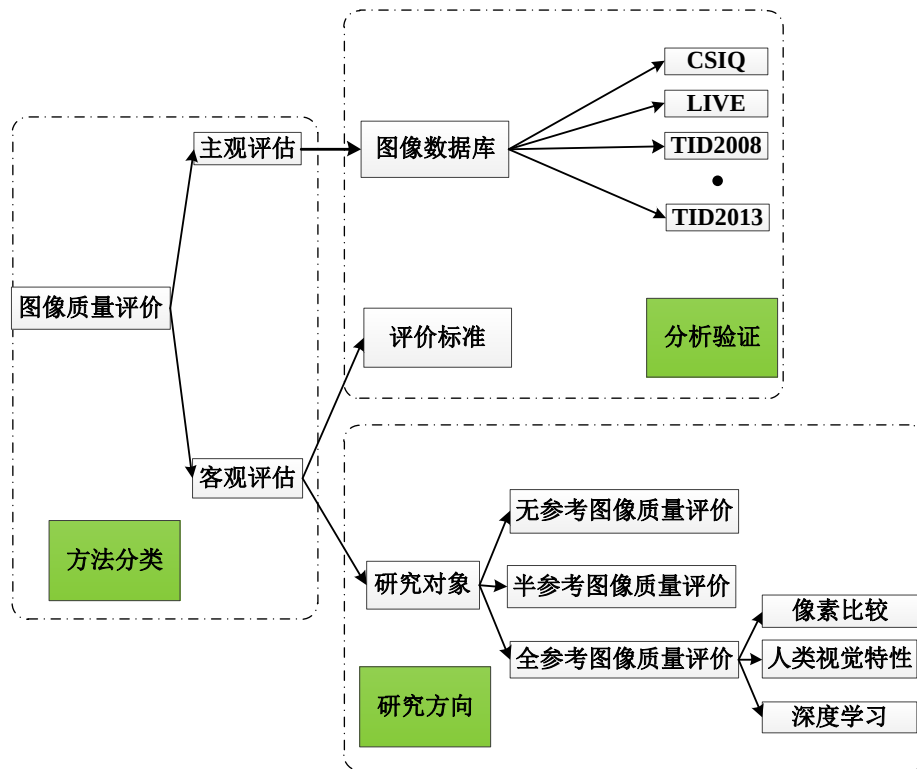


图 1-2 IQA 研究现状框架图

1.2.1 研究对象

图像质量评价的研究对象大体上分为两类^[3]，一类是自然图像(Natural Image, NI)，另一类是屏幕内容图像(Screen Content Image, SCI)。这两类图像的特点是都具有明显的视觉差异性。

自然图像是指现实世界中的场景、物体和人类活动等视觉表示。其一般是通过专业摄像设备获得的，反映了被拍摄物体的外观和特征。自然图像可以包括风景、人物、动物、植物等各种元素，并反映了真实世界中的色彩、形状、纹理、光照等方面。与合成图像或计算机生成图像不同，自然图像是通过直接观察真实场景获得的，并具有真实性和感知上的可信度。

屏幕内容图像一般分为文字和图像两部分。其中图像又分为计算机生成的卡通图形图像和自然图像，例如常见的网页和幻灯片等。文字部分是指含有文字内容的图形图像，如网络小说和网络学报等。

1.2.2 方法分类

目前，根据图像质量的获取方式，图像质量评价可分为主观图像质量评价(Subjective Image Quality Assessment)和客观图像质量评价(Objective Image Quality Assessment)两大类^[4]。主观图像质量评价是指人类在特定条件下对图像质量给出打分，所得结果为平均主观评分值(Mean Opinion Score, MOS)或差分平均主观评分值(DifferenceMean Opinion Score, DMOS)^[5]。图像服务于人类，由人类观测给出的图像质量评分无疑是最准确和符合实际的。主观图像质量评价适用于处理图像较少的场景，当面临大量图像数据时，人为的观测费时费力。因此，主观评价方法在日常生活中很少应用。客观图像质量评价是用客观算法对图像进行打分。其不受地域和时间的限制，能够高效便捷的处理图像。例如，在评价过程中，不需要人的参与，节省了大量的人力物力；在面对海量图像时，客观算法相对于人为观测能够做出更快速的打分，有效的节约了时间。客观图像质量评价的最终目的是为了与主观图像质量评价保持高度一致。

1.2.3 研究方向

根据研究对象的可用性,一般将客观 IQA 分为三个研究类型:完全参考(Full reference, FR)^[6]、部分参考(Reduced reference, RR)^[7]和无参考(No Reference, NR)^[8]IQA。全参考方法是指利用原始未失真图像的全部信息对失真图像进行质量评价;部分参考方法是指有目标的选取部分原始图像信息来评估失真图像质量,例如选取图像对比度、色度和梯度等特征信息;无参考方法是指不借助任何的参考信息,对图像进行质量评价。

(1) 全参考方法

全参考评价方法是目前最热门的研究方向。其研究框架大体分为三步,如图 1-3 所示。首先比较失真图与参考图之间的差异,得到局部差异图,计算差异图的差异分数,然后在空间中汇集局部差异分数,最后给出图像的质量分数。常见的全参考图像评价池化方法有均值、标准差和最大值池化等,通过池化多个图像特征可以有效提高评价算法对不同失真变化的敏感度。全参考评价方法按照不同的计算思路,可大致分为三大类^[9]:基于数学的方法、基于人类视觉系统(Human vision system, HVS)特性的方法和基于机器学习^[10]的方法。

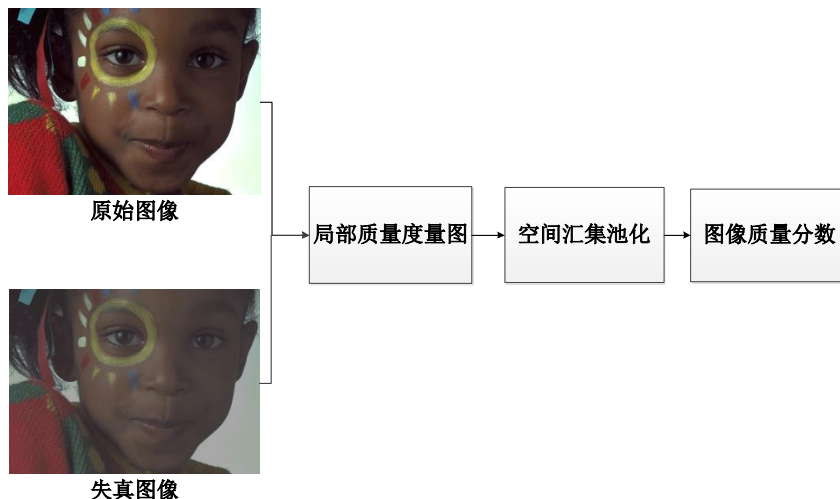


图 1-3 全参考图像质量评价方法基本框架

① 基于数学的方法

基于数学的方法是指直接比较失真图像和参考图像间,所有像素点灰度值的差异。如均方误差(Mean-Square Error, MSE)^[11]和峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)^[12]。MSE 和 PSNR 的计算方法为:

$$MSE = \frac{1}{N} \sum_{n=1}^N (x_n - y_n)^2 \quad (1-1)$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (1-2)$$

其中， n 表示像素位置索引， N 表示像素总数， x 和 y 分别代表原始图像和失真图像。这两种方法在评价噪声失真图像时表现出可靠的预测精度，由于其计算复杂度低，被广泛应用于图像处理环节。但上述两种方法，认为不同位置像素点对人眼的感知贡献是相同的，忽略了人眼的视觉特性，因此未能和人眼感知质量有较好的一致性。

② 基于 HVS 特性的方法

近年来，研究者们开始利用 HVS 特性^[8]（亮度自适应、对比度敏感度、掩盖效应和注意机制等）评价图像质量^[13]，代表性的模型有 Wang^[14]提出的结构相似性指数（Structural Similarity, SSIM）。SSIM 通过计算图像亮度、对比度和结构三个特征相似度来量化图像质量，将这三个分量在空间中汇集池化得到最终质量分数，如图 1-4 所示。图中信号 x 和信号 y 分别代表原始图像信息和失真图像信息， s 为待测图像间的相似质量分数， s 值介于 0-1 之间，其值越接近 1 图像的质量越高，当 $s=1$ 时，失真图像等同于原始图像。SSIM 中的亮度相似性定义为：

$$l(x, y) = \frac{2\mu_x \cdot \mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (1-3)$$

其中， μ_x 和 μ_y 表示 x 和 y 的平均强度， $\mu_x = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$ 。 c_1 是增加稳定性的常数。

对比度相似性定义为：

$$c(x, y) = \frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (1-4)$$

其中， σ_x 和 σ_y 表示 x 和 y 的方差， c_2 是增加稳定性的常数。结构相似度定义为：

$$s(x, y) = \frac{2\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (1-5)$$

其中 σ_{xy} 表示 x 和 y 的协方差, c_3 是正常数, 增加计算稳定性。SSIM 指数可由上述三个相似度值汇集得到, SSIM 指数定义为:

$$SSIM(x, y) = l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma \quad (1-6)$$

其中 α 、 β 和 γ 的值均大于 0, 且表示不同分量的权值。SSIM 具有计算复杂度低和预测精度高等特点, 因此, 在评价图像质量中受到广泛运用。在 SSIM

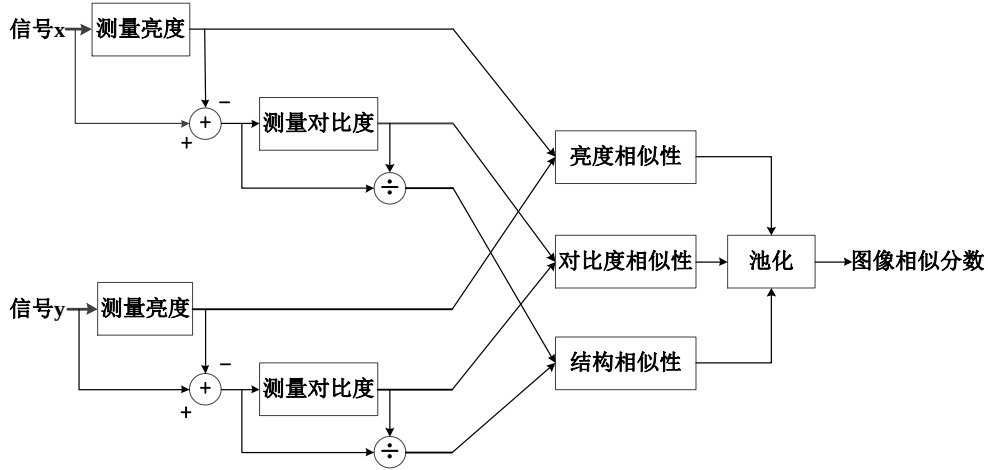


图 1-4 SSIM 流程图

研究基础上, 大量研究者提出了改进方法。例如: Wang 等^[15]提出了多尺度的结构相似性方法 (multi-scale SSIM, MS-SSIM), 实现从单尺度到多尺度的扩展。Li 等^[16]通过对整幅图像的边缘、纹理和平滑区域进行计算, 且赋予其相应权值, 提出了一种名为三元素加权 SSIM (Three-component weighted SSIM, 3-SSIM) 的算法。Zhang 等^[17]则结合了图像视觉显著性相似性和图像对比度相似性^[18]两个特征分量, 提出了 CVSS (Contrast and Visual Saliency Similarity) 算法。上述两种算法预测精度较好, 但泛化性能仍不足。Liu 等^[19]则通过遮掩图像的结构信息来设计 GSM (Gradient Similarity Metric) 算法模型, 该方法的计算精度得到了提升, 但泛化能力较差。Sheikh 等^[20]引入了信息保真度准则 (Information Fidelity Criteria, IFC), 提高了算法的预测精度, 然而这也导致了计算效率上的损失。

相位一致性 (Phase Congruency, PC) 可以在亮度不断变换的情况下精准检测图像中的边缘纹理等特征^[21]。Zhang 等^[22]提出了特征结构相似度 (Features Similarity Index, FSIM) 算法, 该算法通过高斯滤波器获取图像亮度通道上的相位信息, 然后使用 Scharr 算子^[23]计算图像梯度相似度, 并在色度通道上计算图像

色度相似度，融合上述三个特征相似度，最终得到质量分数。输入一对规格相同的 x （原始图像）和 y （失真图像），FSIM 指数则由公式 1-7 给出。

$$FSIM = \frac{\sum_{x \in \Omega} S_p(x, y)^\alpha \cdot S_G(x, y)^\beta \cdot S_q(x, y)^\gamma \cdot PC_{\max}(x, y)}{\sum_{x \in \Omega} PC_{\max}(x, y)} \quad (1-7)$$

其中， $S_p(x, y)$ 表示 x 和 y 间的相位相似性。 $S_G(x, y)$ 表示 x 和 y 间梯度相似性。 $S_q(x, y)$ 表示 x 和 y 间色度相似性。 $PC_{\max}(x, y) = \max(PC(x), PC(y))$ ，用于表征图像 x 和 y 的局部重要性权重。这些图像特征的计算方式和 SSIM 相似，其中相位一致性和梯度相似性的计算过程如下所示。

$$S_p(x, y) = \frac{2PC(x) \cdot PC(y) + c_4}{PC(x)^2 + PC(y)^2 + c_4} \quad (1-8)$$

$$S_G(x, y) = \frac{2G(x) \cdot G(y) + c_5}{G(x)^2 + G(y)^2 + c_5} \quad (1-9)$$

色度相似性计算过程如下所示。

$$S_I(x, y) = \frac{2I(x) \cdot I(y) + T_1}{I(x)^2 + I(y)^2 + T_1} \quad (1-10)$$

$$S_Q(x, y) = \frac{2Q(x) \cdot Q(y) + T_2}{Q(x)^2 + Q(y)^2 + T_2} \quad (1-11)$$

$$S_q(x, y) = S_I(x, y) \cdot S_Q(x, y) \quad (1-12)$$

其中， I 和 Q 表示图像的色度通道，具体计算方法参考 FSIM，不再作过多赘述。相较于 SSIM，FSIM 指数引入了 PC 来描述图像的纹理结构，其结果更加符合人眼的感知特性。

③ 基于机器学习的方法

目前，机器学习技术在图像质量评价领域取得了显著进展。机器学习最大优点是在训练海量数据集时能够表现出可观的泛化性能^[24]。但由于缺乏可训练的数据集，且模型在训练过程中耗时严重，因而结合机器学习的图像质量评价方法依然稀少。Kang^[25]等首次在图像评价领域使用卷积神经网络，该方法通过分割图像来增加训练数据，并为每个图像块分配相同的主观质量分数。图像评价的目的是服务于人眼，据此 Li^[26]等通过计算多个 DoG 频带特征的非线性组合来模拟

HVS, 并采用随机森林回归方法训练特征。Wang^[27]等使用深度神经网络(the Deep Neural Network,DNN)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来获取全局图像质量分数。Wu^[28]等人提取了图像的相位一致性、梯度和熵三种特征分量,接着通过机器学习方法训练特征,最终给出质量分数。与传统的全参考图像质量评价相比,该方法实现了较好的预测性能。

文献[29]设计了一种用于局部图像质量评估的卷积神经网络,提出了基于权重池的多尺度图像质量评估方法。该方法具有可靠的泛化性能。

综上所述,全参考评价方法相较于部分参考和无参考方法拥有更多的可用信息。其预测精度也是三者中最高的。但目前对全参考评价方法的研究仍有不足。例如无法同时满足高预测精度、高泛化性能和高计算效率这三个性能要求。针对该问题,需要在传统研究基础上做出更多的探索与改进。

(2) 部分参考方法

虽然全参考方法预测性能好,但在实际生活中常见还是信息残缺的图像,因而部分参考方法更加适用于实际场景^[30]。部分参考方法一般是指在发送端提取参考图像的某几种特征信息,同时在接收端提取失真图像对应的特征信息,比较相同特征下参考和失真图像间的差异,最终获得质量分数。图 1-5 给出了部分参考方法的基本流程框架。

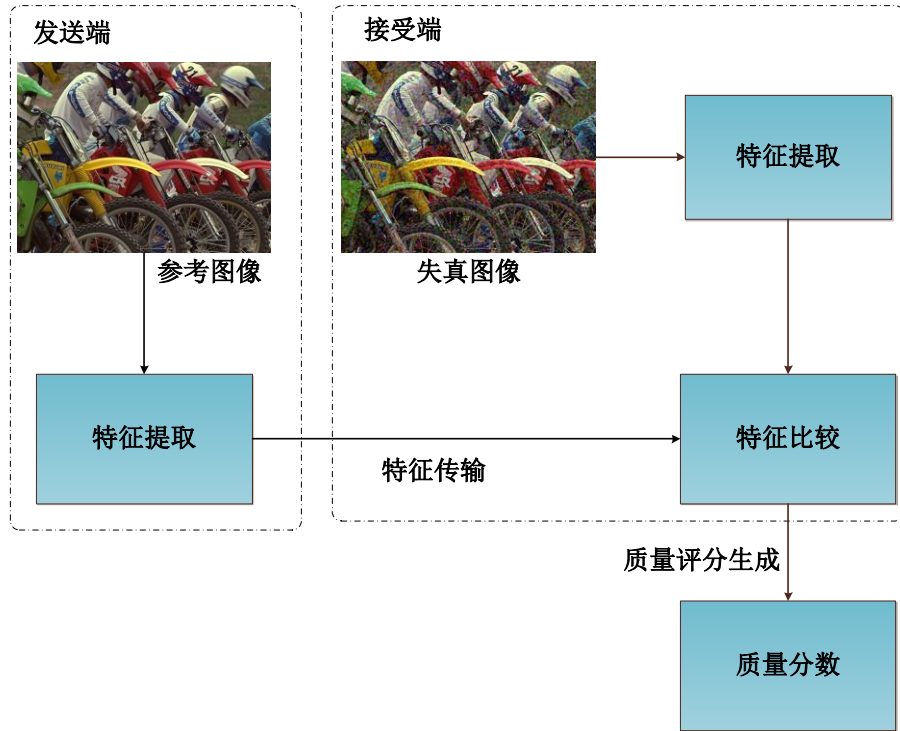


图 1-5 部分参考图像质量评价方法通用框架

理想的部分参考方法具有一个特点，即利用越少的图像信息得到越高的预测精度。视觉显著性失真（Attended Visual Content Degradation, AVCD）^[31]是一个典型的半参考图像质量评价方法。该方法首先提取图像视觉显著性和梯度上的结构直方图，然后直接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，得到最终的质量评价。Wang^[32]等人提出的基于小波变换域自然图像统计模型的半参考图像质量评价算法是一种在图像处理领域颇受关注的方法。这种算法的核心思想是利用小波变换的特性来分析和提取图像的特征参数，进而评估图像的质量。上述两种方法都有较高的计算效率和泛化能力。

总而言之，半参考评价方法相对于全参考评价方法更少的依赖参考信息，且在保证图像质量精度的同时提高了图像的传输效率。

（3）无参考方法

无参考图像质量评价是指在没有原始图像的情况下，对单个图像进行分析和处理。目前，无参考图像质量评价的研究主要集中在以下几个方面。首先是手工设计的特征，如锐度、对比度、颜色饱和度等^[33]。这些方法通过分析图像的低级视觉特征来评估图像的质量，但是面临场景比较复杂的图像其评估效果却并不理

想：基于统计的方法，该方法通过建立大规模图像数据库来提取图像统计特征，并使用机器学习算法进行训练^[34]。这种方法可以在一定程度上提高图像质量评价的准确性，但是应对于不同场景的质量预测仍有一定的挑战。Sheikh 等^[35]人利用小波系数的统计模型来分析图像质量，提出了一种基于自然场景统计模型的 JPEG2000 压缩失真图像质量评价方法；基于深度学习的方法主要是使用卷积神经网络^[36]（CNN）等深度学习模型，提取图像的高级特征，完成对图像质量评价模型的设计。这种方法具有较好的泛化性能，适用于各种复杂图像场景下的无参考图像质量评价。Zhou^[37]等采用小波域幅度谱的下降曲线和相邻尺度系数的同向投影距离作为特征，再结合广义回归神经网络^[38]（General Regression Neural Network, GRNN），对图像直接进行量化。目前出现了几种基于生成对抗网络（GAN）的评价方法，通过对生成器和判别器训练来实现无参考图像质量评价^[39]。上述两种方法在预测精度上都做到了可靠的提升。

总的来说，无参考图像质量评价研究主要集中在手工设计特征、数据统计和深度学习方法上，由于其几乎不依赖于参考信息，故其比其它两个图像质量评价方法更加实用。然而，目前图像质量评价方法的研究仍然存在一些问题。就全参考图像质量评价方法而言，一幅图像往往存在着多种失真，为了衡量图像的失真情况，需选取合适的特征。但特定的特征只会对个别的失真类型做出反馈，因此，为了获得与任意图像失真高度相关的一般特征，有必要进一步研究不同类型失真对图像质量影响的共性。目前基于人眼视觉特征的图像质量评价方法并不能给出与人眼视觉知觉完全一致的评价结果。需要深入探索人类视觉系统的生理特性，改进现有基于人类视觉特性的图像质量计算模型，进一步提高算法的预测精度。

1.2.4 分析验证

在分析不同图像评价模型预测性能的过程中，需要用到公共图像数据库和性能评价标准。公共图像数据库便于研究者对图像质量评估算法进行验证和比较。通过使用公共数据库，研究者可以评估自己的算法在不同失真类型和程度下的性能，并与其他算法进行比较。目前，常用的公共数据库有 LIVE^[40]、CSIQ^[41]、TID2008^[42]和 TID2013^[43]数据库。性能评价标准可以直观反映预测结果和实际结

果之间的相关性，相关性越高，算法的预测精度越好，与人眼感知的结果愈加一致^[47]。下面对公共数据库和性能评价标准作简单的介绍。

（1）公共图像数据库

LIVE (Laboratory for image and Video Engineering, LIVE) 图像数据库^[40]由德州大学奥斯汀分校的图像与视频工程实验室所构建。它包含大量的图像对，每一对图像都有人工标注 0-100 的差异平均主观分数 (Difference Mean Opinion Score, DMOS) 值。所有图像对都经过了失真处理，以模拟真实世界中的图像质量变化。LIVE 数据库包含了 29 幅高分辨率的 RGB 参考图像以及对应的 779 幅失真图像。该数据库共有 5 种不同的失真类型。

CSIQ 图像数据库^[41]由美国奥克拉荷马州立大学计算感知和图像质量实验室研究建立。CSIQ 图像数据库为评估图像质量评价算法提供了一个广泛的测试基准。该数据库包含 30 幅高分辨率的 png 参考图像，6 种不同的失真类型总计为 866 幅的失真图像，每一幅图像都含有 5 种失真等级以及人工标注为 0-1 的 DMOS 值。

TID2008 图像数据库^[42] (Tampere Image Database 2008) 是由乌克兰国家航空航天大学的 N504 信号接收、传输与处理系创建。该数据库包含 25 个参考图像和 1700 个失真图像，共分为 4 个不同的失真等级和 17 种失真类型。此外，数据库中的每个失真图像都经过多个被试者给出质量评分，提供了 0-9 的平均主观分数 (Mean Opinion Score, MOS) 值。该数据库的公开性和多样性使得它成为图像质量评价领域的重要参考资源，为算法研究和性能评估提供了一个标准平台。

TID2013 图像数据库^[43] (Tampere Image Database 2013) 是一个广泛应用于图像质量评估的数据库，在验证客观图像质量评价算法性能方面表现出色。它在 TID2008 基础上增加 7 种新的失真类型，形成了 24 种不同的失真类型。这些失真类型涵盖了实际应用中可能遇到的各种质量损坏情况。TID2013 数据库精心选择了 25 幅参考图像，以确保评估结果具有普适性和可靠性。该数据库包含 5 个失真等级和 24 种失真类型。因此，整个数据库包含了 3000 幅失真图像。TID2013 数据库还提供了每幅图像的 MOS (Mean Opinion Score) 值。MOS 值是通过主观

评价实验获得的，即请一组观察者对图像的质量进行打分，然后取平均分得到 MOS 值。

表 1-1 列举了上四个述数据库的基本信息，数据库的公开性和多样性使得它成为图像质量评价领域的重要参考资源，为评估和比较不同图像质量评价算法的准确性和可靠性提供了一个标准平台。

表 1-1 数据库的主要信息

数据库	年份	参考图像	失真图像	观测人数	主观评分
LIVE	2006	29	799	161	DMOS
CSIQ	2009	30	866	35	DMOS
TID2008	2008	25	1700	838	MOS
TID2013	2013	25	3000	971	MOS

(2) 性能评价标准

不同客观图像质量算法给出的评价结果并不在相同范围，如 CVSS 的质量分数在 0~1 之间，IFC 值在 0~100 之间，且不同数据库的 MOS/DMOS 取值范围也不同，因此为了保证算法评估的严谨性，需要对客观评分结果进行非线性拟合，以保持不同算法在量纲上的一致性和公平性。这里采用 Sheikh 给出的非线性拟合公式^[44]，如式 1-7

$$f(s) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(s - \beta_3))} \right] + \beta_4 s + \beta_5 \quad (1-13)$$

其中 $\beta_1, \dots, \beta_5$ 值都是拟合参数， s 代表原始客观评价分数， $f(s)$ 是映射到主观分数后的客观评价分数。对原始客观分数完成映射后，论文给出了四种常用的性能指标来衡量不同图像评价算法性能，分别为 Spearman 秩序相关系数 (Spearman rank-order correlation coefficient, SROCC)、Kendal 秩序相关系数 (Kendal rank-order correlation coefficient, KROCC)、Pearson 线性相关系数 (Pearson linear correlation coefficient, PLCC) 和均方根误差 (root mean squared error, RMSE)。前两个指标可以展现算法预测的趋势，SROCC 计算方法可由下式给出：

$$SROCC = 1 - \frac{6 \sum_{n=1}^N d_n^2}{N(N^2 - 1)} \quad (1-14)$$

其中, N 代表数据库中失真图像的数量。 d_n^2 表示每对主客观分数间的序列差异。KROCC 可定义为:

$$KROCC = \frac{N_c - N_d}{0.5 \times (N - 1) \times N} \quad (1-15)$$

其中 N_c 和 N_d 代表了图像数据库中相符和不相符的图像对。PLCC 反映了算法的预测精度, 可由公式 1-10 计算得到:

$$PLCC = \frac{\sum_i (q_i - \bar{q}) \cdot (o_i - \bar{o})}{\sqrt{\sum_i (q_i - \bar{q})^2 \cdot \sum_i (o_i - \bar{o})^2}} \quad (1-16)$$

其中, o_i 是第 i 幅图像的主观分数, $\bar{o}$ 是主观分数的均值。 q_i 为第 i 幅图像的拟合客观分数, $\bar{q}$ 是客观分数均值。上述三个标准值越接近 1, 算法的预测精度越高。

RMSE 值与之相反, 其值越接近 1, 算法的预测精度越低。RMSE 定义为:

$$RMSE = \sqrt{\frac{1}{N} \sum (q_i - o_i)^2} \quad (1-17)$$

1.3 研究内容与结构安排

人类视觉系统是一个高度复杂的非线性集合。如何将人类视觉系统特性结合到图像质量评价算法中, 并依据现代图像处理技术对算法的需求, 获得高精度和高效率的图像质量评价算法是论文研究的关键。

1.3.1 研究内容

依据 1.2.3 小节的内容, 全参考图像质量评价可分为基于数学、基于 HVS 特性和基于机器学习的方法。论文主要研究基于 HVS 特性的全参考图像评价方法, 具体内容如下:

(1) **基于色貌尺度相位一致性的全参考评价方法研究** 理想的 FR-IQA 算法可以很好的表征图像结构信息的变化。PC 值对比较图像结构信息差异有着显著的优势, 但所有传统 PC 无法全面的表征彩色图像结构信息变化。针对这一问题,

本文通过引入色貌算子计算彩色图像的相位值，再结合图像对比度和色度特征，提出了基于色貌尺度相位一致性的全参考评价的算法，该方法补偿了相位一致性对颜色信息的缺失，提高了算法的精度。

(2) **基于多通道视觉信息融合的全参考图像质量评价方法研究** 根据视觉感知特性，图像质量评价需要考虑颜色信息对人眼感知的影响，高斯拉普拉斯滤波器可以模拟人眼的神经机制，但传统的方法一般将高斯拉普拉斯滤波器用到图像亮度通道上，本课题对图像的三个通道分别使用高斯拉普拉斯滤波，再综合滤波后的三个分量得到了新的图像特征，该方法弥补了高斯拉普拉斯滤波器的单一性，提高了图像评价算法对图像质量评估的准确性。

(3) **基于 PSO-GRNN 池化策略的全参考图像质量评价方法研究** 图像特征池化是全参考评价框架中重要的一个步骤。理想的池化方法，可以最大程度上优化客观算法，计算出符合人眼视觉感知的评价结果。传统的池化方法都不能同时满足较高泛化性和准确性。本课题提出了一种基于 PSO-GRNN 池化策略的全参考图像质量评价方法，该方法是将与 HVS 相关的多个图像特征输入到已经训练完成的 PSO-GRNN 网络中，输出的值作为最终的客观分数。实验结果表明，所提方法在无论是在泛化性还是预测准确度都有不错的表现。

1.3.2 结构安排

论文主要以人眼视觉机制为基础，对与人眼特性相关的图像特征进行改进和提升。论文的整体结构安排如下。

第一章是绪论。本章首先阐述了图像质量评价研究的背景和意义。接着描述了图像质量评价的研究过程，最后介绍了本文的研究内容和成果。

第二章是人类视觉系统及其特性。本章从人眼和视觉中枢神经系统的生理结构入手，描述了从信号采集到视觉成像的过程。然后重点阐述了人类视觉系统的视觉特性及其应用，为后面的工作留下了理论基础

第三章是基于色貌尺度相位一致性的全参考评价方法。本章首先介绍了色貌指标的来源，接着提出了色貌尺度相位一致性方法，并将其与传统的相位一致性作比较。然后依据视觉多通道特性将色貌尺度相位一致性、对比度和色度三者融

合，提出了基于色貌尺度相位一致性的全参考评价方法。最后对本章的内容做出总结。

第四章是基于多通道视觉信息融合的全参考图像质量评价方法。本章从高斯拉普拉斯特性出发，提出了基于高斯拉普拉斯算子的图像多通道信息融合方法。接着，将上述方法与图像对比度和色度结合，提出了基于多通道视觉信息融合的全参考图像质量评价方法。最后分析了所提方法在公共图像数据库中的预测性能。

第五章是基于 PSO-GRNN 池化策略的全参考图像质量评价。本章将传统的图像评价方法与机器学习方法结合在一起，提取出了基于 PSO-GRNN 池化策略的全参考图像质量评价方法。文章首先介绍了 PSO-GRNN 的网络架构和训练流程。接着，提出了融合多个图像特征相似分数的 PSO-GRNN 池化方法。最后对所提方法进行实验验证。

第六章是总结与展望。本章阐述了论文的研究成果，并对图像质量评价领域的研究前景做出展望。

第 2 章 人类视觉系统及其特性

视觉是人类最重要的感觉之一。统计学研究表明，人类从外部环境获取信息的途径 80% 来源于视觉系统。图像是服务于人眼的，而图像质量评价的最终目的是替代 HVS 为图像质量做出准备判断。因此，研究人类视觉系统的组成与功能，探索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。本章将分别从人类视觉系统的生理构造和视觉特性两方面介绍人类视觉系统的基础知识，并为后续工作提供理论来源。

2.1 人类视觉系统的生理结构

人类视觉系统是一种复杂的感知系统，包括人眼、视神经和大脑中的视觉皮层等多个部分。如图 2-1 所示，人眼成像是一个及其复杂的过程。首先视觉刺激通过眼球系统收集，再由视觉通路传输到视觉神经，接着被大脑视觉皮层分析解码，最终生成所见的图像。

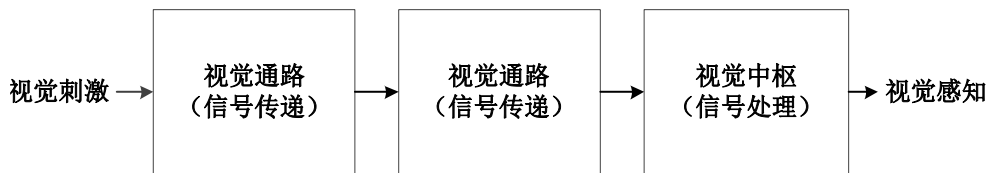


图 2-1 人眼视觉系统的生理结构

2.1.1 人眼与视觉中枢神经系统

(1) 人眼结构

人眼的构造非常精密，每个部分都有其独特的功能，共同协作以实现视觉功能。它由角膜、瞳孔、晶状体、玻璃体和视网膜等组成^[45]，如图 1-6 所示。角膜是眼睛的最外层，透明且有弧度，它的主要功能是聚焦进入眼内的光线。瞳孔是眼角膜和晶状体之间的孔，位于眼球中间。射入眼睛的光量可以通过瞳孔的直径来控制。晶状体是介于瞳孔和玻璃体之间的透明双凸结构，主要功能是聚焦不同场景中的物体，使不同距离的物体清晰可见。玻璃体是一种透明的粘合剂材料。它的主要功能是维持眼部形状，支撑视网膜，稳定眼压。视网膜是眼球的感光部

位,其上分布着大多数锥状和杆状光敏细胞。锥状细胞主要负责色觉和日间视觉,杆状细胞主要负责夜视和弱光视觉。

人眼在接收到光线后,光线会通过角膜进入眼睛,再经过晶状体和瞳孔的调节聚焦后,映射到视网膜上。然后,视网膜受到光学信号刺激产生相应的生物电信号,这些信号将在后续的视觉神经和视觉中枢作用下产生图像。

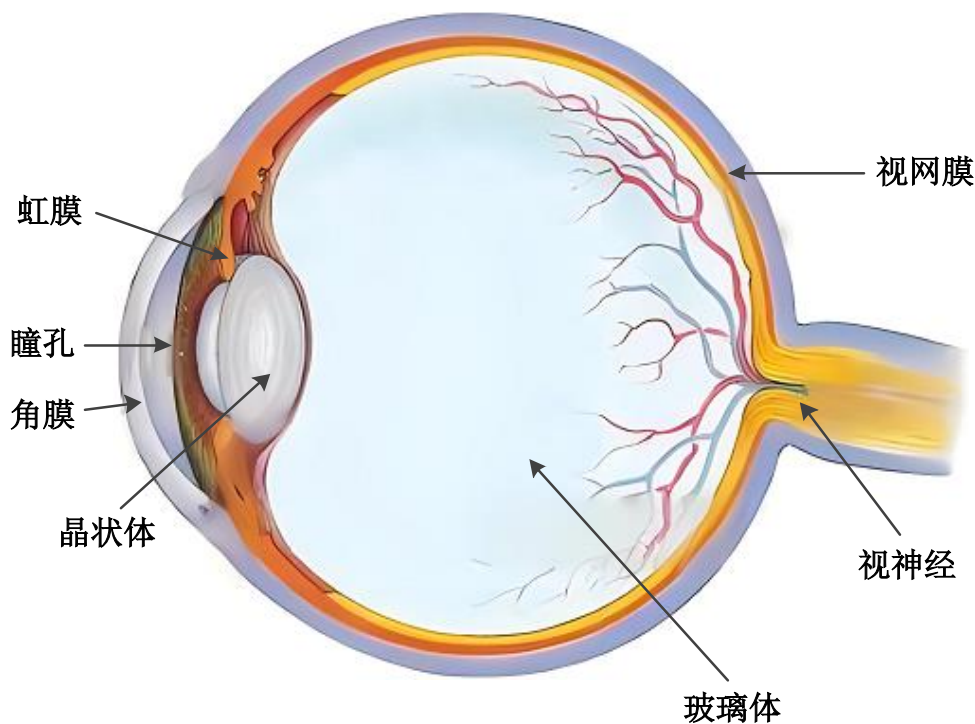


图 2-2 人眼生理结构图

(2) 视觉中枢神经系统

视觉中枢神经系统是负责处理和解释视觉信息的大脑区域,包括视觉皮层和视觉通路等生理结构。如图 2-3 所示。

视皮层 (Visual Cortex) 是人体大脑皮层中主要处理视觉信息的部分,位于大脑的后叶。该区域分为初级视皮层 (Striate Cortex, V1 区) 和纹状体外皮层^[46] (Extrastriate Cortex, V2 和 V3 区等)。初级视皮层 (V1) 是第一个接收和分类所有视觉输入的区域,同时处理分析颜色、亮度等信息的变化。而纹外皮层则包括 V2、V3、V4、V5 等区域,这些区域对视觉信息做进一步的处理和分析。

视觉信息从眼睛通过视觉神经传递到大脑的视觉皮层,沿着特定的神经通路进行处理。主要的视觉通路包括视觉传感通路和视觉联合通路。视觉传感通路负责传递原始的感知信息,包括形状、颜色和运动等方面的信息。而视觉联

合通路则负责整合和分析这些信息，形成对物体、场景和运动的综合认知。

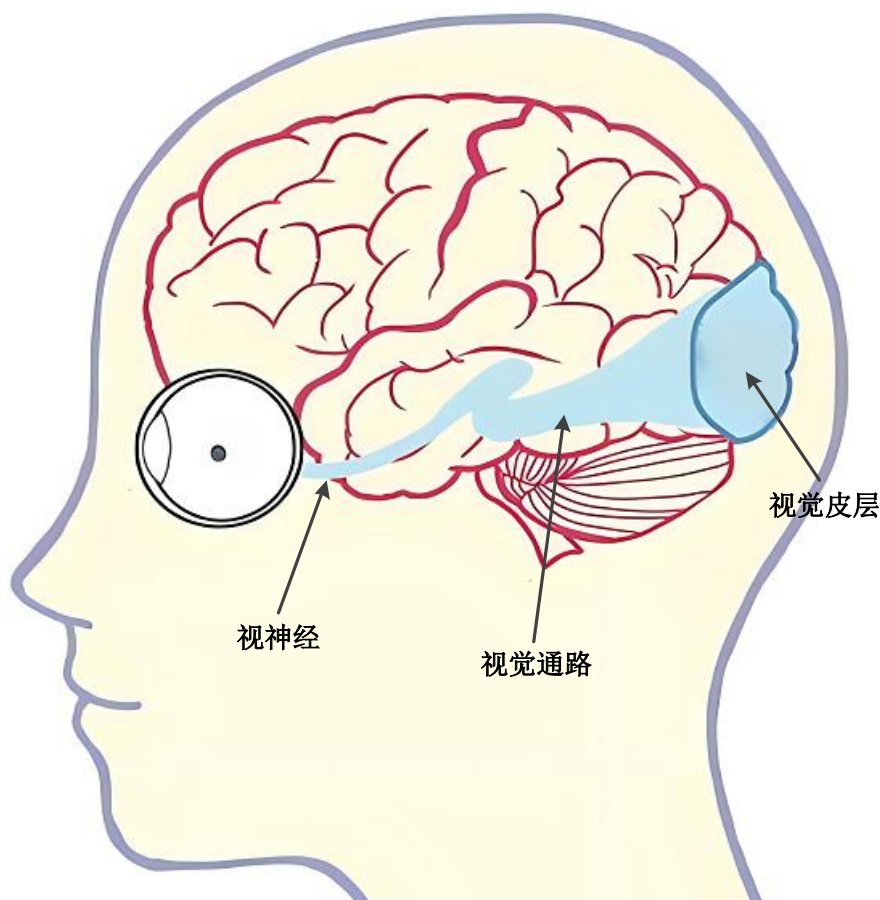


图 2-3 视觉中枢神经系统简图

2.2 人类视觉系统的感知特性

随着人类对 HVS 研究的深入，越来越多的 HVS 的特性也随之出现。这些视觉感知特征模型一般是以心理物理学（Psychophysics）知识为基础的。在这些假设数学模型基础上，对 HVS 的视觉过程进行数值模拟，接着完成对经验公式的实验测量。根据论文的研究内容，本章主要介绍四个视觉特性，分别为对比敏感度、视觉掩盖、最小可觉差和视觉多通道特性。

2.2.1 对比敏感度

对比敏感度一般是指区分物体边界的能力^[47]。人类视觉系统在给定场景中所能检测到的最低亮度变化称为对比度阈值。而对比度灵敏度是对比度阈值的倒数。阈值越小，对比度灵敏度越高，人眼识别边界的能力就越好。

影响人眼感知对比度变化的因素有很多,起到最显著影响的是空间频率和亮度。空间频率对人眼感知对比度的影响可用坎贝尔—罗宾森^[48] (Campbell and Rob-son) 的对比敏感度图来描述。如图 2-4 所示,对比敏感度图为一排竖直放置的正弦光栅。在水平方向,随着空间频率的增加,图像灰度值呈现正弦变化。视觉最敏感的位置是在频率中间区域。亮度也是影响对比度敏感度感知的重要因素,这是因为人眼通过调节感光灵敏度可以分辨一定亮度范围内的物体,但是对于超出或者低于此范围的亮度变化人眼是无法分辨的。对比度敏感度反映了图像边界纹理的变化,是图像质量评价领域最常用的视觉特性。

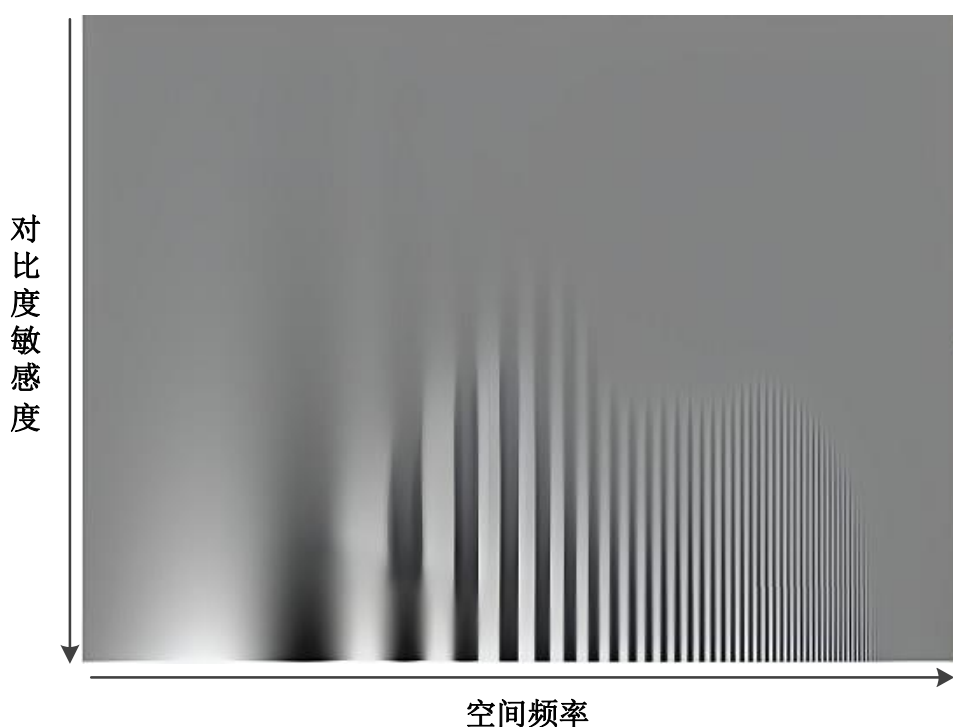


图 2-4 对比敏感度图

2.2.2 视觉掩盖

视觉信息间的相互作用或相互干扰将引起视觉掩盖效应 (Visual Masking)^[49]。例如人眼在识别某一信号时,同时受到一个或者多个其它信号干扰,结果导致人眼识别该信号的能力受到抑制或降低。如图 2-5 所示,图 2-5 (a) 为原始未失真图像,图 2-5 (b) 是添加了噪声失真的原始图像。从图 2-5 (b) 中可以看出,上方天空的明显比下方大地的失真效果更为明显,此时发生了掩盖效应。掩盖效

应揭示了人类视觉系统对于信息处理的局限性。研究发现视觉遮掩效应和信号的带宽、相位、和频率有关。

视觉掩盖效应可以分为两种主要类型^[50]，分别为时空掩盖效应和特征掩盖效应。时空掩盖效应是指掩蔽刺激在时间和空间上干扰目标刺激的感知。在时空掩盖效应中，掩蔽刺激可以在目标刺激之前、之后或同时出现。时空掩盖效应表明，人类视觉系统对于连续和快速的刺激序列具有较低的感知敏感性。特征掩盖效应是指掩蔽刺激与目标刺激在某些特征上相互干扰，导致目标刺激的感知能力降低。这些特征可以是颜色、形状、方向等。特征掩盖效应揭示了人类视觉系统对于特定特征的感知优先级和竞争机制。

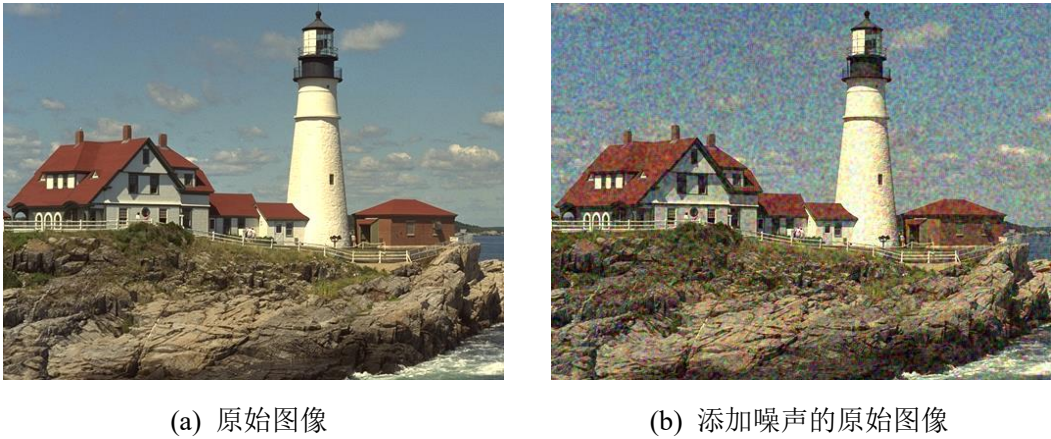


图 2-5 图像遮掩效应示例

2.2.4 最小可觉差和视觉多通道分解

(1) 最小可觉差

人眼在观察不同物体时刚刚能够引起差异感的最小差异量就叫做最小可觉差（Just Noticeable Difference, JND）。在视觉感知中，人眼对不同物体间的差异感比物体本身的绝对差异还要敏感，这个比绝对差异值多出的部分就是 JND。JND 可由下式给出^[51]

$$K = \frac{\Delta I}{I} \quad (2-1)$$

其中 I 表示原始刺激的强度。 ΔI 显示刺激的增加。图像质量评价中最常见的 JND 是亮度自适应（Light Adaption）^[52]。此时式 2-1 中， I 为输入背景亮度， ΔI 代表目标场景的亮度变化值的能被检测到的最小差值， K 为基于亮度灵敏度的阈值

JND。图像质量评价方法中 JND 常被用于衡量灰度图像的亮度变化^[53]。

（2）视觉多通道分解

研究表明，初级视皮层神经元的感受野调制具有特定的空间位置、空间频率、颜色和相位等的视觉刺激^[54]。这表明，神经元只会对与自己适配的信号做出反应，而不会对其它信号有所反应。进一步，HVS 对于不同特征的视觉信号是平行独立处理的，该处理过程被称为视觉多通道（multi-channel）分解模型。

在视觉感知中，可以假设视觉信号被分解成了多个独立的频率波段，每个波段都由特定的响应通道进行处理。由此，常见的分解方法有傅里叶分解^[55]和 Gabor 分解^[56]等。此外，还有很多图像质量评价方法会融合各种互补的特征来描述图像的信息。

2.3 本章小结

HVS 特性不仅与它的生理结构有关，还与某些心理活动有着密不可分的关系。本章主要介绍了 HVS 的生理结构和 HVS 的感知特性。首先从人眼结构和视觉中枢神经系统这两方面出发，描绘了光学信号从采集到人眼中成像的过程，接下来简要的介绍了对比度敏感度、视觉遮掩、最小可觉差和视觉多通道分解四个视觉感知特性。上述的视觉特性有充分详细的心理学和生物学实验支持，本文将基于这些视觉特性对图像质量评价算法进行深入研究。

第3章 基于色貌尺度相位一致性的全参考图像质量评价

上一章中,介绍了人类视觉系统及其特性。本章基于视觉多通道特性,融合三个图像特征,提出了基于色貌尺度相位一致性的全参考图像质量评价方法。论文首先介绍了主要特征即色貌尺度相位一致性的计算过程,然后再对次要特征图像对比度和色度作简单概述,最后依据图像公共数据中的实验结果,将本章算法和其它图像评价算法进行了详细的性能比较。

3.1 色貌尺度相位一致性相似图

色貌尺度相位一致性指的是在色貌图上提取图像相位信息,然后将提取的相位信息进行相似性处理。该方法补偿了传统相位一致性对图像颜色信息的忽视。接下来将分别从获取色貌算子和图像相位两方面来介绍色貌尺度相位一致的构建思路。

3.1.1 色貌算子

Vividness (V_{ab}^*) 是由 CIELAB 匀色空间新定义的一个色貌坐标^[57]。它表示由 L^* 和 C_{ab}^* (见公式 3-4, L^* 和 C_{ab}^* 代表了 CIELAB 匀色空间里的亮度和色度) 定义的颜色到 C_{ab}^* 为 0 的欧几里得距离,当 $L^*=0$ 表示 V_{ab}^* 。 V_{ab}^* 可以代表彩色图像在 HVS 感知方式下的颜色偏差,同时反映了颜色在明度和色度上的变化。

为计算图像的色貌指标 V_{ab}^* 值,首先需要将 RGB 图像映射到 CIELAB 匀色空间中。如 Shi 等^[58]提出的基于图像色貌和梯度特征的图像质量客观评价 (color appearance and gradient similarity, CAGS) 算法,该算法将 RGB 图像转化到对立 CIELAB 匀色空间中,再通过匀色空间中的色度通道计算出对应的色貌指标值。对于一幅相同分辨率的 RGB 参考图像 X 和对应的失真图像 Y,其映射到 CIELAB 均匀色空间的过程如下:

首先将 RGB 图像转为 H, J, K (过渡算子),其中线性变换公式为:

$$\begin{aligned}
H &= 0.412453 \cdot R + 0.357580 \cdot G + 0.180423 \cdot B \\
J &= 0.212671 \cdot R + 0.715160 \cdot G + 0.072169 \cdot B \\
K &= 0.019334 \cdot R + 0.119193 \cdot G + 0.950227 \cdot B
\end{aligned} \tag{3-1}$$

接着归一化处理 H 、 J 和 K 算子，结果用 h 、 j 和 k 表示。然后分别将 h 、 j 和 k 进行非线性映射，其公式为：

$$t = \begin{cases} t^{1/3}, & \text{if } t > (\frac{6}{29})^3 \\ (\frac{1}{3})(\frac{29}{6})^2 \cdot t + \frac{16}{116}, & \text{else} \end{cases} \tag{3-2}$$

其中 t 代表 h 、 j 和 k 。

最后再通过公式 3-3 得到 Lab 图像：

$$\begin{aligned}
L &= 116 \cdot j - 16 \\
a &= 500 \cdot (h - j) \\
b &= 200 \cdot (j - k)
\end{aligned} \tag{3-3}$$

其中 L^* 表示亮度通道、 a^* 表示红绿通道、 b^* 表示蓝黄通道。 V_{ab}^* 数值可由公式 3-4 计算得到

$$V_{ab}^* = \sqrt{(L^*)^2 + (a^*)^2 + (b^*)^2} = \sqrt{(L^*)^2 + (C_{ab}^*)^2} \tag{3-4}$$

其中 V_{ab}^* 为图像的色貌算子，由公式 3-4 可以看出， V_{ab}^* 同时反映亮度和色度的变化。

3.1.2 获取色貌尺度相位一致性相似图

由上节可知，色貌算子 V_{ab}^* 值可由公式 3-4 计算得到。而本章所研究的色貌尺度相位一致性是在 V_{ab}^* 图上提取图像的相位信息。因此，为了获取图像色貌尺度相位一致性特征，接下来需要计算 V_{ab}^* 图上的相位值。根据文献[59]，目前有多种方法可以计算给定图像的相位值。本文参照文献[60]中的方法去计算相位值。一维图像信号用 $g(x)$ 表示， M_n^e 和 M_n^o 为尺度 n 上的奇偶对称滤波器，由其生成的每个正交对在 x 处的响应矢量为： $[e_n(x), o_n(x)] = [g(x) * M_n^e, g(x) * M_n^o]$ 。设 $F(x) = \sum_n e_n(x)$ 以及 $H^2(x) = \sum_n o_n(x)$ ，则一维相位可计算为：

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_n A_n(x)} \quad (3-5)$$

其中, $A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2}$ 为 x 处的局部振幅, $E(x) = \sqrt{F^2(x) + H^2(x)}$ 为 x 处的局部能量。 ε 是一个小的正常数。这里选择 log-Gabor 滤波器来处理此处产生的正交对^[22], 其传递函数为 $G(w) = \exp\left(-(\log(w/w_0))^2 / 2\sigma_r^2\right)$, 其中 w_0 是滤波器的中心频率, σ_r 为控制滤波器带宽。

二维 Gabor 传递函数可由一维叠加扩展得到。高斯函数是一个应用广泛的扩展函数, 且 PC 值对高斯运算不敏感^[24]。故二维 Gabor 传递函数可以表示为:

$$G_2(\omega, \theta_j) = \exp\left(-\frac{\left(\log\left(\frac{\omega}{\omega_0}\right)\right)^2}{2\sigma_r^2}\right) \cdot \exp\left(-\frac{(\theta - \theta_j)^2}{2\sigma_\theta^2}\right) \quad (3-6)$$

其中 $\theta_j = j\pi / J, j = \{0, 1, \dots, J-1\}$ 为滤波器的方向角, J 为方向数, σ_θ 决定滤波器的角度带宽。选择合适的 θ_j 和 ω , 然后将 G_2 与二维图像卷积, 则图像的相位图可表示为:

$$PC = \frac{\sum_j E_{\theta_j}(X)}{\varepsilon + \sum_n \sum_j A_{n, \theta_j}(X)} \quad (3-7)$$

其中, X 表示被卷积的图像, $A_{n, \theta_j}(v) = \sqrt{e_{n, \theta_j}(v)^2 + o_{n, \theta_j}(v)^2}$ 为局部振幅、 $E_{\theta_j}(v) = \sqrt{F_{\theta_j}(v)^2 + H_{\theta_j}(v)^2}$ 为延 θ_j 方向的局部能量, ε 为小的正常数。将式中的 X 替换为 v_{ab}^* 后, 色貌尺度相位图可表示为

$$CPC = \frac{\sum_j E_{\theta_j}(v)}{\varepsilon + \sum_n \sum_j A_{n, \theta_j}(v)} \quad (3-8)$$

对于一个图像对 (X 和 Y), 从 v_{ab}^* 图上提取的相位映射分别定义为 CPC_1 和

CPC₂。CPC 相似性计算公式为：

$$S_p = \frac{2CPC_1 \cdot CPC_2 + C_1}{CPC_1^2 + CPC_2^2 + C_1} \quad (3-9)$$

其中 C 是提高等式稳定性的常数。由于色空间映射可能导致不可避免的色差问题，从而影响了在颜色通道中提取 CPC 相似图的准确性。为此，绘制了两对原始失真图像的 PC 相似图和 CPC 相似图，并提供了对应的 MSE 值。图 3-1 展示了这一点，其中图 3-1 (e) 和图 3-1 (f) 亮点较少，这说明本文提出算法能够有效地表征图像差异性能。进一步分析，图 3-1 (e) 和图 3-1 (f) 之间 PC 图的 MSE 值分别为图 3-1 (g) 和图 3-1 (h) 之间 CPC 图的 2.25% 和 0.17%，这表明 CPC 算法比 PC 算法对于捕捉图像差异具有更高敏感度。同时根据主观评分结果可以看出，CPC 算法与人眼感知结果有很好的 consistency，且色空间映射不会对 CPC 相似性的精度造成影响。

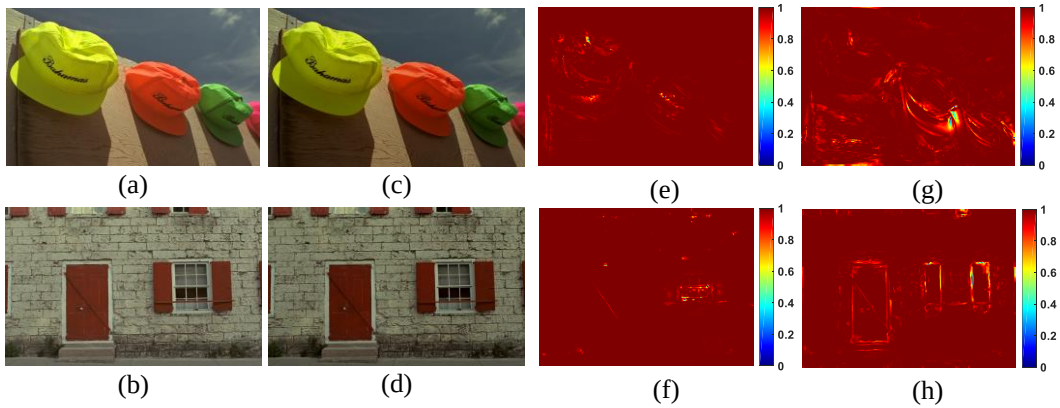


图 3-1 CPC 算法和传统 PC 算法性能对比图。(a)和(b)为原始图像。(c)和(d)为失真图像,对应主观评依次为 6.4333 和 5.6286。(e)和(f)为传统 PC 相似图。(g)和(h)为 CPC 相似图。(a)和(c)的 PC 图之间 MSE 为 0.000014, CPC 图之间 MSE 为 0.000622。(b)和(d)的 PC 图之间 MSE 为 0.000002, CPC 图之间 MSE 为 0.001200。

3.2 对比度和色度相似图

(1) 对比度相似图

图像对比度相似性是表征图像间对比度差异的尺度。对比度是指图像不同灰度度间的差异。根据文献[61]所述，在计算对比度时常采用 RMS 方法作为理想选

择。本文也选择了 RMS 对比度作为全局对比度指标来衡量结构变化程度，并通过实验结果证明其与自然图片主观感知一致性较高。全局对比度表达式为：

$$LC_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2} \quad (3-10)$$

其中 $\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$ 为图像的平均强度。原始图像 (X) 和失真图像 (Y) 的对比度相似图可表示为：

$$S_{LC}(X, Y) = \frac{2LC_X \cdot LC_Y + C_2}{LC_X^2 + LC_Y^2 + C_2} \quad (3-11)$$

其中 C_2 是增加稳定性的常数。进一步验证对比度特性对图像质量度量的重要性，这里绘制了一幅原始失真图像对的对比度图，如图 3-2 所示。其中图 3-2 (X) 和图 3-2 (Y) 表示参考图像和失真图像。图 3-2 (a1) 和图 3-2 (b1) 为对应的对比度图。从图 3-2 中可以观察到不同对比度图之间存在明显的差异，表明选择 RMS 算法计算图像对比度符合人眼感知结果。

(2) 色度相似图

图像色度相似性是用于比较两幅图像间色彩分布差异度的尺度。色度是指图像颜色信息，包括颜色的亮度、饱和度和色调等特征^[62]。图像色度相似性的计算可以帮助我们了解两幅图像在颜色方面的相似性或差异性本章使用 a^* 和 b^* 两个颜色通道计算图像色度相似度，其表达式为：

$$S_q = I(X, Y) = \frac{2a_1 \cdot a_2 + C_3}{a_1^2 + a_2^2 + C_3} \cdot Q(X, Y) = \frac{2b_1 \cdot b_2 + C_4}{b_1^2 + b_2^2 + C_4} \quad (3-12)$$

其中 C_3 和 C_4 是增加稳定性的常数。 a_1 、 b_1 和 a_2 、 b_2 分别为 X 和 Y 上的色度通道。这里提供了一对原始失真图像的色度图，如图 3-2 所示。其中，图 3-2 (a2) 和图 3-2 (b2) 为对应色度图。可以观察到两个色度图间存在明显差异。这表明选用的色度特征可以量化图像颜色变化。

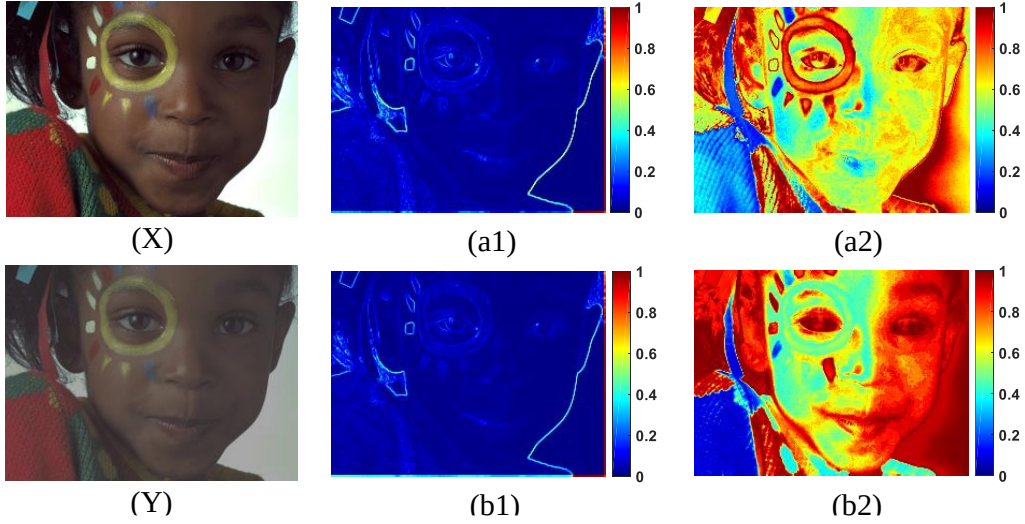


图 3-2 图像对比度以及色度的差异比较

3.4 基于色貌尺度相位一致性的全参考图像质量评价算法

结合上述计算得到的特征分量,可获得本章所提出的基于色貌尺度相位一致性的全参考图像质量评价(color appearance-based phase consistency, CPCC)算法,计算过程如下。

通过公式 3-13 汇集池化空间中的上述三个分量:

$$S = w_1 \cdot SD(\sum_{\Omega} S_{LC}) + w_2 \cdot SD(\sum_{\Omega} S_q) + w_3 \cdot SD(\sum_{\Omega} S_p) \quad (3-13)$$

其中, w_1 、 w_2 和 w_3 是代表权值系数,且 $w_1 + w_2 + w_3 = 1$ 。标准差(Standard Deviation,

SD)^[63]计算公式为: $SD(i) = \sqrt{\frac{1}{N} \sum_{i=1}^N (i - \bar{i})^2}$, $\bar{i} = \frac{1}{N} \sum_{i=1}^N i$ 。其中 i 表示图像特征相似度值。

为了验证所选特征可靠性,这里选用 3.3 节中的测试图像作为输入图像进行验证,如图 3-3 所示。可以观察到,图片 3-3 (a) 左下角的衣领处为浅红色,这表明相似度较低,表征差异性的能力较高。因此,本章方法可以提供关于图像结构的更全面的信息。

在图 3-3 (b) 中,人的面部特征鲜艳,特别是左眼,表示面部特征的相似度较低,符合实际场景。在图 3-3 (c) 中,整幅图像的左侧为蓝色,相似度较低。因此本文的特征选择是合理可靠的。

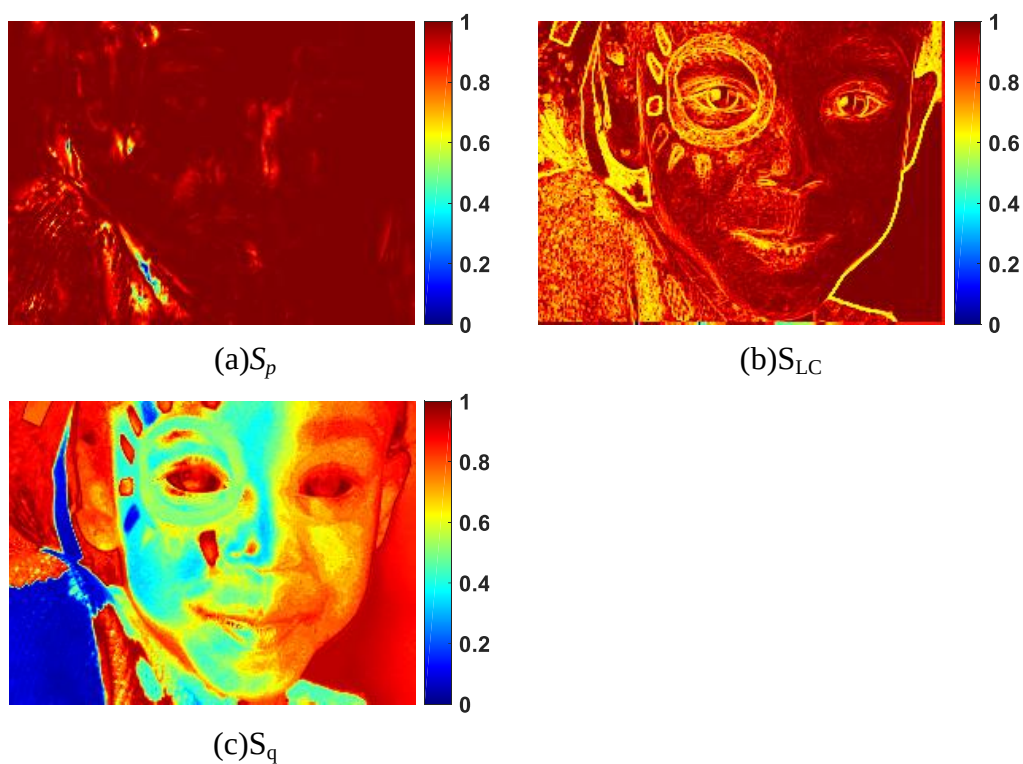


图 3-3 图像的特性特征相似图

所提算法的流程图如图 3-4 所示。首先输入图像信号 X 和 Y ，接着分别将其转换到对立的颜色空间 CIELab 中，然后在 L 、 a 和 b 三个图像信号分量上计算色貌尺度上相位相似度、对比度相似度和色度相似度。最后池化这三个特征相似度，得到最终的图像质量分数。

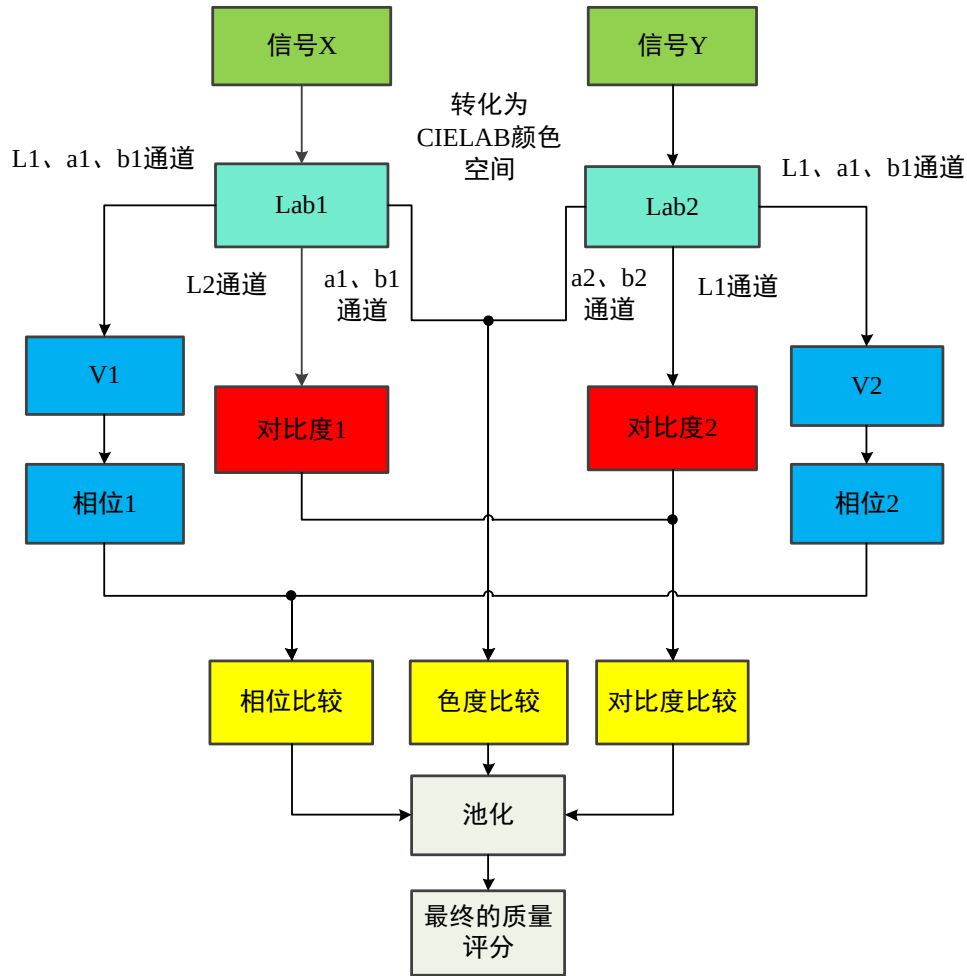


图 3-4 算法流程图

3.5 实验结果与分析

3.5.1 数据库和模型评价标准

这里选择了 TID2008, CSIQ, LIVE 和 TID2013 这四个常用的图像公共数据库来测试不同图像质量评价方法的预测性能,以上四个数据库在前文中已详细介绍,为避免文章冗长,这里不再过多的描述(下同)。表 3-1 列举了这四个数据库的基本信息。

表 3-1 IQA 基准测试数据库

数据库	参考图像	失真图像	失真类型	观测者
TID2013	25	3000	24	971
TID2008	25	1700	17	838
CSIQ	30	866	6	35
LIVE	29	779	5	161

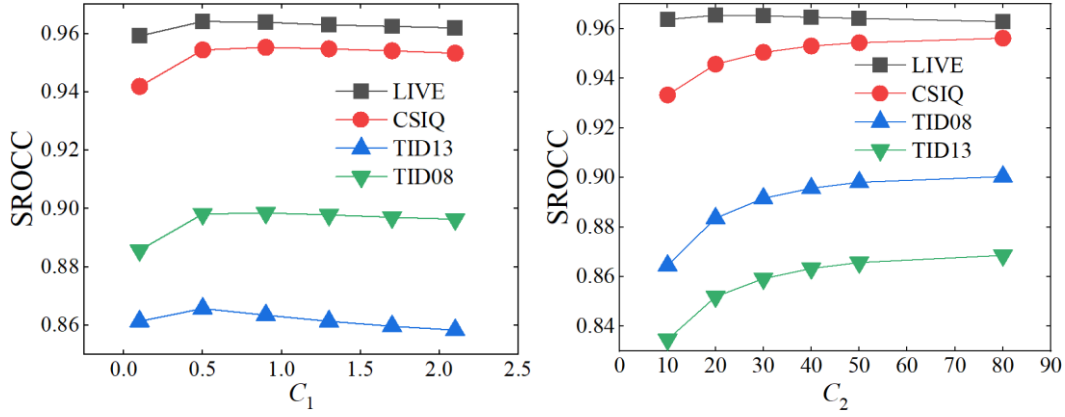
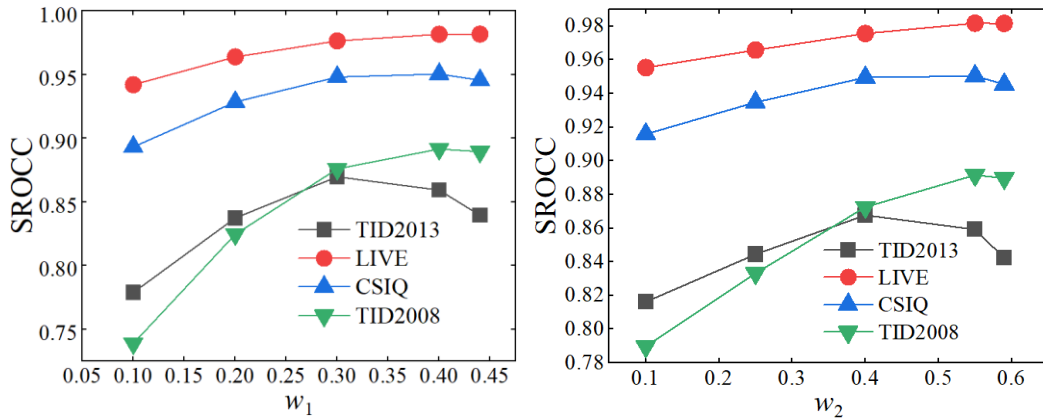
为验证数据精度，采用四个常用的评价标准 SROCC、KROCC、PLCC 和 RMSE 来分析主客观分数之间的相关性。此外，为了保证评估过程在同一量纲下，这里通过式 1-7 将客观分数向主观分数映射。

3.5.2 算法参数优化

根据前文所述，本章的模型参数包括 C_1 、 C_2 、 C_3 、 C_4 、 w_1 、 w_2 和 w_3 。其中，直接采用 CVSS 中的修正数值作为 C_3 和 C_4 的值。为了确定 C_1 、 C_2 、 w_1 、 w_2 和 w_3 的取值，论文使用了控制变量法。通过计算和比较四个数据库中的 SROCC 值，能够确定这些参数的最佳取值。

首先，在保持 C_1 不变的情况下，调整 C_2 的取值范围。由于 C_2 值过大或过小都可能影响特征表现和计算稳定性，因此将其限定在[10,80]区间内。根据图 3-5 显示的结果，在四个数据库中 SROCC 值达到最理想状态时， C_2 处于区间[25,45]内。经过进一步细化选择后，确定最佳参数为 $C_2=30$ 。接下来采用类似方法优化 C_1 。首先调整 C_1 处于[0,2.5]范围内。当各数据库中 SROCC 值达到最优时， C_1 处于区间[0.5,0.8]内。这里设置 $C_1=0.5$ 。

对于 w_1 、 w_2 和 w_3 的取值也遵循相同逻辑和方法进行优化选择。考虑到这三个变量之间满足 $w_1+w_2+w_3=1$ 的关系，只需确定其中两个变量即可推导出第三个变量。图 3-6 展示了在固定 w_2 值时通过调整 w_1 取得结果。根据图表分析，在四个数据库中 SROCC 值达到最理想状态时，将 w_1 设置在[0.3,0.4]区间内效果较好。因此决定选取 $w_1=0.35$ 作为最佳参数。同样地保持 w_1 不变，当将 w_2 设置在[0.4,0.55]区间内时，所提模型 SROCC 值达到最优状态，故选取 $w_2=0.5$ 。


 图 3-5 不同的 C_1 和 C_2 值在四个基准数据库上的 SROCC 值

 图 3-6 不同的 w_1 和 w_2 值在四个基准数据库上的 SROCC 值

3.5.3 算法性能比较

(1) 整体性能比较

为了验证该模型的性能优势,这里选择了十个经典算法进行比较,包括 SSIM、CVSS、GSM、FSIMc、CAGS 以及最明显的失真^[64](Most Apparent Distortion,MAD)、梯度幅相似性偏差^[65](Gradient magnitude similarity deviation,GMSD)、图像信息与视觉质^[66](Image information and visual quality, VIF)、视觉显著性指数^[67](A Visual Saliency-Induced index for perceptual image quality assessment,VSI)和基于 Riesz 变化和视觉对比敏感度的特征相似性指数^[68](Riesz transform and Visual contrast sensitivity-based feature SIMilarity index,RVSIM)。四个数据库中表现最好的前三名已加粗标出,且对每个数据库中的指标结果还进行了直接平均和加权平均处理。

根据表 3-2 中的数据,可以看出本文模型在 TID2013、TID2008 和 CSIQ 这

三个常用的图像公共质量评估数据库上的表现确实非常出色，其整体性能位列前三，这充分证明了模型的有效性和可靠性。尤其在 CSIQ 数据库中，本文模型的性能评价标准值与前两名极为接近，显示出模型在该数据库上的强大竞争力。与同样采用了 PC 特征的 FSIMc 算法相比，本文模型在四个数据库中的数据都具有较明显优势。这表明本文模型在提取和利用图像特征信息方面可能更加有效，或者模型在算法设计和实现上更加符合人类视觉系统的感知特性。此外，在前三个指标

的加权平均和直接平均值中，所提模型的性能仅次于 VSI，位列第二。在 TID2008 数据库中，所提模型的 PLCC 值位列第一，说明模型在该数据库上的预测结果与真实值之间的线性相关性最好，进一步验证了模型的有效性和准确性。

根据表 3-2 的结果，本文模型在前三名中获得了最高的次数（18 次），其次是 CVSS 和 VSI 模型（各 14 次）。此外，从表 3-2 可以观察到，本文模型的 SROCC 和 PLCC 值均超过 0.8592，这说明该模型具备出色的预测性能和良好的泛化能力。综上所述，本文模型在多个图像质量评估数据库上均表现出良好的性能和精度值，充分证明了模型的有效性和竞争力。这些结果也为后续的研究和改进提供了有力的支持和参考。

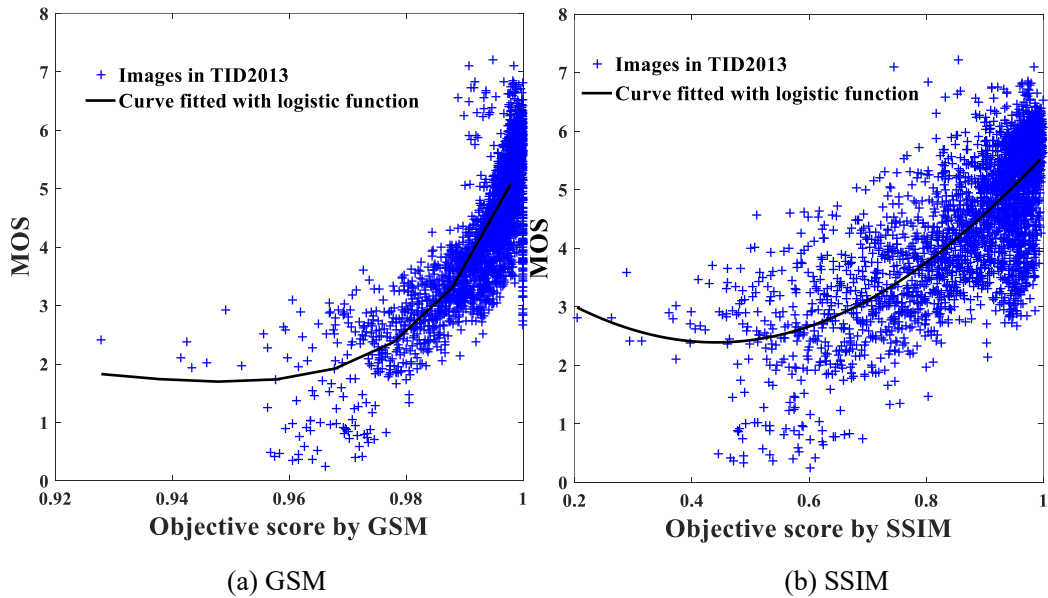
表 3-2 不同的 IQA 模型在 4 个数据库中的性能比较

数据库	SSIM	VIF	MAD	FSIMc	GMSD	VSI	GSM	RVISM	CVSS	CAGS	proposed	
TID2013	SROCC	0.7417	0.6769	0.7807	0.8510	0.8044	0.8965	0.7946	0.6757	0.8069	0.8316	0.8592
	KROCC	0.5588	0.5147	0.6035	0.6665	0.6339	0.7183	0.6255	0.5146	0.9001	0.6469	0.6738
	PLCC	0.7895	0.7720	0.8267	0.8769	0.8583	0.9000	0.8464	0.7825	0.8406	0.8445	0.8781
	RMSE	0.7608	0.7880	0.6975	0.5959	0.6360	0.5404	0.6603	0.7719	0.6715	0.6639	0.5932
TID2008	SROCC	0.7749	0.7491	0.8340	0.8840	0.8907	0.8979	0.8504	0.7375	0.8961	0.8231	0.8916
	KROCC	0.5768	0.5860	0.6445	0.6991	0.7092	0.7123	0.6596	0.5628	0.9580	0.6289	0.7040
	PLCC	0.7732	0.8084	0.8308	0.8762	0.8788	0.8762	0.8422	0.7954	0.5956	0.8091	0.8868
	RMSE	0.8511	0.7899	0.7468	0.6468	0.6404	0.6466	0.7235	0.8133	0.7215	0.7886	0.6201
CSIQ	SROCC	0.8756	0.9195	0.9466	0.9310	0.9570	0.9423	0.9108	0.8979	0.9580	0.9198	0.9505
	KROCC	0.6907	0.7537	0.7970	0.7690	0.8129	0.7857	0.7374	0.7234	0.8171	0.7487	0.8009
	PLCC	0.8613	0.9277	0.9502	0.9192	0.9541	0.9279	0.8964	0.9236	0.9589	0.9014	0.9537
	RMSE	0.1334	0.0980	0.0818	0.1034	0.0786	0.0979	0.1164	0.1007	0.0745	0.1137	0.0790

	SROCC	0.9479	0.9636	0.9669	0.9645	0.9603	0.9524	0.9561	0.9600	0.9672	0.9734	0.9653
LIVE	KROCC	0.7963	0.8282	0.8421	0.8363	0.8268	0.8058	0.8150	0.8203	0.8406	0.8658	0.8356
	PLCC	0.9449	0.9604	0.9675	0.9613	0.9595	0.9482	0.9512	0.9570	0.9651	0.9640	0.9616
	RMSE	8.9455	7.6137	6.9073	7.5296	7.6937	8.6816	8.4327	7.9274	7.1573	8.3251	7.4983
	SROCC	0.7942	0.7646	0.8405	0.8847	0.8544	0.9100	0.8452	0.7638	0.8751	0.8588	0.8964
权重平均	KROCC	0.6108	0.6049	0.6702	0.7101	0.7022	0.7366	0.6732	0.6006	0.7115	0.6828	0.7300
	PLCC	0.8140	0.8261	0.8619	0.8928	0.8892	0.9033	0.8650	0.8307	0.8893	0.8575	0.9010
	SROCC	0.8350	0.8273	0.8821	0.9076	0.9031	0.9223	0.8780	0.8178	0.9080	0.8870	0.9167
直接平均	KROCC	0.6557	0.6707	0.7218	0.7427	0.7457	0.7555	0.7094	0.6553	0.7531	0.7226	0.7536
	PLCC	0.8422	0.8671	0.8938	0.9084	0.9127	0.9131	0.8841	0.8646	0.9152	0.8798	0.9201

图 3-7 直观地展示了不同模型与主观观测值之间的拟合性能，其中主观分数离散点由 TID2008 提供。

通过离散点的分布，可以清晰地看到所提模型预测值与主观观测值之间的相关性较好。这一观察结果验证了所提模型在图像质量评价方面的优势。通过离散点比较结果，进一步证明了所提模型在拟合性能上的可靠性。这一优势来源于模型对 HVS 特性的融合、多种信息的提取以及亮度和色度协调变化的考虑。这些因素共同提升了模型的精度和泛化性。



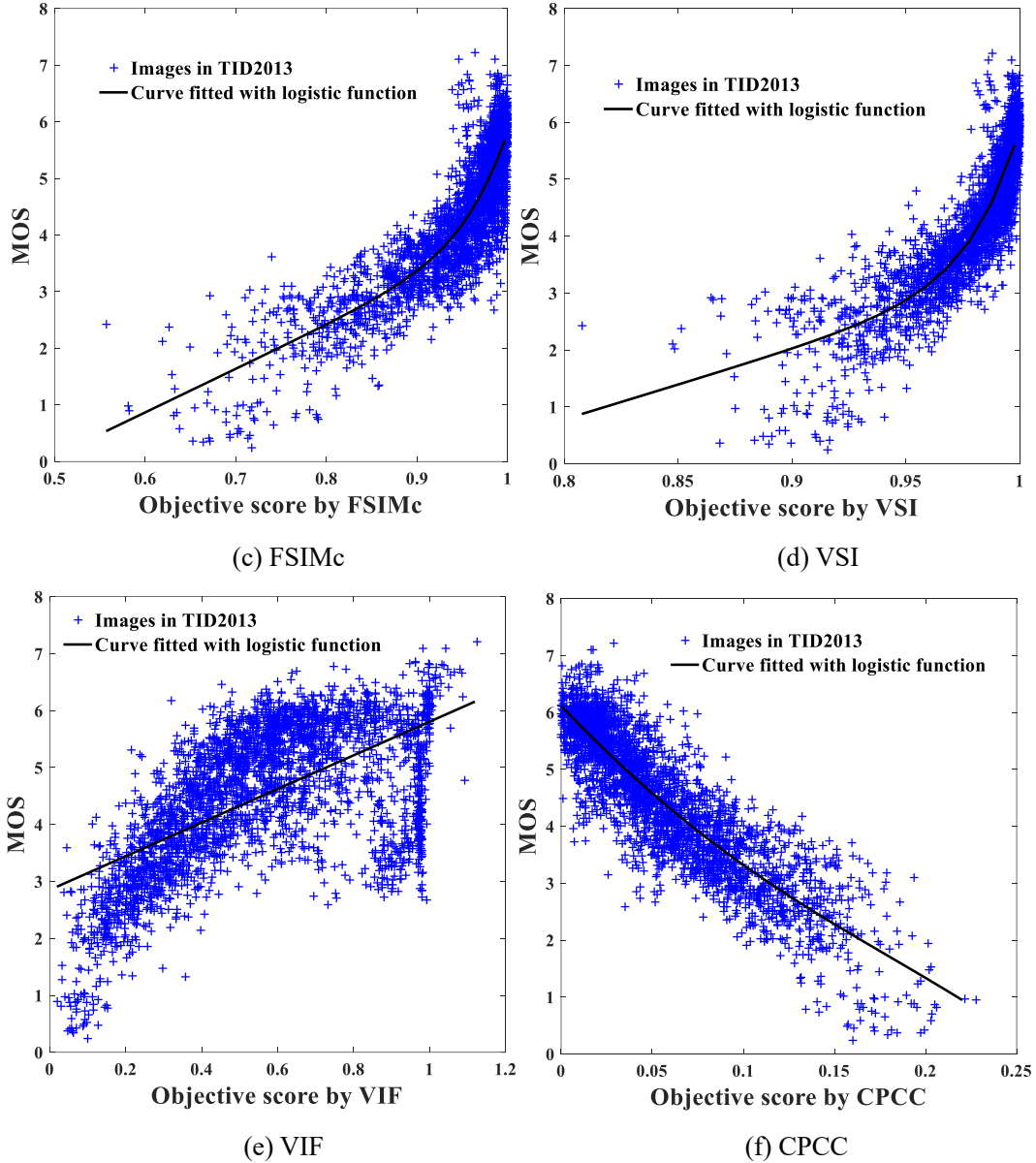


图 3-7 TID2013 数据库上的主观 MOS 和模型预测得分的散点图

(2) 单个失真类型的性能比较

分析模型在不同失真类型下的预测性能是评估图像质量评估模型全面性的重要步骤。考虑到 TID2008 数据库与 TID2013 数据库中的失真类型存在重叠，为避免冗余，选择 TID2013 数据库进行性能分析（下同）。表 3-3 提供了在单个失真类型下各模型的 SROCC 值比较数据，SROCC 是一种非参数统计方法，用于衡量预测排序与真实排序之间的相关性，其值范围在-1 到 1 之间，值越大表示相关性越好。根据表 3-3 中的数据，可以看到 VSI 模型在 17 个失真类型中表现最好，紧随其后的是 CVSS 模型在 15 个失真类型中表现出色。而本文模型在 14 个失真类型中的 SROCC 值较为理想，在表 3-3 所列的 11 个算法中位列前三，这

一结果显示了本文模型在不同失真类型下的预测性能较为稳定。值得注意的是，虽然本文模型在某些失真类型下并未获得最佳性能，但其总体表现仍然非常出色。这表明本文模型在提取和利用图像特征以评估图像质量方面具有良好的通用性和鲁棒性。此外，与其他模型相比，本文模型在多个失真类型下都取得了不错的成绩，显示出其在应对不同类型失真时的适应性。综上所述，本文模型在不同失真类型下的预测性能表现稳定且优秀，与 VSI 和 CVSS 等先进模型相比具有相当的竞争力。这一结果进一步验证了本文模型的有效性和可靠性，为图像质量评价领域提供了有益的探索和实践。

深入分析表 3-3 中的数据，可以看到在对比度变化（Contrast change, CTC）和颜色饱和度变化（Change of Color Saturation, CCS）这两种失真类型上，本文所提模型的 SROCC 值虽然相对较低，但仍然优于 VSI 和 CVSS 模型。具体来说，在 CTC 失真类型上，本文模型的 SROCC 值为 0.3147，而 VSI 和 CVSS 的 SROCC 值分别为 0.3054 和 0.2940。在 CCS 失真类型上，本文模型的 SROCC 值为 0.5840，优于 VSI 的 SROCC 值 0.2748 和 CVSS 的 SROCC 值 0.2614。这表明本文模型在应对这两种较难解决的失真类型时，仍具有一定的竞争力。此外，在局部块状扭曲（Local block-wise distortions of different intensity,Block）失真类型上，本文模型的 SROCC 值为 0.7562，远高于 VSI 的 0.1713 和 CVSS 的 0.5152。这一结果显示本文模型在 Block 失真类型上的预测性能显著优于其他两个模型。综上所述，尽管在某些特定失真类型上本文模型的 SROCC 值相对较低，但对于较难解决的失真类型如 CTC 和 CCS，本文模型展现出了一定的竞争力，为后续的研究提供了新的思路和方法。

表 3-3 IQA 模型在不同失真类型下的 SROCC 值

数据库		SSIM	VIF	MAD	FSIMc	GMSD	VSI	GSM	RVISM	CVSS	CAGS	proposed
TID2013	AGN	0.8671	0.8994	0.8843	0.9101	0.9462	0.9460	0.9064	0.8645	0.9401	0.9359	0.9325
	ANC	0.7726	0.8299	0.8019	0.8537	0.8684	0.8705	0.8175	0.8124	0.8639	0.8653	0.8434
	SCN	0.8515	0.8835	0.8911	0.8900	0.9350	0.9367	0.9158	0.8412	0.9077	0.9276	0.9119
	MN	0.7767	0.8450	0.7380	0.8094	0.7075	0.7697	0.7293	0.8228	0.7715	0.7526	0.7815
	HFN	0.8634	0.8972	0.8876	0.9040	0.9162	0.9200	0.8869	0.8837	0.9097	0.9159	0.9080
	IN	0.7503	0.8537	0.2769	0.8251	0.7637	0.8741	0.7965	0.8823	0.7457	0.8361	0.7814
	QN	0.8657	0.7854	0.8514	0.8807	0.9049	0.8748	0.8841	0.7519	0.8869	0.8718	0.8993
	GB	0.9668	0.9650	0.9319	0.9551	0.9113	0.9612	0.9689	0.9742	0.9348	0.9614	0.9268
	DEN	0.9254	0.8911	0.9252	0.9330	0.9525	0.9484	0.9432	0.8947	0.9427	0.9466	0.9466
	JPEG	0.9200	0.9192	0.9217	0.9339	0.9507	0.9541	0.9284	0.9306	0.9521	0.9585	0.9463
	JP2K	0.9468	0.9516	0.9511	0.9589	0.9657	0.9706	0.9602	0.9521	0.9587	0.9620	0.9547
	JGTE	0.8493	0.8409	0.8283	0.8610	0.8403	0.9216	0.8512	0.8434	0.8613	0.8644	0.8832
	J2TE	0.8828	0.8761	0.8788	0.8919	0.9136	0.9228	0.9182	0.8854	0.8851	0.9250	0.8805
	NEPN	0.7821	0.7720	0.8315	0.7937	0.8140	0.8060	0.8130	0.7502	0.8201	0.7833	0.8318
	Block	0.5720	0.5306	0.2812	0.5532	0.6625	0.1713	0.6418	0.6037	0.5152	0.6015	0.6620
	MS	0.7752	0.6276	0.6450	0.7487	0.7351	0.7700	0.7875	0.6079	0.7150	0.7441	0.7107
	CTC	0.3775	0.8386	0.1972	0.4679	0.3235	0.4754	0.4857	0.1526	0.2940	0.4514	0.3147
	CCS	0.4141	0.3099	0.0575	0.8100	0.2948	0.2748	0.3578	0.3994	0.2614	0.3711	0.5840
	MGN	0.7803	0.8468	0.8409	0.8569	0.8886	0.9117	0.8348	0.8233	0.8799	0.8700	0.8688
	CN	0.8566	0.8946	0.9064	0.9243	0.9298	0.9121	0.9124	0.8984	0.9351	0.9168	0.9345
LCNI	0.9057	0.9204	0.9443	0.9485	0.9629	0.9564	0.9563	0.9156	0.9629	0.9574	0.9656	
ICQD	0.8542	0.8414	0.8745	0.8815	0.9102	0.8839	0.8973	0.8036	0.9108	0.9060	0.9061	
CHA	0.8775	0.8848	0.8310	0.8925	0.8530	0.8906	0.8823	0.9154	0.8523	0.8768	0.8577	
SSR	0.9461	0.9353	0.9567	0.9576	0.9683	0.9628	0.9668	0.9439	0.9605	0.9580	0.9603	
CSIQ	AGWN	0.8974	0.9575	0.9541	0.9359	0.9680	0.9636	0.9440	0.9377	0.9670	0.9652	0.9490
	JPEG	0.9546	0.9705	0.9615	0.9664	0.9650	0.9618	0.9632	0.9599	0.9689	0.9573	0.9682
	JP2K	0.9606	0.9672	0.9752	0.9704	0.9720	0.9694	0.9648	0.9693	0.9777	0.9545	0.9730
	AGPN	0.8922	0.9511	0.9570	0.9370	0.9500	0.9638	0.9387	0.9319	0.9516	0.9492	0.9326
	GB	0.9609	0.9745	0.9602	0.9729	0.9710	0.9679	0.9589	0.9715	0.9789	0.9574	0.9754
	GCD	0.7922	0.9345	0.9207	0.9438	0.9040	0.9504	0.9354	0.8792	0.9324	0.9273	0.9274
LIVE	JP2K	0.9614	0.9696	0.9676	0.9724	0.9710	0.9604	0.9700	0.9622	0.9719	0.9822	0.9707
	JPEG	0.9764	0.9846	0.9764	0.9840	0.9780	0.9761	0.9778	0.9822	0.9836	0.9836	0.9804
	AWGN	0.9694	0.9858	0.9844	0.9716	0.9740	0.9835	0.9774	0.9692	0.9809	0.9837	0.9669
	GB	0.9517	0.9728	0.9465	0.9708	0.9570	0.9527	0.9518	0.9643	0.9662	0.9641	0.9689
	FF	0.9556	0.9650	0.9569	0.9519	0.9420	0.9430	0.9402	0.9581	0.9592	0.9633	0.9680

(3) 模型计算效率比较

为了更全面地评估本文模型在实际应用中的性能,这里对其计算复杂度进行了详细的分析和比较。为此,选用了 TID2008 数据库中的一对分辨率为 512×512 的彩色图像作为测试对象,并在内存为 2.3 GHz Intel Core i5 的计算机上运行仿真,仿真软件为 MATLAB R2016a。

实验结果表明,本章所研究的评价模型计算复杂度较低。如表 3-4 所示,与其它六种模型相比,所提模型的计算效率仅次于 PSNR、SSIM 和 CAGS,位列第四。但在忽略 RGB 到 CIELAB 颜色空间转换所需要的时间,所提模型仅需要 0.371 秒的运行时间。这意味着在实际应用中,只需提前完成颜色空间转换的预处理工作,即可迅速完成相关任务,满足高实时性要求的应用场景。

表 3-4 不同 FR-IQA 模型的计算效率对比

IQA Index	Time cost /s	IQA Index	Time cost/s
SSIM	0.278	RVSIM	0.939
PSNR	0.156	MAD	1.256
CAGS	0.483	CPCCs	0.531
VSI	0.536		

3.6 本章小结

研究图像相位一致性,对图像质量评价领域有着深远的意义。针对彩色图像,本章考虑到图像颜色分量对获取图像相位信息的贡献,引入色貌指标 Vividness 计算图像相位值,提出了色貌尺度相位一致性方法。经过滤波处理后的图像对比度发生改变,故本章将色貌尺度一致性、对比度和色度三种图像特征进行融合,提出了基于色貌尺度相位一致性的全参考图像质量评价方法。在四个常用公共图像数据库上,所提方法与其它九个经典的 FR-IQA 模型运行结果进行了比较分析。实验结果表明,所提算法无论在整体性能预测还是在单一失真性能预测都表现出可靠的稳定性和精度。

第4章 基于多通道融合的全参考图像质量评价

前一章中，使用高斯低通滤波器提取了彩色图像中的相位信息，并给出了质量评分，然而对于高斯滤波器并未做进一步描述。事实上，高斯滤波器可以减少图像噪声，适用于含有大量噪声的图像。但面对边缘纹理复杂的图像，显然并不适用。高斯拉普拉斯算子（Laplacian of Gaussian, LoG）^[69]能够检测图像边缘纹理和降低图像噪声，是图像质量评价领域中常用的高通滤波器。如感知均匀色域中的多分辨率图像质量评估（Multi-resolution image quality assessment in the perceptually uniform color domain, PerSIM）^[70]和 FSIM 等。但是上述算法仅使用 LoG 滤波器提取图像亮度通道上的结构信息，忽略了颜色分量对结构信息的贡献。基于此，本章通过在图像三个颜色分量上分别使用 LoG 滤波处理，提出了基于多通道融合的全参考图像质量评价方法。接下来将介绍多通道 LoG 滤波的提取过程，然后与现有的经典模型进行性能比较，最后对本章内容做出总结。

4.1 多通道信息融合

4.1.1 LoG 的定义

LoG 是在高斯函数的基础上利用拉普拉斯算子提取边缘得出的一个算子。主要用于图像边缘和纹理的检测。文献[71]中提到，高斯差分（Difference of Gaussian, DoG）可用于模拟猫的视觉特性，但在构造 DoG 模型时，通常需要调试不同的标准差和尺度差以获得理想的特征差异，但不同尺度下模型的融合非常复杂。当调整尺度时，高斯算子的差可以用高斯二阶导数来近似，高斯二阶导数对应于高斯的拉普拉斯算子，该算子可以模拟人眼的神经机制。因此，为了简化计算，这里采用 LoG 来计算图像差异值，LoG 算子由下式给出

$$LoG = \frac{1}{\sqrt{2\pi}\sigma^2} \frac{m^2 + n^2 - 2\sigma^2}{\sigma^4} e^{-(m^2+n^2)/(2\sigma^2)} \quad (4-1)$$

其中， σ 是 LoG 的标准偏差， m 、 n 表示各个像素点位置（为了方便书写，这里用 x 表示像素点位置，即 $x = (m, n)$ ，下同）。

4.1.2 多通道 LoG 滤波相似性

在构造多通道 LoG 滤波器之前，首先需要对图像进行预处理，即将 RGB 图像映射到对立的匀色空间。目前，色空间转化的方法有很多，例如 CIELab、LMN^[72]、YIQ^[73]和 YUV^[74]等。由于图像在进行色空间转化的过程中可能会产生误差，从而影响到 LoG 算子对结构信息表达精度。因此，先计算上述色空间在四个公共数据库中的评价标准值，通过比较数值结果来确定理想的匀色空间（即 C1、C2 和 C3）。排在前列的评价标准值已用黑体标出，如表 4-1 所示。

表 4-1 四种颜色空间在不同数据库中的总体性能

数据库		CIELab	YIQ	LMN	YUV
TID2013	SROCC	0.7805	0.7983	0.8123	0.7983
	KROCC	0.5828	0.6060	0.6192	0.6061
	PLCC	0.7899	0.8158	0.8241	0.8159
	RMSE	0.7603	0.7169	0.7023	0.7168
TID2008	SROCC	0.7735	0.8103	0.8138	0.8103
	KROCC	0.5717	0.6123	0.6172	0.6123
	PLCC	0.7713	0.8130	0.8161	0.8130
	RMSE	0.8541	0.7813	0.7755	0.7813
CSIQ	SROCC	0.8585	0.9075	0.9093	0.9076
	KROCC	0.6620	0.7269	0.7298	0.7269
	PLCC	0.8595	0.9069	0.9086	0.9069
	RMSE	0.1342	0.1106	0.1097	0.1106
LIVE	SROCC	0.9462	0.9656	0.9665	0.9656
	KROCC	0.8107	0.8586	0.8612	0.8586
	PLCC	0.9152	0.8772	0.8766	0.8772
	RMSE	12.6080	15.0241	15.0581	15.0244
Direct average	SROCC	0.8397	0.8704	0.8755	0.8705
	KROCC	0.6568	0.7010	0.7069	0.7010
	PLCC	0.8340	0.8532	0.8564	0.8533

从表 4-1 中的数据可以看出，四个均匀色空间中 LMN 整体性能最优，故将 RGB 图像转化为 LMN 均匀颜色空间。其转化方法如式 4-2。

$$\begin{bmatrix} L \\ M \\ N \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4-2)$$

其中 L^* 表示亮度， M^* 和 N^* 代表色度。

然后将由公式 4-1 计算得到的 LoG 算子分别与均匀色空间中的 L 、 M 和 N 三个颜色分量进行卷积，该过程可由等式 4-3 实现。

$$\begin{aligned}
LLoG_i(x) &= L_i(x) * LoG(x) \\
MLoG_i(x) &= M_i(x) * LoG(x) \\
NLoG_i(x) &= N_i(x) * LoG(x)
\end{aligned} \tag{4-3}$$

其中， L 、 M 和 N 分别为色空间中的三个颜色分量， i 表示图像类型（即原始图像 f_1 ，失真图像 f_2 ，下同）。下面对卷积后的三个分量进行融合处理，该过程命名为多通道 LoG 滤波（Multi-channel LoG filtering, MLG）。则 MLG 图可由下式给出：

$$MLG(x) = \left(LLoG_i^2(x) + MLoG_i^2(x) + NLoG_i^2(x) \right)^{\frac{1}{2}} \tag{4-4}$$

则 MLG 相似性可表示为

$$SIM_{LoG}(x) = \frac{2 \cdot MLG_1(x) \cdot MLG_2(x) + T_1}{MLG_1^2(x) + MLG_2^2(x) + T_1} \tag{4-5}$$

其中 T_1 为增加公式稳定性的常数。

为了验证 MLG 的性能，从 TID2008 数据库选择两组规格相同的原始失真图像对，通过使用传统 LoG 滤波和 MLG 方法分别计算该组图像质量评分，并附带计算过程对应的相似图，如图 4-1 所示。

从图 4-1 中可以观察到，图 4-1 (g) 与图 4-1 (h) 中亮点较少，而图 4-1 (e) 和图 4-1 (f) 亮点相对较多，表明本文算法表征图像差异性能强。此外，对评分结果进行分析，其中主观评分值越高，参考和失真图像的差异就越小，质量就越高；客观评分的值越低，参考和失真图像的差异就越小，质量就越高，反之亦然。如图 4-1，图 4-1 (d) 的主观评分值低于图 4-1 (c)，图 4-1 (e) 的客观评分值低于图 4-1 (f)，图 4-1 (g) 的客观评分值高于图 4-1 (h)。这表明，MLG 在预测性能上优于传统的 LoG 滤波，能够更加准确的表征图像结构信息的差异。从图 4-1 (c) 和图 4-1 (d) 的主观评分可以看出 MLG 评价结果与人眼的感知结果保持高度一致。可见 MLG 可更好地表达失真程度。

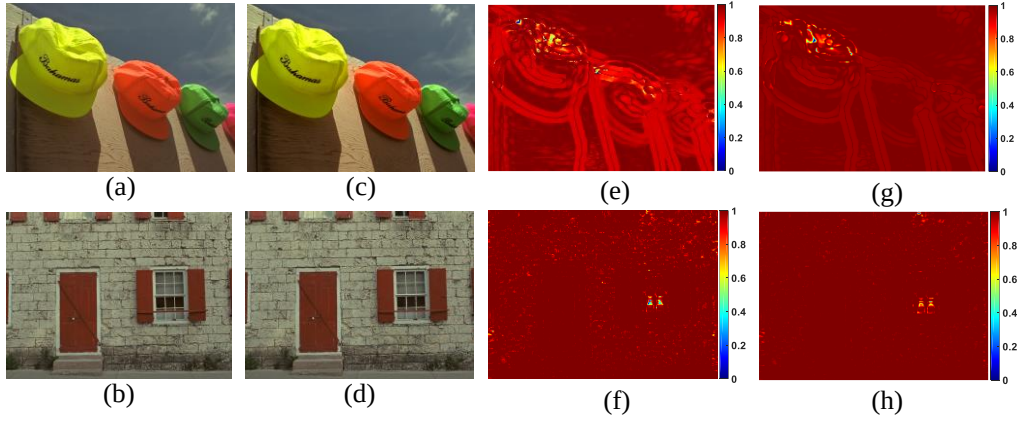


图 4-1 MLG 和传统 LoG 滤波性能对比。(a)和(b)为原始图像。(c)和(d)为失真图像,对应主观评分依次为 7.2 和 4.8286。(e)、(f)为 MLG 相似图,其客观评分依次为 0.008526 和 0.008884。(g)和(h)为传统 LoG 滤波相似图,其客观评分依次为 0.015842 和 0.01499。

4.2 对比度和色度相似图

(1) 对比度相似图

经过 LoG 滤波后的图像对比度是不变的,但图像对比度的变化会影响对图像质量的衡量。故这里引入图像对比度作为本章算法另一特征。对比度计算方法和第 3.3 节一致,全局对比度由公式 4-6 给出:

$$LC_i(x) = \sqrt{\frac{1}{H-1} \sum_{j=1}^H (x_j - \mu_x)^2} \quad (4-6)$$

其中 $\mu_x = \frac{1}{H} \sum_{j=1}^H x_j$ 为图像的平均强度。

根据等式 4-6,对比度相似图的计算方式为:

$$SIM_{LC}(x) = \frac{2LC_1(x) \cdot LC_2(x) + T_2}{LC_1^2(x) + LC_2^2(x) + T_2} \quad (4-7)$$

其中 T_2 是增加等式稳定性的常数, $LC_1(x)$ 和 $LC_2(x)$ 分别为 f_1 和 f_2 上的对比度图。

(2) 色度相似图

图像颜色的变化会对图像的质量产生影响,故引入色度相似性作为客观模型的第三特征,如 FSIM、VSI 和 PerSIM 等。根据表 4-1,将 RGB 图像映射至 LMN

均匀颜色空间，其中 M^* 代表红绿通道， N^* 代表蓝黄通道^[72]。则色度相似图定义为

$$SIM_{MN}(x) = \frac{2M_1(x) \cdot M_2(x) + T_3}{M_1^2(x) + M_2^2(x) + T_3} \cdot \frac{2 \cdot N_1(x) \cdot N_2(x) + T_3}{N_1^2(x) + N_2^2(x) + T_3} \quad (4-8)$$

其中 T_3 是增加等式稳定性的常数， M_1 、 M_2 和 N_1 、 N_2 分别为 f_1 和 f_2 上的色度通道。

4.4 基于多通道融合的全参考图像质量评价算法

通过融合前文提到的三个分量，最终可以计算出图像质量分数。所提模型命名为基于多通道融合的全参考图像质量评价（Full-Reference Image Quality Assessment Based on Multi-Channel Visual Information Fusion, MVFC）算法。该指数可由公式 4-9 给出。

$$S = w_1 \cdot SD(SIM_{LoG}) + w_2 \cdot SD(\sum_{\Omega} SIM_{MN}) + w_3 \cdot SD(\sum_{\Omega} SIM_{LC}) \quad (4-9)$$

其中， w_1 、 w_2 和 w_3 是表示重要性的权重，且 $w_1 + w_2 + w_3 = 1$ 。 SD 计算公式为：

$$SD(k) = \sqrt{\frac{1}{H} \sum_{k=1}^H (k - \bar{k})^2}, \quad \bar{k} = \frac{1}{H} \sum_{k=1}^H k, \quad \text{其中 } k \text{ 表示图像的特征相似分数。}$$

为了更加直观的理解本章图像评价方法的构造流程，在图 4-2 中给出了流程图。如图 4-2 所示，原始图像 f_1 和失真图像 f_2 在经过滤波处理后获得了 MLG 相似图，再通过将 MLG 相似图、对比度相似图和色度相似图在空间中汇集池化，最终得到所提相似分数。

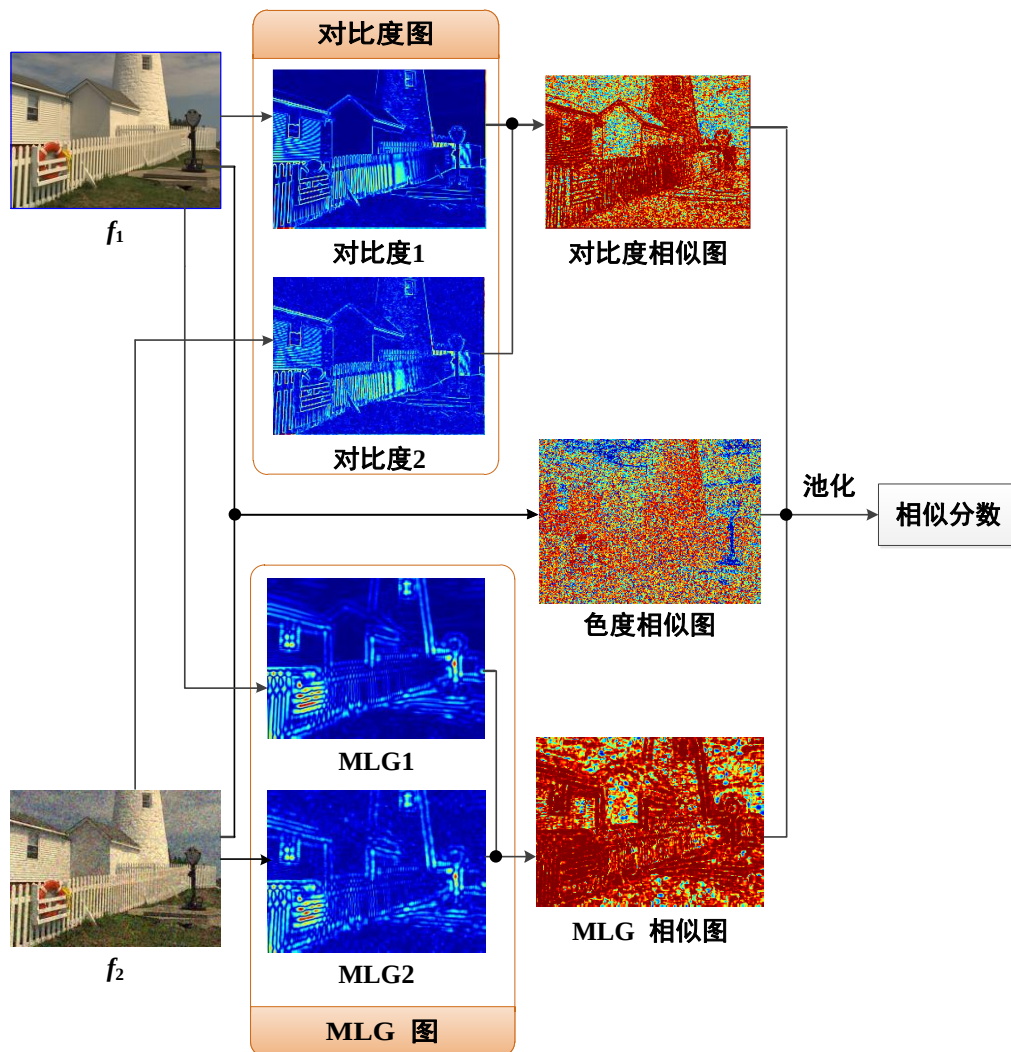


图 4-2 本章算法流程图

4.5 实验结果与分析

4.5.1 数据库和模型评价标准

本章选用四个公共图像数据库作为比较不同评价算法性能的图像集，其分别为 TID2008，CSIQ，LIVE 和 TID2013 数据库。以上四个数据库在前文中已详细介绍，为避免文章冗长，这里不再过多的描述（下同）。

本章采用 SROCC、KROCC、PLCC 和 RMSE 这四个性能指标来判别不同评价算法的性能好坏。评价方法的结果越准确对应 SROCC、KROCC 和 PLCC 值越大；RMSE 值越小。此外，在计算性能指标之前，需要通过 logistic 回归分析

方法将原始客观分数映射为与主观分数同一量级。

4.5.2 算法参数优化

前文所提的计算公式中，涉及的参数变量有 T_1 、 T_2 、 w_4 、 w_5 和 w_6 。为了使模型的性能达到最优，需要对变量值进行修正。这里选择采用控制变量的方法寻找最优解，并依据对应的评价标准系数 PLCC 值确定最终目标值。

首先，保持 T_1 、 w_4 、 w_5 和 w_6 值不变，在 $[20,70]$ 区间内调整 T_2 的取值范围。根据图 4-3 显示的结果，在四个数据库中的 PLCC 值达到最佳时， T_2 处于区间 $[30,60]$ 内。经过进一步细化选择后，最佳参数设置为 $T_2=50$ 。同样的方法，保持其它变量不变。将 T_1 设置在区间 $[0.00003, 0.00008]$ 内，当 T_1 值在 $[0.00004, 0.00006]$ 时，所提模型在各数据库中的 PLCC 值最理想，选定 $T_1=0.00006$ 。

权重系数 w_4 、 w_5 和 w_6 的修正方式也遵循上述逻辑。考虑到这三个变量之间满足关系 $w_4+w_5+w_6=1$ 。因此，只需确定其中两个变量即可推导出第三个变量。图 3-6 展示了在固定 w_6 值时通过调整 w_4 值取得的结果。 w_4 值初始取值范围为 $[0.44, 0.54]$ 。从图中可知，当 w_4 处于 $[0.49, 0.51]$ 区间内时，四个数据库中 PLCC 值达到最理想状态，故选取 $w_4=0.5$ 。同样保持 w_4 不变，将 w_6 设置在区间 $[0.41, 0.46]$ 内。当 w_6 设置在区间 $[0.43, 0.45]$ 时，所提模型 PLCC 值达到最佳状态，故选取 $w_6=0.45$ 。

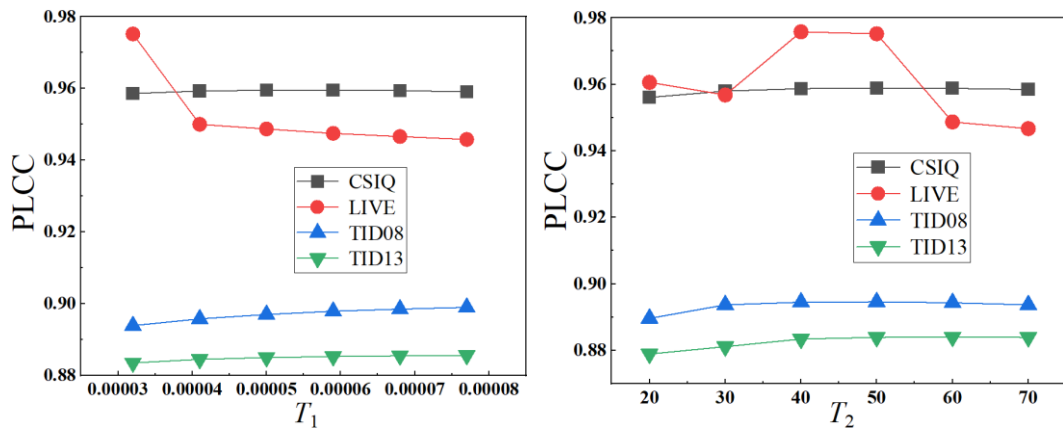


图 4-3 不同 T_1 和 T_2 值在四个基准数据库上的 PLCC 值

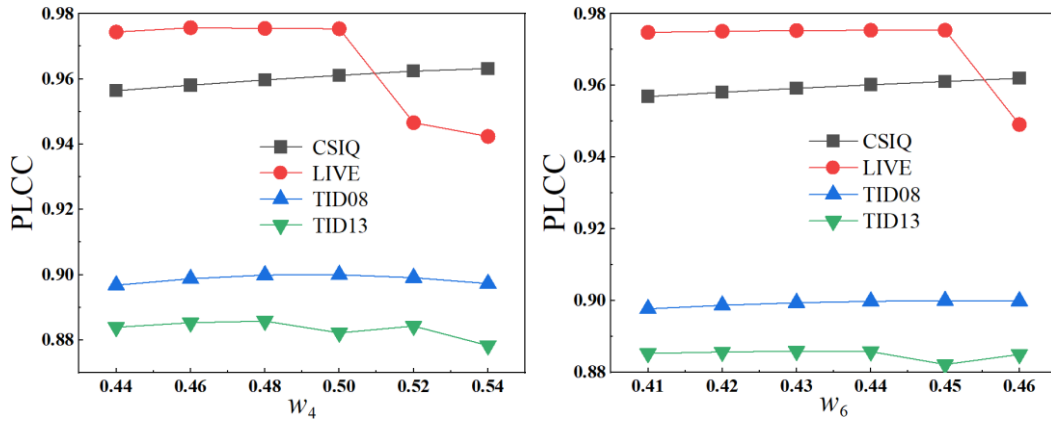


图 4-4 不同 w_4 和 w_6 值在四个基准数据库上的 PLCC 值

4.5.3 算法性能比较

(1) 整体性能比较

为了验证本文模型在图像质量评估方面的性能, 选用 SSIM、CVSS、GSM、VIF、FSIMc、VSI、CAGS、RVSIM、MAD 和 GMSD 十个计算模型与本文模型作比较。这些模型在图像质量评估领域具有一定的代表性和影响力, 通过与其进行比较, 可以更全面地评估本文模型的性能。在实验中, 分别提取了图像的 MLG、对比度和色度三个特征信息, 作为本模型特征。这些特征信息涵盖了图像结构、对比度和色彩等方面的信息, 能够更好的量化图像质量。如表 4-3 所示, 评价结果排列前三的数据均已加粗标出。为了更全面地比较不同模型的性能, 这里对四个数据库中的性能评价标准值进行了加权和直接平均处理。

由表 4-3 可见, 本文模型在四个数据库中都有很好的表现, 整体性能位列第一。与同样采用 LoG 滤波提取图像结构信息的 FSIMc 和 VSI 相比, 本文模型在 LIVE 和 CSIQ 数据库中表现优势更加明显。在前三个指标的加权平均和直接平均值中, 所提模型的性能位列第一。此外, 在 TID2008 数据库中, 所提模型的 SROCC 值为 0.9048, 在 11 个计算模型中位列第一。综上所述, 本文模型具有很好的竞争力和预测精度。

表 4-3 不同的 IQA 模型在四个数据库中的性能比较

Database Criteria	SSIM	VIF	MAD	FSIMc	GMSD	VSI	GSM	RVISM	CVSS	CAGS	Proposed
SROCC	0.7417	0.6769	0.7807	0.8510	0.8044	0.8965	0.7946	0.6757	0.8069	0.8316	0.8606
TID2013 KROCC	0.5588	0.5147	0.6035	0.6665	0.6339	0.7183	0.6255	0.5146	0.6331	0.6469	0.6810
PLCC	0.7895	0.7720	0.8267	0.8769	0.8590	0.9000	0.8464	0.7825	0.8406	0.8445	0.8822

	RMSE	0.7608	0.7880	0.6975	0.5959	0.6346	0.5404	0.6603	0.7719	0.6715	0.6639	0.5838
TID2008	SROCC	0.7749	0.7491	0.8340	0.8840	0.8907	0.8979	0.8504	0.7375	0.9001	0.8231	0.9048
	KROCC	0.5768	0.5860	0.6445	0.6991	0.7092	0.7123	0.6596	0.5628	0.7215	0.6289	0.7278
	PLCC	0.7732	0.8084	0.8308	0.8762	0.8788	0.8762	0.8422	0.7954	0.8961	0.8091	0.9000
	RMSE	0.8511	0.7899	0.7468	0.6468	0.6404	0.6466	0.7235	0.8133	0.5956	0.7886	0.5850
	SROCC	0.8756	0.9195	0.9466	0.9310	0.9570	0.9423	0.9108	0.8979	0.9580	0.9198	0.9570
CSIQ	KROCC	0.6907	0.7537	0.7970	0.7690	0.764	0.7857	0.7374	0.7234	0.8173	0.7487	0.8127
	PLCC	0.8613	0.9277	0.9502	0.9192	0.930	0.9279	0.8964	0.9236	0.9589	0.9014	0.9611
	RMSE	0.1334	0.0980	0.0818	0.1034	0.0786	0.0979	0.1164	0.1007	0.0745	0.1137	0.0725
	SROCC	0.9479	0.9636	0.9669	0.9645	0.9603	0.9524	0.9561	0.9600	0.9672	0.9734	0.9798
LIVE	KROCC	0.7963	0.8282	0.8421	0.8363	0.8268	0.8058	0.8150	0.8203	0.8406	0.8658	0.8863
	PLCC	0.9449	0.9604	0.9675	0.9613	0.9595	0.9482	0.9512	0.9570	0.9651	0.9640	0.9754
	RMSE	8.9455	7.6137	6.9073	7.5296	7.6937	8.6816	8.4327	7.9274	7.1573	8.3251	6.9051
	SROCC	0.7942	0.7646	0.8405	0.8847	0.8544	0.9100	0.8452	0.7638	0.8721	0.8588	0.9027
Weighted Average	KROCC	0.6108	0.6049	0.6702	0.7101	0.7022	0.7366	0.6732	0.6006	0.7074	0.6828	0.7414
	PLCC	0.8140	0.8261	0.8619	0.8928	0.8892	0.9033	0.8650	0.8307	0.8869	0.8575	0.9112
	SROCC	0.8350	0.8273	0.8821	0.9076	0.9031	0.9223	0.8780	0.8178	0.9081	0.8870	0.9256
Direct Average	KROCC	0.6557	0.6707	0.7218	0.7427	0.7457	0.7555	0.7094	0.6553	0.7531	0.7226	0.7770
	PLCC	0.8422	0.8671	0.8938	0.9084	0.9127	0.9131	0.8841	0.8646	0.9152	0.8798	0.9297
	SROCC	0.8350	0.8273	0.8821	0.9076	0.9031	0.9223	0.8780	0.8178	0.9081	0.8870	0.9256

为了进一步体现所提模型在图像质量评估方面的性能优势,选用 TID2013 公共数据库提供的离散点来比较不同模型的拟合性能。TID2013 数据库是一个广泛使用的图像质量评估数据库,其包含了多种类型的失真图像以及对应的主观观测值。如图 4-5 所示,通过对比模型间回归曲线与主观观测值的相关性,可以直观地观察到模型的拟合性能。

图 4-5 中,所提模型的回归曲线与主观观测值具有更好的相关性。这一结果表明,本文模型所提取的图像特征,能够更准确地捕捉到图像质量的变化,并与主观观测值保持较高的一致。这基于本文模型融合了 HVS 特性,同时提取了多通道 LoG 滤波、对比度信息和色度信息。具体而言,多通道 LoG 滤波的提取方法充分考虑了亮度和色度的协调变化,以及色度分量对图像结构的影响。此外,通过偏差和融合多个图像特征,让本文模型能够更全面地评估图像质量,避免单一特征带来的偏差。综上所述,所提模型在 TID2013 数据库上的拟合性能优于其它模型,进一步验证了其可靠性。

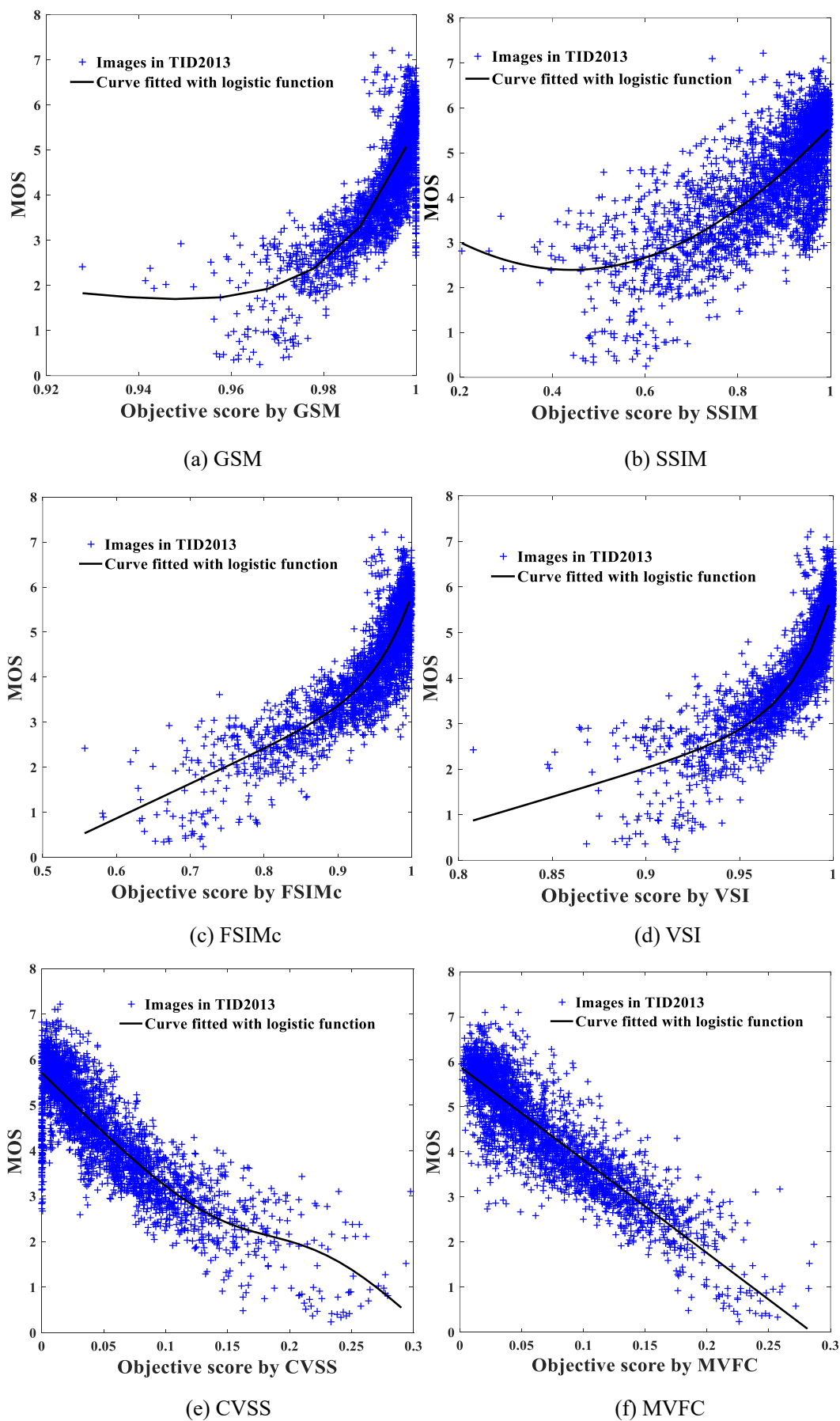


图 4-5 TID2013 数据库上的主观 MOS 和模型预测得分的散点图

(2) 单个失真类型的性能比较

分析模型在不同失真类型下的预测性能是评估图像质量算法稳定性和泛化性的重要步骤。表 4-4 展示了各个模型在不同失真类型下的 PLCC 值，这一指标能够衡量模型预测结果与主观评分之间的线性相关性，从而反映模型的预测性能。在每个失真类型中，都用黑体标记了 PLCC 值排名前三的模型，以便更直观地比较不同模型的性能。从表 4-4 中可以看出，本文模型的黑体标记共计出现 21 次位列第一，显示出较高的稳定性和泛化能力。VSI 模型紧随其后，共标记 16 次位列第二，而 CVSS 模型则位列第三，共标记 12 次。这一结果进一步证明了本文模型在图像质量评估领域的优势。不同失真类型下的良好预测性能表明，本文模型能够有效地提取图像的特征信息，并将其与主观评分进行准确的映射。同时，本文模型在多个失真类型中均能获得较高的 PLCC 值，如 TID2013 数据库中的加性高斯噪声 (Additive Gaussian noise)、CSIQ 数据库中的全局对比度衰减 (Global contrast decrement) 和 LIVE 数据库中的 JPEG 压缩 (JPEG compression) 对应 PLCC 值依次为 0.9537、0.9478 和 0.9912。综上所述，本文模型在不同失真类型下的预测性能均表现出色，具有较高的稳定性和泛化能力。

对比表 4-4 中数据，本文模型在 LIVE 和 CSIQ 数据库中具有很好的表现，其 PLCC 值最高为 0.9912，明显优于那些同样使用了 LoG 的计算模型 (VSI 和 FSIMc)。进一步分析本文模型、VSI 和 CVSS 在表 4-4 中的数据，本文模型 PLCC 值较低的是 CTC 和 Block 两种失真类型，其值分别为 0.5466 和 0.6253；VSI 模型 PLCC 值较低的是 Block 和 CTC 两种失真类型，其结果分别是 0.4875 和 0.6595；CVSS 模型 PLCC 值较低的是 Block、CTC 和 CCS 三种失真类型，其结果依次为 0.4983、0.4278 和 0.0893。以上数据可知，面对难以处理的 CTC 和 Block 两种失真类型，本文模型的表现优于 CVSS。特别地，在评价 Block 失真类型时，本文模型 PLCC 值远优于 VSI 和 CVSS。综上所述，本文模型在每个数据库上都具有可靠的稳定性和泛化性。

表 4-4 IQA 模型在不同失真类型下的 PLCC 值

数据库		SSIM	VIF	MAD	FSIM	GMSD	VSI	GSM	RVISM	CVSS	CAGS	MVFC
TID2013	AGN	0.8599	0.8555	0.8846	0.9151	0.9494	0.9515	0.9129	0.8674	0.9454	0.9395	0.9537
	ANC	0.7811	0.8169	0.8302	0.8869	0.9133	0.9170	0.8576	0.8331	0.9048	0.8835	0.9138
	SCN	0.8502	0.8680	0.8835	0.8240	0.9389	0.9441	0.9257	0.8474	0.9206	0.8637	0.9470
	MN	0.8231	0.8627	0.7858	0.8478	0.7514	0.811	0.7834	0.8600	0.8039	0.7975	0.8231
	HFN	0.9037	0.9083	0.9223	0.9432	0.9548	0.9639	0.929	0.9161	0.948	0.876	0.8358
	IN	0.7649	0.8346	0.331	0.7216	0.7573	0.8637	0.8168	0.8804	0.7426	0.8344	0.7971
	QN	0.7693	0.8515	0.8365	0.8096	0.9109	0.8726	0.8659	0.7740	0.8944	0.8763	0.9072
	GB	0.9662	0.9602	0.9336	0.9549	0.9094	0.9564	0.9629	0.9733	0.9292	0.9592	0.9028
	DEN	0.9609	0.8892	0.9611	0.9647	0.9756	0.9704	0.9759	0.9343	0.9649	0.9743	0.9802
	JPEG	0.9493	0.9198	0.9612	0.973	0.9834	0.9854	0.9765	0.9626	0.9753	0.9831	0.9846
	JP2K	0.9681	0.9366	0.9738	0.9741	0.9812	0.9831	0.9801	0.9689	0.9699	0.9734	0.9787
	JGTE	0.8859	0.8922	0.8955	0.9178	0.8786	0.9454	0.9217	0.9024	0.8856	0.9098	0.8980
	J2TE	0.8725	0.8650	0.8755	0.8010	0.9087	0.9194	0.8994	0.8835	0.8884	0.8315	0.8884
	NEPN	0.7755	0.7989	0.8552	0.8057	0.8162	0.8154	0.8063	0.7368	0.8405	0.7675	0.8154
	Block	0.5622	0.5620	0.4216	0.5488	0.6351	0.4875	0.6061	0.6393	0.4983	0.6636	0.6253
	MS	0.7849	0.7081	0.6726	0.7870	0.7661	0.8025	0.7721	0.6911	0.7537	0.7577	0.7176
	CTC	0.4691	0.8362	0.2970	0.6838	0.5098	0.6595	0.6470	0.2129	0.4278	0.6346	0.5466
	CCS	0.4998	0.3541	0.1601	0.8377	0.1516	0.8106	0.7384	0.4374	0.0893	0.408	0.7324
	MGN	0.7768	0.7908	0.8528	0.8654	0.891	0.913	0.8430	0.8291	0.8825	0.8768	0.8961
	CN	0.9043	0.8866	0.9343	0.9466	0.9562	0.9544	0.9492	0.9333	0.9480	0.9390	0.9643
	LCNI	0.9141	0.8761	0.9561	0.9564	0.9699	0.9634	0.9570	0.9256	0.9671	0.9639	0.9783
	ICQD	0.7976	0.8147	0.8694	0.8053	0.9192	0.8963	0.900	0.8169	0.9124	0.9176	0.9346
	CHA	0.9599	0.9381	0.9591	0.9798	0.9709	0.9752	0.9683	0.9764	0.9697	0.9713	0.9580
	SSR	0.9690	0.9161	0.9762	0.9766	0.9848	0.9800	0.9840	0.9657	0.9785	0.9739	0.9847
CSIQ	AGWN	0.9148	0.9563	0.9514	0.9383	0.4556	0.9636	0.9505	0.9389	0.9671	0.9646	0.9601
	JPEG	0.9774	0.9833	0.9827	0.9848	0.7621	0.9808	0.9804	0.9823	0.9865	0.9791	0.9860
	JP2K	0.9698	0.9782	0.9842	0.9818	0.7883	0.9745	0.9696	0.9799	0.9856	0.9616	0.9848
	AGPN	0.8916	0.9612	0.9548	0.8172	0.5573	0.8698	0.8499	0.9194	0.9581	0.8514	0.9553
	GB	0.9363	0.9562	0.9713	0.8933	0.6558	0.8761	0.8573	0.9737	0.9815	0.8394	0.9811
	CTC	0.7867	0.9233	0.9306	0.8914	0.8775	0.8686	0.8595	0.8581	0.9449	0.8672	0.9478
LIVE	JP2K	0.8848	0.9392	0.9754	0.9794	0.9771	0.8662	0.8564	0.9619	0.978	0.9776	0.9853
	JPEG	0.9761	0.9817	0.9806	0.9849	0.9832	0.9816	0.984	0.9853	0.9523	0.9861	0.9912
	AWGN	0.9765	0.9907	0.9881	0.9772	0.9655	0.9819	0.8904	0.9809	0.982	0.9857	0.9863
	GB	0.9237	0.9626	0.9415	0.9737	0.9618	0.8544	0.8565	0.9682	0.9709	0.9572	0.9751
	FF	0.8510	0.9527	0.952	0.8515	0.9408	0.8151	0.7925	0.9623	0.9585	0.9541	0.9625

为了进一步验证所提模型在单个失真类型中的评价性能,添加了五篇近期发表的期刊与之比较,即 DISIS^[75]、SG-ESSIM^[76]、VSPSI^[77]、LGV^[78]和 SWLGV^[78]。这里选择包含失真类型最多的 TID2013 数据库作为模型性能测试库,并设置 SROCC 为性能评价标准值。结果如表 4-5 所示,其中 SROCC 值排名前三的模型,已用黑体标出。

在表 4-5 中，黑体标记出现最多的模型依次为 VSPSI（20 次）、本文模型（17 次）和 SG-ESSIM（14 次）。通过上述数据，可以看出本文模型具有较高的预测性能。进一步分析表中数据，所提模型在所有失真类型（24 种）中的平均 SROCC 值为 0.8531，与模型 VSPSI（平均 SROCC 值为 0.8532）和 SWLGV（平均 SROCC 值为 0.8552）非常接近，且明显高于剩下三个计算模型。上述结果表明本文模型具有较好的泛化性和良好的竞争力。

表 4-5 IQA 模型在不同失真类型下的 SROCC 值

Database	Type	DISTS	SG- ESSIM	VSPSI	LGV	SWLGV	MVFC
TID2013	AGN	0.8450	0.9360	0.9471	0.9210	0.9360	0.9500
	ANC	0.7860	0.8550	0.8722	0.9110	0.9040	0.8859
	SCN	0.8590	0.9350	0.9385	0.8870	0.9300	0.9438
	MN	0.8140	0.7150	0.7712	0.8210	0.8420	0.7480
	HFN	0.8680	0.9200	0.9218	0.9160	0.8740	0.9211
	IN	0.6740	0.8330	0.8758	0.7590	0.7950	0.8082
	QN	0.8100	0.9110	0.8765	0.8660	0.9560	0.9063
	GB	0.9260	0.9690	0.9631	0.9520	0.9610	0.9074
	DEN	0.8990	0.9630	0.9502	0.9840	0.9760	0.9592
	JPEG	0.8970	0.9500	0.9650	0.9700	0.9520	0.9532
	JP2K	0.9310	0.9490	0.9725	0.9450	0.9680	0.9625
	JGTE	0.9060	0.8230	0.9234	0.8950	0.9000	0.8804
	J2TE	0.8650	0.8990	0.9246	0.9180	0.8740	0.8960
	NEPN	0.8330	0.8010	0.8076	0.7190	0.7950	0.8140
	Block	0.3020	0.6230	0.1716	0.6030	0.6010	0.6427
	MS	0.7520	0.7060	0.7715	0.6770	0.7560	0.6566
	CTC	0.4640	0.4520	0.4763	0.6590	0.6670	0.3666
	CCS	0.7890	0.0100	0.8116	0.7500	0.7580	0.7350
	MGN	0.7900	0.9000	0.9135	0.8190	0.8410	0.8903
	CN	0.9070	0.9160	0.9231	0.8380	0.8590	0.9356
	LCNI	0.9320	0.9520	0.9583	0.8730	0.9170	0.9754
	ICQD	0.8320	0.9280	0.8856	0.8450	0.8640	0.9261
	CHA	0.8790	0.8350	0.8923	0.7930	0.7880	0.8396
	SSR	0.9440	0.9640	0.9632	0.8000	0.8100	0.9695
	Average	0.8127	0.8227	0.8532	0.8384	0.8552	0.8531

（3）算法的计算效率比较

为了更全面地评估本文模型在实际应用中的性能，这里选用七个评价模型与其进行计算复杂度比较。实验选用的待测图像、实验设备和仿真平台与 3.5.3 小节中的一致，不再赘述。

结果如表 4-6 所示，本文模型运行时间仅需 0.14s，具有较低的计算复杂度。这表明，在实际应用中，本算法可以快速完成相关任务。

表 4-6 不同 FR-IQA 模型的计算效率对比

IQA Index	Time cost /s	IQA Index	Time cost/s
SSIM	0.126	VSI	0.536
GSM	0.271	RVSIM	0.939
CVSS	0.436	MAD	1.256
CAGS	0.483	Proposed	0.148

4.6 本章小结

论文从图像质量评价算法的高精度、高效率出发,提出了基于多通道视觉信息融合的全参考模型(MVFC 模型)。本文通过 LoG 滤波来分解图像信息,模拟视觉神经细胞的多通道特性。此外,融合了多通道 LoG、图像对比度和色度三种特征分量,并通过控制变量法修正算法参数,最终得到本文的计算模型。实验阶段,选用了四个常用数据库来验证图像质量评价算法的性能,并将所提算法与多个经典的 FR-IQA 算法分析比较。结果表明,本章方法的泛化性能和预测性能优于大多数 FR-IQA 模型;此外,考虑到算法的实用性,这里选用了七种模型与本文模型进行计算复杂度比较。综上,本文所提的全参考评价方法具有较可靠的性能,适用于对图像质量预测精度和预测效率有要求的图像评价场景。

第 5 章 基于 PSO-GRNN 池化策略的全参考图像质量评价

在前一章中，论文充分利用了颜色分量中的结构信息，并完成了对多个特征的融合池化。但面对池化方法的选择问题，并未做详细的解释。而图像特征池化方法的研究，对图像质量评价有着重要意义。传统池化方法通常只适用于特定的特征组合，且池化结果并不理想。而机器学习通过训练可以准确的寻找到图像特征分数间的非线性关系，基于此论文提出了基于 PSO-GRNN 池化策略的全参考图像质量评价方法。

本章首先介绍了 PSO-GRNN 池化方法的理论框架。接着描述了 PSO-GRNN 池化方法的实现过程。然后在 5.3 节中，完成了方法的性能验证。最后对本章的研究内容和结果进行了总结。

5.1 PSO-GRNN 池化方法

5.1.1 GRNN 网络结构

GRNN 网络结构包含输入层、模式层、求和层和输出层^[38]。其中，输入层负责将输入数据传递到模式层。模式层中采用高斯函数对输入数据进行处理，节点数目等于训练样本数目。求和层包含两个节点，分别为上层输出的加权和以及算术和。求和层的两个节点分别对应式 5-1 中的分子和分母。输出节点计算求和层的两个输出的商，最终得到预测值，如图 5-1 所示。

对于输入的向量 X ，GRNN 的输出 Y 为^[79]

$$\hat{Y}(X) = \frac{\sum_{i=1}^n Y_i \exp(-D_i^2 / 2\sigma^2)}{\sum_{i=1}^n \exp(-D_i^2 / 2\sigma^2)} \quad (5-1)$$

其中， n 为样本观测数量， $D_i^2 = (X - X_i)^T (X - X_i)$ ， X_i 和 Y_i 是样本值， σ 模式层中高斯函数的宽度系数。 σ 值越大，函数拟合曲线越平滑。

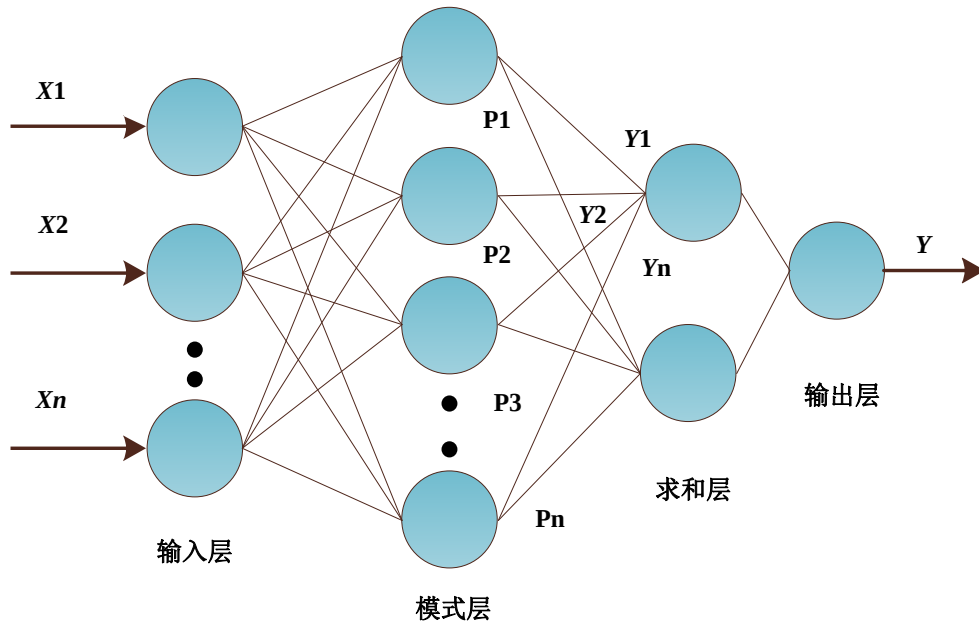


图 5-1 GRNN 回归网络

5.1.2 PSO 方法原理

粒子群优化（Particle Swarm Optimization, PSO）将待解问题的搜索空间视为多维空间中的一个区域，并将潜在解称为粒子^[80]。这些粒子通过信息共享来寻找问题的最优解。每个粒子都有位置和速度，位置表示潜在解，速度决定了粒子如何移动以寻找更好的解。粒子的移动受到个体最优解（Pbest）和全局最优解^[81]（Gbest）两个因素影响。个体最优解是指代粒子自身找到的最佳解，而全局最优解则是指整个群体当前所找到的最优解。粒子通过两个最优解的引导，来调整自己的速度和位置以逐渐接近更优秀的解决方案。PSO 算法具有简单和高效等优点。此外，PSO 算法还具有很好的并行性和扩展性，这使得它在面临大规模优化问题时能够稳定且准确的计算出理想解。

PSO 算法大体分为四个步骤^[82]。首先初始化粒子群的位置和速度。其次计算每个粒子的适应度值。然后更新每个粒子的速度和位置。最后，判断迭代终止条件。下面简要介绍 PSO 算法中两个关键组成部分：适应度函数和粒子位置与速度更新函数。

（1）适应度函数

适应度值是用来判断粒子当前位置的好坏程度，确保粒子朝着最优解的方向。

粒子的适应度函数由下式给出^[83]:

$$I_i = \sum_j (Y_{ij} - y_{ij})^2 \quad (5-2)$$

$$I_{popIndex} = \frac{1}{n} \sum_{i=1}^n I_i \quad (5-3)$$

式中, n 为样本个数; Y_{ij} 和 y_{ij} 分别为第 i 个样本的第 j 个主观评分输出值和客观评分输出值; $popIndex=1, 2, \dots, popSize$, 其中 $popSize$ 为粒子个数。

(2) 粒子位置和速度更新函数

依据粒子适应度值大小, 可确定 $Pbest$ 和 $Gbest$ 。粒子的速度更新公式则由下式给出^[81]:

$$v_i = w \cdot v_i + c_1 \cdot r_1 \cdot (pbest_i - x_i) + c_2 \cdot r_2 \cdot (gbest_i - x_i) \quad (5-4)$$

$$x_i = x_i + v_i \quad (5-5)$$

其中, w 为惯性权重。 v_i 是粒子速度, x_i 是粒子当前位置, r_1 和 r_2 是一个产生 $[0,1]$ 数值的随机数函数^[82]。 c_1 和 c_2 是非负的学习因子, 用来调节粒子移动步长, 通常 $c_1=c_2=2$ 。

5.1.2 基于 PSO 优化的 GRNN 网络

这里 PSO 优化算法主要是优化径向基函数的扩展速度 σ 值^[84], 如图 5-2 所示。首先将图像特征相似分数输入到 GRNN 网络中, 产生了一组预 σ 测值和 GRNN 响应面; 将预测值和对应的实测值输入到适应度函数中, 记忆最接近实测值的预测值; 再将最优预测值对应的 σ 值用到粒子位置和速度更新函数中, 得到了一组新的 σ 值; 更新 GRNN 网络中 σ 值, 训练生成一组新的预测值和响应面。随着 σ 值不断更新, 最终找到全局最优解。

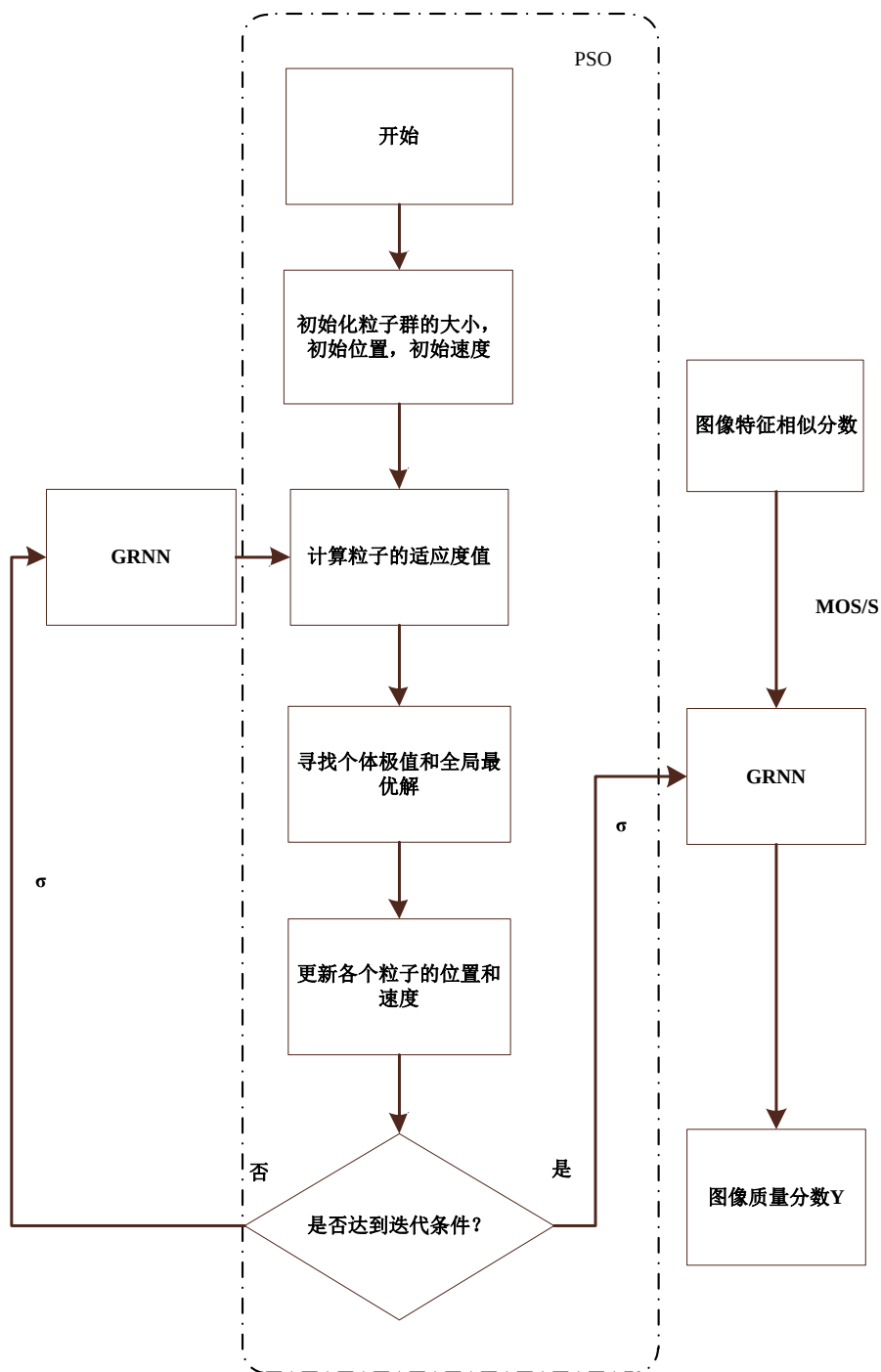


图 5-2 PSO-GRNN 回归网络

5.1.3 图像预处理和 PSO-GRNN 池化策略

(1) 图像预处理

图像特征一般是通过滤波、卷积等方法获得，例如视觉显著性、梯度和对比度特征等。由该方法提取的图像特征图 PF_n 和特征相似图 FS_n 均为二维信号，并

不能作为 PSO-GRNN 的输入值。因此，这里需要将 FS_i 进行预处理。不同的预处理方法会对网络训练的精度造成影响，为了选择合适的预处理方法，这里分别用 MEAN^[85]、STD^[17]、mad^[86]、基于 MEAN 的 PSO-GRNN 池化、基于 STD 的 PSO-GRNN 池化和基于 mad 的 PSO-GRNN 池化方法计算图像对比度相似分数，并比较不同方法对应的 SROCC 值大小。如图 5-3 所示。

由图 5-3 可见，基于 mad 的 PSO-GRNN 池化方法在所有数据库中是表现最好的，与人眼感知的结果也是最相符的。因此，本文选择 mad 计算图像特征相似分数。

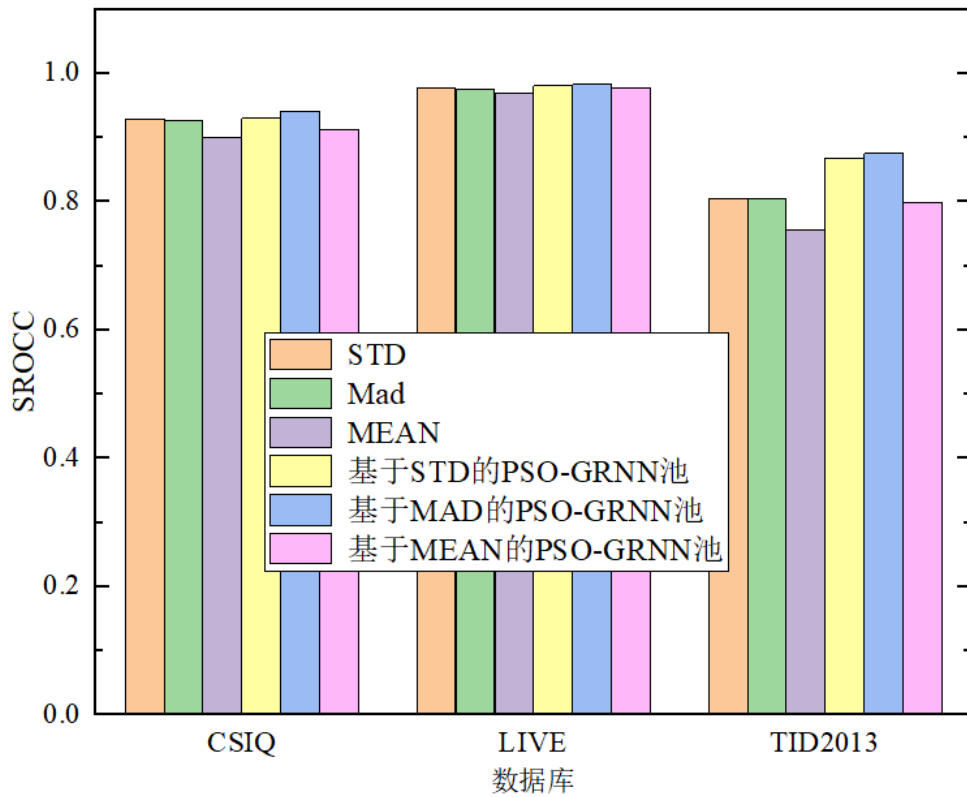


图 5-3 不同池化方法的性能比较

图 5-4 描述了图像预处理的流程图，首先从原始失真图像对中获得相应的特征图，然后计算原始失真图像间的特征相似度，最后通过 mad 数学方法处理 FS_n ，最终得到特征相似分数 S_n 。

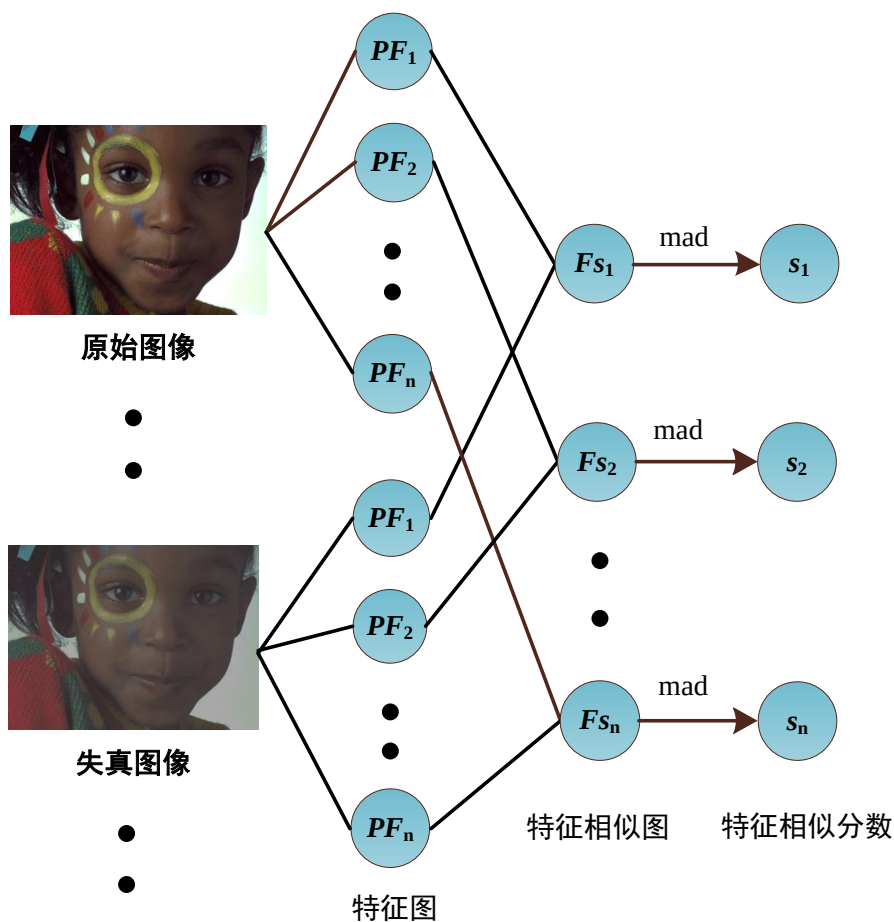


图 5-4 图像特征相似分数

(2) PSO-GRNN 池化策略

根据 5.1.2 节中的内容，这里对 PSO-GRNN 网络进行训练，然后将预处理后得到的多个图像特征相似分数输入到已经训练好的 PSO-GRNN 网络中，最终输的结果即为图像质量分数。如图 5-5 所示。

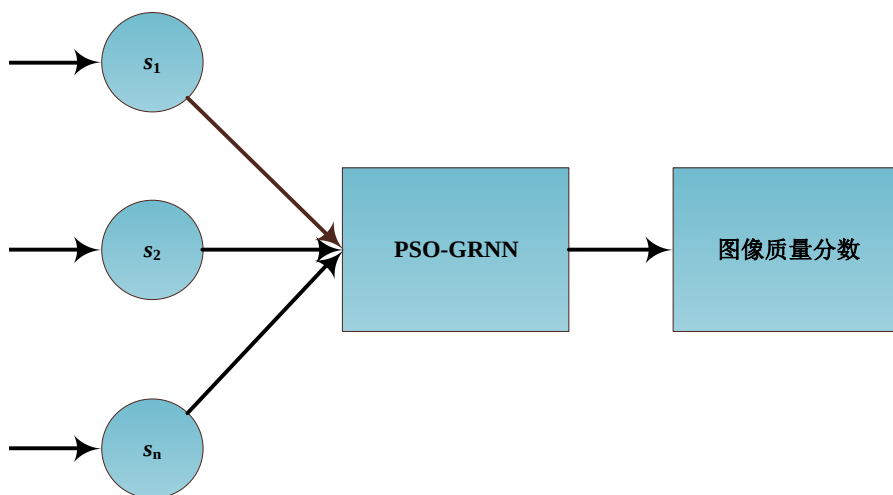


图 5-5 PSO-GRNN 池化简图

5.2 实验结果与分析

5.2.1 数据库和模型评价标准

这里选用 SROCC、KROCC、PLCC 和 RMSE 这四个性能指标衡量本文提出基于 PSO-GRNN 池化策略的全参考图像质量方法。前三个指标越接近于 1，代表图像评价方法预测的结果与人眼的感知结果愈加一致。RMSE 表示主观评价和客观评价的差异，RMSE 越小，评价算法的结果越准确。本章选择在 TID2008、CSIQ、LIVE 和 TID2013 图像公共数据库上比较不同算法的预测性能，其中 LIVE、CSIQ 和 TID2013 这三个数据库主要用作网络的训练和验证，而 TID2008 数据库是用来验证在上述三个数据库中已经训练完成网络的性能。以上四个数据库在前文中已详细介绍，为避免文章冗长，这里不再过多的描述（下同）。表 3-1 列举了这四个数据库的基本信息。

5.2.2 算法参数设置

根据失真图像按照 4:1 比例随机分为训练集与测试集，并将其对应的主观分数随机分为训练集与测试集。由于不同数据集的主观分数并不在一个量级，这里将每个数据集的主观分数缩放到[0-1]范围内。本章提出的 PSO 主要是优化 GRNN 网络中的变量 σ 值，而 PSO 的维度是和变量数目保持一致的。因此，这里将 PSO 的维度设置为 1。权重设为随着迭代次数递减的函数。学习因子默认为 1.5。每个粒子的位置和速度设置在 0.001-1 之间。

由前文可知，本章算法只有一个变量 σ 值，为了选择符合人眼感知结果的 σ 值，将 PSO-GRNN 池在三个数据库中训练 10 次，迭代次数设置为 500，选择最高 SROCC 值所对应的点作为 σ 值修正结果，如图 5-6 所示。从图 5-6 可以看出，PSO-GRNN 池在 LIVE 数据库中的第八次训练结果最为理想，其对应的 σ 值是 0.005108，这里选择 σ 值为 0.005108。在 CSIQ 数据库中，第一次训练的结果最为理想，其对应的 σ 值是 0.006456，本章选择 CSIQ 数据库中的 σ 值为 0.006456。TID2013 数据库中 SROCC 值的最佳结果是在第九次训练，其对应的 σ 值为 0.0035，这里选择 0.0035 作为 PSO-GRNN 网络在 TID2013 数据库中的最佳 σ

值。

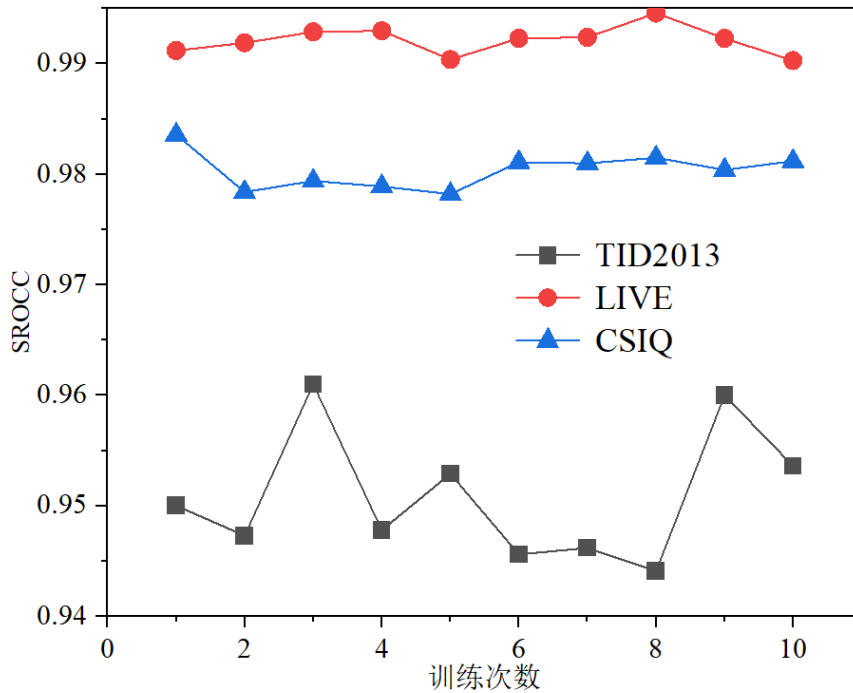


图 5-6 不同训练次数下 σ 值对应的 SROCC 值

5.2.3 算法性能比较

(1) 整体性能比较

为了证明所提方法的性能优势，将本文 PSO-GRNN 池化方法分别用到三个经典的图像质量评价算法 SSIM, FSIMc 和 VSI 以及两个近年发表的 CAGS 和 MVFC 算法中，并将得到的结果与原算法进行比较。如表 5-3 所示。在表格中给出了评价标准值的加权平均和直接平均，其中权值大小是根据不同数据库中的失真图像数量所确定。考虑到 TID2008 数据库与 TID2013 数据库中的失真类型存在重叠，为避免冗余，故仅选择两者中的 TID2013 数据库进行性能分析。此外，后续会使用 TID2008 数据库中所有的图像标签来验证 PSO-GRNN 池化的可靠性。

在表 5-3 中，由左向右，后五个算法是对前五个算法的改进，即通过使用 PSO-GRNN 池来代替原池化方法融合图像的多个感知特征。如表 5-3 所示，通过 PSO-GRNN 池化后的算法在预测性能和泛化性能上都得到了提升，如 SSIM 在 CSIQ 数据库中的 SROCC 值为 0.8756，而 PG-SSIM 在 CSIQ 数据库中的 SROCC 值为 0.9513；VSI 在 LIVE 数据库中的 SROCC 值为 0.9524，与之对应的 PG-VSI

在 LIVE 数据库中的 SROCC 值为 0.9937。使用 PSO-GRNN 池化的优势在 TID2013 数据库中表现最明显, SSIM 在 TID2013 数据库中的 SROCC 值为 0.7417, 而 PG-SSIM 在 TID2013 数据库中的 SROCC 值为 0.9184, 提升了近 30% 的 SROCC 值。在 10 个算法中表现最好的是 PG-FSIMc, 其 SROCC 值在三个数据库中最高为 0.9945, 最低为 0.9700。综上所述, 本文方法具有很好的竞争力和预测精度。

表 5-3 IQA 模型在三个数据库中的整体性能比较

Database	Criteria	SSIM	FSIMc	VSI	CAGS	MVFC	PG-SSIM	PG-FSIMc	PG-VSI	PG-CAGS	PG-MVFC
TID2013	SROCC	0.7417	0.8510	0.8965	0.8316	0.8606	0.9184	0.9700	0.9607	0.9511	0.9612
	KROCC	0.5588	0.6665	0.7183	0.6469	0.6810	0.7660	0.8668	0.8431	0.8293	0.8450
	PLCC	0.7895	0.8769	0.9000	0.8445	0.8822	0.9218	0.9740	0.9672	0.9584	0.9667
	RMSE	0.7608	0.5959	0.5404	0.6639	0.5838	0.0480	0.0281	0.0315	0.0354	0.0317
CSIQ	SROCC	0.8756	0.9310	0.9423	0.9198	0.9570	0.9513	0.9893	0.9878	0.9793	0.9847
	KROCC	0.6907	0.7690	0.7857	0.7487	0.8127	0.8055	0.9245	0.9103	0.8781	0.8995
	PLCC	0.8613	0.9192	0.9279	0.9014	0.9611	0.9555	0.9909	0.9894	0.9770	0.9856
	RMSE	0.1334	0.1034	0.0979	0.1137	0.0725	0.0774	0.0353	0.0382	0.0559	0.0444
LIVE	SROCC	0.9479	0.9645	0.9524	0.9734	0.9798	0.9871	0.9945	0.9937	0.9908	0.9940
	KROCC	0.7963	0.8363	0.8058	0.8658	0.8863	0.9207	0.9499	0.9427	0.9305	0.9435
	PLCC	0.9449	0.9613	0.9482	0.9640	0.9754	0.9856	0.9945	0.9935	0.9899	0.9934
	RMSE	8.9455	7.5296	8.6816	8.3251	6.9051	0.0530	0.0329	0.0356	0.0443	0.0358
Weighted Average	SROCC	0.7942	0.8847	0.9100	0.8588	0.9027	0.9382	0.9784	0.9722	0.9642	0.9720
	KROCC	0.6108	0.7101	0.7366	0.6828	0.7414	0.8044	0.8939	0.8753	0.8585	0.8747
	PLCC	0.8140	0.8928	0.9033	0.8575	0.9112	0.9407	0.9812	0.9765	0.9681	0.9755
Direct Average	SROCC	0.8350	0.9076	0.9223	0.8870	0.9256	0.9523	0.9846	0.9807	0.9737	0.9800
	KROCC	0.6557	0.7427	0.7555	0.7226	0.7770	0.8307	0.9137	0.8987	0.8793	0.8960
	PLCC	0.8422	0.9084	0.9131	0.8798	0.9297	0.9543	0.9865	0.9834	0.9751	0.9819

为了进一步验证 PSO-GRNN 算法的可靠性, 依据 CSIQ、LIVE 和 TID2013 数据库中所提供的 MOS 离散点, 绘制出 PG-FR 算法的预测散点图, 如图 5-7 所示。

可见, PSO-GRNN 算法符合 HVS 特性, 且绘制的回归曲线与主观观测值具有较好的相关性。

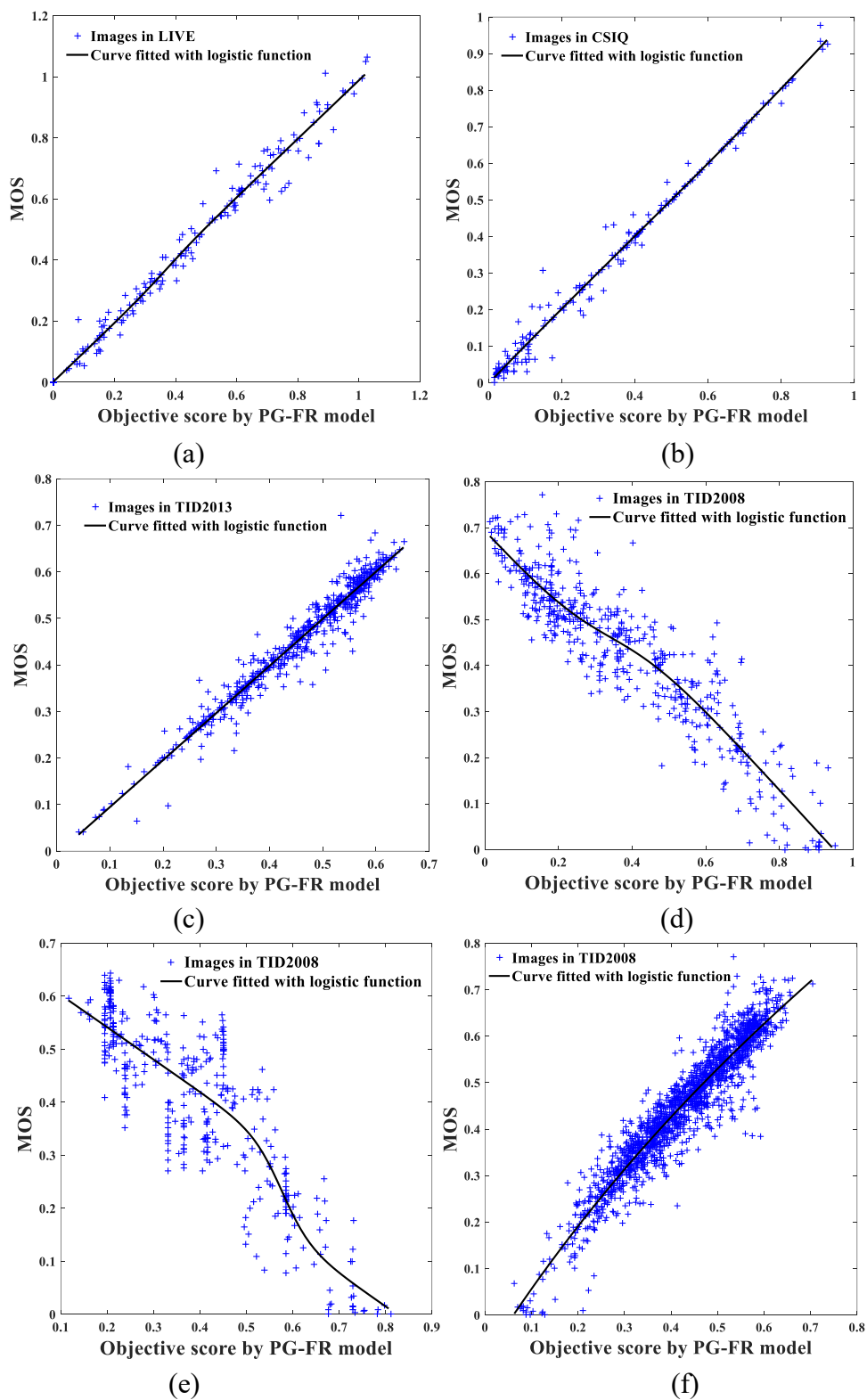


图 5-7 不同数据库上的 MOS 和模型预测得分的散点图。图 5-7 (a)、图 5-7 (b) 和图 5-7 (c) 分别表示在 CSIQ、LIVE 和 TID2013 数据库中测试集的测试结果，图 5-7 (d)、图 5-7 (e) 和图 5-7 (f) 分别为 LIVE、CSIQ 和 TID2013 数据库中已经训练好的网络模型在 TID2008 数据库中的预测结果。

(2) 单一失真类型的性能比较

这里比较不同算法在单一失真情况下的泛化性能。如表 5-4 所示,选用的数据库以及用作性能对比的算法都和表 5-3 一致,这里性能评价标准采用 SROCC 值。

由表 5-4 可见,经过 PSO-GRNN 池化处理后的算法相较于原算法在预测性能上有了稳定的提升。如 SSIM 在失真类型 AGN 下的 SROCC 值为 0.8671,而 PG-SSIM 在相同失真类型下的 SROCC 值为 0.9059; VSI 在失真类型 AGWN 下的 SROCC 值为 0.9636,与之对应的算法 PG-VSI 在 AGWN 下 SROCC 值为 0.9783。进一步分析 SSIM、VSI、PG-SSIM 和 PG-VSI 这 4 个算法在表 4 中的数据,SSIM 在表 5-4 中 SROCC 较低的是 Block、CTC 和 CCS 三种失真类型,其结果是 0.5720、0.3775 和 0.4141; VSI 结果较低的是 Block 和 CTC 两种失真类型,其结果是 0.1713 和 0.4754; PG-SSIM 在 Block、CTC 和 CCS 三种失真类型下的 SROCC 值分别为 0.6628、0.9253、0.8605; PG-VSI 在 Block 和 CTC 两种失真类型下的 SROCC 值分别为 0.8431 和 0.8408。以上数据可知,所提算法的 SROCC 值比原算法提升超过了 50%,有效处理较难预测的失真类型。在 52 个失真类型中预测性能最好的是 PG-VSI,其 SROCC 最高为 0.9928,最低为 0.8408,故本文算法的泛化性能也是较为可靠的。

表 5-4 IQA 模型在不同失真类型下的 SROCC 值

Database	Distortion type	SSIM	FSIMc	VSI	CAGS	MVFC	PG- SSIM	PG- FSIMc	PG- VSI	PG- CAGS	PG- MVFC
	AGN	0.8671	0.9101	0.9460	0.9359	0.9325	0.9059	0.9618	0.9562	0.9602	0.9567
	ANC	0.7726	0.8537	0.8705	0.8653	0.8434	0.7911	0.9496	0.9136	0.9104	0.9079
	SCN	0.8515	0.8900	0.9367	0.9276	0.9119	0.9256	0.9645	0.9737	0.9653	0.9557
	MN	0.7767	0.8094	0.7697	0.7526	0.7815	0.8306	0.8868	0.8885	0.8528	0.8584
	HFN	0.8634	0.9040	0.9200	0.9159	0.9080	0.9398	0.9686	0.9532	0.9647	0.9635
	IN	0.7503	0.8251	0.8741	0.8361	0.7814	0.8445	0.9466	0.9331	0.9126	0.8509
TID2013	QN	0.8657	0.8807	0.8748	0.8718	0.8993	0.9258	0.9651	0.9302	0.9660	0.9504
	GB	0.9668	0.9551	0.9612	0.9614	0.9268	0.9463	0.9853	0.9735	0.9789	0.9719
	DEN	0.9254	0.9330	0.9484	0.9466	0.9466	0.9223	0.9637	0.9658	0.9773	0.9642
	JPEG	0.9200	0.9339	0.9541	0.9585	0.9463	0.9217	0.9665	0.9532	0.9619	0.9808
	JP2K	0.9468	0.9589	0.9706	0.9620	0.9547	0.9559	0.9826	0.9787	0.9795	0.9754
	JGTE	0.8493	0.8610	0.9216	0.8644	0.8832	0.9290	0.9573	0.9370	0.9262	0.9531
	J2TE	0.8828	0.8919	0.9228	0.9250	0.8805	0.9197	0.9805	0.9623	0.9741	0.9227

	NEPN	0.7821	0.7937	0.8060	0.7833	0.8318	0.8182	0.9427	0.9155	0.8952	0.9124
	Block	0.5720	0.5532	0.1713	0.6015	0.6620	0.6628	0.7825	0.8431	0.7007	0.7707
	MS	0.7752	0.7487	0.7700	0.7441	0.7107	0.9095	0.9183	0.8538	0.7832	0.8817
	CTC	0.3775	0.4679	0.4754	0.4514	0.3147	0.9253	0.9036	0.8408	0.8897	0.9482
	CCS	0.4141	0.8100	0.2748	0.3711	0.5840	0.8605	0.9572	0.9660	0.9229	0.9433
	MGN	0.7803	0.8569	0.9117	0.8700	0.8688	0.8989	0.9366	0.9389	0.9413	0.9078
	CN	0.8566	0.9243	0.9121	0.9168	0.9345	0.9205	0.9668	0.9506	0.9592	0.9680
	LCNI	0.9057	0.9485	0.9564	0.9574	0.9656	0.8997	0.9125	0.9103	0.9162	0.9182
	ICQD	0.8542	0.8815	0.8839	0.9060	0.9061	0.9261	0.9410	0.9494	0.9522	0.9556
	CHA	0.8775	0.8925	0.8906	0.8768	0.8577	0.8927	0.9260	0.9220	0.9445	0.9574
	SSR	0.9461	0.9576	0.9628	0.9580	0.9603	0.9622	0.9856	0.9750	0.9752	0.9787
CSIQ	AGWN	0.8974	0.9359	0.9636	0.9652	0.9490	0.9331	0.9811	0.9783	0.9749	0.9781
	JPEG	0.9546	0.9664	0.9618	0.9573	0.9682	0.9639	0.9854	0.9875	0.9846	0.9796
	JP2K	0.9606	0.9704	0.9694	0.9545	0.9730	0.9772	0.9921	0.9825	0.9757	0.9880
	AGPN	0.8922	0.9370	0.9638	0.9492	0.9326	0.9564	0.9872	0.9860	0.9669	0.9702
	GB	0.9609	0.9729	0.9679	0.9574	0.9754	0.9774	0.9888	0.9874	0.9881	0.9875
	CTC	0.7922	0.9438	0.9504	0.9273	0.9274	0.9588	0.9776	0.9713	0.9687	0.9685
	JP2K	0.9614	0.9724	0.9604	0.9822	0.9707	0.9868	0.9935	0.9928	0.9923	0.9915
LIVE	JPEG	0.9764	0.9840	0.9761	0.9836	0.9804	0.9884	0.9925	0.9912	0.9926	0.9927
	AWGN	0.9694	0.9716	0.9835	0.9837	0.9669	0.9793	0.9945	0.9971	0.9890	0.9919
	GB	0.9517	0.9708	0.9527	0.9641	0.9689	0.9704	0.9895	0.9873	0.9855	0.9917
	FF	0.9556	0.9519	0.9430	0.9633	0.9680	0.9895	0.9906	0.9905	0.9853	0.9897

5.3 本章小结

本章基于提出的 PSO-GRNN 池化策略对多个与视觉特性相关的图像特征进行融合，并将融合的结果应用于图像质量评价当中，提出基于 PSO-GRNN 池化策略的全参考图像质量评价。实验结果表明 PG-FR 方法与人眼的感知结果有很好的一致性，该方法无论是在整体性能还是单个失真性能上都表现出很高的精度和泛化性。特别的，已经训练完成的 PSO-GRNN 网络只需要输入现有的图像特征就能迅速作出高精度评价。因此，本章提出的方法在实际中无需消耗过多时间。

论文在第三章、第四章和本章分别提出了 CPCC、MVFC 和 PG-FR 三个评价模型。这三个模型在四个公共数据库中测得的最高 SROCC 值分别为 0.9653、0.9798 和 0.9945。这表明，随着模型不断迭代，其预测性能也在不断提升。

第 6 章 总结与展望

6.1 工作总结

图像质量评价算法的目的是代替人眼视觉感知为图像质量做出评价,研究符合人眼主观感受的图像感知质量评价方法对于改进图像处理算法的性能有重大意义。

针对图像评价算法的高质量和高效率要求,论文提出了三种基于 HVS 特性的全参考图像质量评价方法。主要内容总结如下:

(1) 本文将色貌图上的相位信息作为研究对象,提出了色貌尺度相位一致性方法,该方法弥补了色度变化对相位信号的精度影响。视觉多通道分解特性表明,神经元对不同特征信号是平行处理的。据此,论文将色貌尺度相位一致性、对比度和色度三种特征进行融合,提出了基于色貌尺度相位一致性的全参考图像质量评价方法。实验结果表明,该方法不仅优于基于传统相位一致性的图像评价算法,还在大多数数据库中表现出可靠的结果。

(2) 人眼处理复杂的视觉信号是按照特定通路平行编码的。据此,本文采用 LoG 滤波器同时与图像上的三个颜色通道进行卷积,提出了多通道 LoG 滤波方法。该方法与仅在亮度通道上使用滤波分解所有不同,其考虑到了色度通道上的结构信息,有效的提高了信息的完整性。经过滤波后的图像对比度改变的,因此这里引入对比度比较作为算法的第二特征。本文是针对彩色图像的质量评价,故加入色度比较作为算法的第三特征。由此提出了基于多通道信息融合的全参考图像质量评价方法。实验表明,该方法无论在整体性还是在单个失真评价性能上都有着不错的表现。

(3) 人眼融合平行视觉信号的过程是高度复杂的,现有池化方法并不能很好的表达它。据此,本文使用图像特征分数和主观分数训练 PSO-GRNN 网络,然后将训练完成的 PSO-GRNN 响应面作为图像特征分数空间池,为此提出了基于 PSO-GRNN 池化策略的全参考图像质量评价方法。基于机器学习的评价方法一般都依赖数据集的标签数量,有些评价方法通过将图像分成多个模块来增加标

签的总量，但是忽略了图像块之间的联系。本文是通过分解出多个图像特征分量来为网格训练提供更多数据。实验结果表明，所提的池化方法大大的提升了评价模型的预测性能和泛化性能。

6.2 未来展望

图像质量评价一直是图像处理中一个困难但活跃的研究领域。本文提出的方法还有许多地方可以改进和延续。下面将从两点展望论文工作。

第一点是针对色空间转化问题。论文第三章所提的色空间转换方法只考虑了 CIELAB 匀色空间，而 RGB 转换到 CIELAB 颜色空间的计算复杂度较高。此外，提取图像相位一致性所花费的时间也有所偏高。从实验结果可以看出，所提方法的计算效率并不理想。未来将从颜色空间转化效率出发，探索出高精度和高效率的评价算法。

第二点是针对池化方法的选择问题。论文在第四章中采用了标准偏差池化图像的三个特征，并未对其它池化方法做出深入探索。不同的池化方法会影响所提模型评价单个失真类型的性能。实验数据表明，本章模型在评价特点失真类型时，表现的并不理想。因此，未来将进一步研究不同池化方法和所提模型评价性能之间的关系。

总的来说，未来图像质量评价领域将继续深入研究和探索，以提高图像质量评价的预测精度、泛化性和预测精度，满足不断增长的现实需求。

参考文献

- [1] Madhusudana P C, Birkbeck N, Wang Y, et al. Image quality assessment using contrastive learning[J]. IEEE Transactions on Image Processing, 2022, 31: 4149-4161.
- [2] 卢鹏, 刘楷赞, 邹国良. 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] 胡波, 谢国庆, 李雷达, 等. 图像重定向质量评价的研究进展[J]. 中国图象图形学报, 2024, 29(1): 22-44.
- [4] 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.
- [5] 单月, 刘段, 万晓霞. 无参考图像质量评价研究现状与前景分析[J]. 包装工程, 2022, 43(13): 296-304.
- [6] Gao M J, Dang H SH, Wei L L. Review and Prospect of Full Reference Image Quality Assessment[J]. Acta. Electronica. Sinica. 2021, 49(11): 2261-2272.
- [7] Wei L, Zhao L, Peng J. Reduced Reference Quality Assessment for Image Retargeting by Earth Mover's Distance[J]. Applied Sciences. 2021, 11(20): 9776.
- [8] Varga D. No-Reference Video Quality Assessment Using the Temporal Statistics of Global and Local Image Features[J]. Sensors, 2022, 22(24): 9696.
- [9] 高敏娟, 党宏社, 魏立力. 全参考图像质量评价回顾与展望[J]. 电子学报, 2021, 49(11): 2261.
- [10] Varga D. No-Reference Image Quality Assessment with convolutional neural networks and decision fusion[J]. Applied Sciences, 2021, 12(1): 101.
- [11] Eskicioglu A M, Fisher P S. Image quality measures and their performance[J]. IEEE Transactions on communications, 1995, 43(12): 2959-2965.
- [12] Avcibas I, Sankur B, Sayood K. Statistical evaluation of image quality measures[J]. 2001.
- [13] 丰明坤, 孙丽慧, 葛丁飞. 深度视觉频谱残余融合的图像质量评价[J]. 光电子激光, 2021, 32(10): 1055-1064.
- [14] Wang Z, Bovik A, Sheikh H. Image Quality Assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [15] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for Image Quality Assessment[C]. IEEE The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, 2: 1398-1402.
- [16] Li C, Bovik A C. Three-component weighted structural similarity index[C]. SPIE Image quality and system performance VI, 2009, 7242: 252-260.
- [17] 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.
- [18] Kukkonen H, Rovamo J, Tiippana K, et al. Michelson contrast, RMS contrast and energy of various spatial stimuli at threshold[J]. Vision research, 1993, 33(10): 1431-1436.
- [19] Liu A M, Lin W S, Narwaria M. Image Quality Assessment Based on Gradient Similarity[J]. IEEE Transactions on Image Processing, 2012: 1500-1512.
- [20] Sheikh H R, Bovik A C, De Veciana G. An information fidelity criterion for Image Quality Assessment using natural scene statistics[J]. IEEE Transactions on image processing, 2005, 14(12): 2117-2128.
- [21] Mokayed H, Shivakumara P, Woon H H, et al. A new DCT-PCM method for license plate number detection in drone images[J]. Pattern Recognition Letters, 2021, 148: 45-53.
- [22] Zhang L, Zhang L, Mou X, et al. FSIM: A feature similarity index for Image Quality Assessment[J]. IEEE transactions on Image Processing, 2011, 20(8): 2378-2386.
- [23] Yao K, Cheng J, Chen Y, et al. Multi-exposure fusion light field image quality assessment with motion region detection[C]. International Conference on Frontiers of Applied Optics and Computer Engineering (AOCE 2024). SPIE, 2024, 13080: 78-85.
- [24] Liu L, Gao M, Zhang Y. Application of machine learning in intelligent encryption for digital information of real-time image text under big data[J]. EURASIP Journal on Wireless Communications and Networking, 2022, 2022(1): 1-16.
- [25] Kang L, Ye P, Li Y. Convolutional neural networks for No-Reference Image Quality Assessment[C]. IEEE Computer Vision and Pattern Recognition, 2014: 1733-1740.
- [26] 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.
- [27] Gao F, Wang Y, Li P. Deepsim: Deep similarity for Image Quality Assessment[J]. Neurocomputing, 2017, 257: 104-114.
- [28] Li C, Bovik A C, Wu X. Blind Image Quality Assessment Using a General Regression Neural Network[J]. IEEE Transactions on Neural Networks, 2011(5): 22.
- [29] 朱惠娟, 宗平, 丛玉华. 基于权重池的多尺度图像质量评估方法[J]. 计算机工程与应用, 2021, 57(03): 215-221.
- [30] 贾曦然, 隆海燕. 双通道立体图像质量评价方法的研究[J]. 科学技术与工程, 2023, 23(4): 1589-1597.
- [31] 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.
- [32] Wang Z, Wu G, Sheikh H R, et al. Quality-aware images[J]. IEEE transactions on image processing, 2006, 15(6): 1680-1689.

- [33] 杨春玲,杨雅静.基于多尺度特征逐层融合深度神经网络的无参考图像质量评价方法[J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(4).
- [34] 尹海涛, 岳勇赢. 基于半监督学习和生成对抗网络的医学图像融合算法[J]. Laser & Optoelectronics Progress, 2022, 59(22): 2215005-2215005-10.
- [35] Sheikh H R, Bovik A C, Cormack L. No-Reference Quality Assessment using natural scene statistics: JPEG2000[J]. IEEE Transactions on image processing, 2005, 14(11): 1918-1927.
- [36] Pan Z, Yuan F, Wang X, et al. No-reference image quality assessment via multibranch convolutional neural networks[J]. IEEE Transactions on Artificial Intelligence, 2022, 4(1): 148-160.
- [37] Zhou J, Xiao B, Li Q. A no reference Image Quality Assessment method for JPEG2000[C]. IEEE International Joint Conference on Neural Networks, 2008: 863-868.
- [38] Ren H, Deng J, Xie X. Grnn: generative regression neural network—a data leakage attack for federated learning[J]. ACM Transactions on Intelligent Systems and Technology, 2022, 13(4): 1-24.
- [39] Ma J, Wu J, Li L, et al. Active inference of GAN for no-reference Image Quality Assessment[C]. IEEE International Conference on Multimedia and Expo, 2020: 1-6.
- [40] Sheikh H. LIVE Image Quality Assessment database release 2[J]. <http://live.ece.utexas.edu/research/quality>, 2005.
- [41] Yang X, Wang T, Ji G. Image Quality Assessment via colour information fluctuation[J]. Signal, Image and Video Processing, 2022: 1-11.
- [42] Ponomarenko N, Lukin V, Zelensky A. TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics[J]. adv modern radioelectron, 2009.
- [43] Ling Y, Zhou F, Guo K. ASSP: An adaptive sample statistics-based pooling for Full-Reference Image Quality Assessment[J]. Neurocomputing, 2022: 493.
- [44] Ninassi A, Le Callet P, Atrousseau F. Pseudo no reference image quality metric using perceptual data hiding[C]. SPIE Human vision and electronic imaging XI, 2006, 6057: 146-157.
- [45] 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.
- [46] Jones J P, Palmer L A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex[J]. Journal of neurophysiology, 1987, 58(6): 1233-1258.
- [47] 闫乐乐, 李辉, 邱聚能. 基于区域对比度和 SSIM 的图像质量评价方法[J]. 应用光学, 2015(1): 58-63.
- [48] Campbell F W, Robson J G. Application of Fourier analysis to the visibility of gratings[J]. The Journal of physiology, 1968, 197: 551-566.

- [49] 范媛媛, 桑英军, 沈湘衡. 视觉掩盖效应下的无参考图像信噪比评价[J]. 应用光学, 2012, 33(4): 711-716.
- [50] 肖亮, 韦志辉. 一种利用人眼视觉掩盖的小波域数字小印[J]. 通信学报, 2002, 23(3): 100-106.
- [51] Hecht S. The visual discrimination of intensity and the Weber-Fechner law[J]. The Journal of general physiology, 1924, 7(2): 235.
- [52] Suthaharan, S. Perceptual quality metric for digital video coding[J]. Electronics Letters, 2003, 39(5): 431-433.
- [53] Shen D F, Wang S C. Measurements of JND property of HVS and its applications to image segmentation, coding, and requantization[C]. SPIE Digital Compression Technologies and Systems for Video Communications, 1996, 2952: 113-121.
- [54] Matheson H E, Barsalou L W. Embodiment and grounding in cognitive neuroscience[J]. Stevens' handbook of experimental psychology and cognitive neuroscience, 2018, 3: 1-27.
- [55] Mannos J, Sakrison D. The effects of a visual fidelity criterion of the encoding of images[J]. IEEE transactions on Information Theory, 1974, 20(4): 525-536.
- [56] Taylor C C, Pizlo Z, Allebach J P, et al. Image Quality Assessment with a Gabor pyramid model of the human visual system[C]. SPIE Human Vision and Electronic Imaging II, 1997, 3016: 58-69.
- [57] Berns R S. Extending CIELAB: Vividness, depth, and clarity[J]. Color Research & Application, 2014, 39(4): 322-330.
- [58] 史晨阳, 林燕丹. 基于图像色貌和梯度特征的图像质量客观评价[J]. 物理学报, 2020, 69(22): 401-412.
- [59] Cheng J, Ren B, Zhao L, et al. Analysis of phase response of fiber Fabry-Pérot cavity microphones[J]. The Journal of the Acoustical Society of America, 2021, 149(2): 1012-1020.
- [60] Megahed Y, Shaker A, Yan W Y. A Phase-Congruency-Based Scene Abstraction Approach for 2D-3D Registration of Aerial Optical and LiDAR Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, (99): 1-1.
- [61] Bhuiyan M A A, Khan A R. Image Quality Assessment employing RMS contrast and histogram similarity[J]. Int. Arab J. Inf. Technol, 2018, 15(6): 983-989.
- [62] 明军, 梁彦, 王新. 一种与色度相关的图像质量评价方法[J]. 安徽大学学报:自然科学版, 2008, 32(5): 37-40.
- [63] Ahmed I T, Der C S, Hammad B T, et al. Contrast-distorted Image Quality Assessment based on curvelet domain features[J]. International Journal of Electrical and Computer Engineering, 2021, 11(3): 25-95.

- [64] Larson E C, Chandler D M. Most apparent distortion: full-reference Image Quality Assessment and the role of strategy[J]. Journal of electronic imaging, 2010, 19(1): 011006-011006.
- [65] Xue W, Zhang L, Mou X. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index[J]. IEEE transactions on image processing, 2013, 23(2): 684-695.
- [66] Sheikh H R, Bovik A C. Image information and visual quality[J]. IEEE Transactions on Image Processing, 2006, 15(2): 430-444.
- [67] 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.
- [68] Yang G, Li D, Lu F. RVSIM: a feature similarity method for Full-Reference Image Quality Assessment[J]. EURASIP Journal on Image and Video Processing, 2018, 2018(1): 6.
- [69] Young, Richard A. The Gaussian derivative model for spatial vision: I. Retinal mechanisms[J]. Spatial Vision, 1987, 2(4): 273.
- [70] Temel D, Alregib G. PerSIM: Multi-resolution Image Quality Assessment in the perceptually uniform color domain[C]. IEEE International Conference on Image Processing. 2015: 1682-1686.
- [71] Enroth-Cugell C, Robson J G. The contrast sensitivity of retinal ganglion cells of the cat[J]. The Journal of physiology, 1966, 187(3): 517-552.
- [72] Li C, Zhang Z, Sun W, et al. A full-reference quality assessment metric for cartoon images[C]. 2022 IEEE 24th International Workshop on Multimedia Signal Processing. IEEE, 2022: 1-6.
- [73] Yang C C, Kwok S H. Efficient gamut clipping for color image processing using LHS and YIQ[J]. Optical Engineering, 2003, 42(3): 701-711.
- [74] Wen X, Pan Z, Hu Y. Generative Adversarial Learning in YUV Color Space for Thin Cloud Removal on Satellite Imagery[J]. Remote Sensing, 2021, 13(6): 1079.
- [75] Ding K, Ma K, Wang S, et al. Image quality assessment: Unifying structure and texture similarity[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 44(5): 2567-2581.
- [76] Varga D. Saliency-Guided Local Full-Reference Image Quality Assessment[J]. Signals, 2022, 3(3): 483-496.
- [77] Wang X, Zheng B J, Kong L J. Full Reference Image Quality Assessment Based on Visual Saliency and Perception Similarity Index[J]. Packaging Engineering, 2022, 43(9): 239-248.
- [78] Varga D. Full-Reference Image Quality Assessment Based on Grünwald–Letnikov Derivative, Image Gradients, and Visual Saliency[J]. Electronics, 2022, 11(4): 559.
- [79] Specht D F. A general regression neural network[J]. IEEE transactions on neural networks,

1991, 2(6): 568-576.

- [80] Zaman H R R, Gharehchopogh F S. An improved particle swarm optimization with backtracking search optimization algorithm for solving continuous optimization problems[J]. Engineering with Computers, 2022, 38(4): 2797-2831.
- [81] Fang J, Liu W, Chen L, et al. A survey of algorithms, applications and trends for particle swarm optimization[J]. International Journal of Network Dynamics and Intelligence, 2023: 24-50.
- [82] Gad A G. Particle swarm optimization algorithm and its applications: a systematic review[J]. Archives of computational methods in engineering, 2022, 29(5): 2531-2561.
- [83] Jain M, Saihjpal V, Singh N. An overview of variants and advancements of PSO algorithm[J]. Applied Sciences, 2022, 12(17): 8392.
- [84] Kang Y, Song J, Jia D, et al. An improved GRNN model and prediction of noctiluca density[J]. Journal of Computational Methods in Sciences and Engineering, 2022, 22(4): 1131-1139.
- [85] Gu J, Cai H, Dong C, et al. NTIRE 2022 challenge on perceptual image quality assessment[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 951-967.
- [86] Yang S, Wu T, Shi S, et al. Maniqa: Multi-dimension attention network for no-reference image quality assessment[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 1191-1200.

攻读学位期间发表的学术论文目录

一、论文发表情况

- [1] Jiang Benchu, **Bian Shilei**, Shi Chenyang, et al. Full-Reference Image Quality Assessment Based on Multi-Channel Visual Information Fusion[J]. Applied Sciences, 2023, 13(15): 8760. (SCI 收录)
- [2] 江本赤, **卞仕磊**, 史晨阳, 吴路路. 基于色貌尺度相位一致性的全参考图像质量评价[J]. Optics and Precision Engineering, 2023, 31(10): 1509-1521. (EI 收录)
- [3] **Bian Shilei**, Jiang Benchu, Shi Chenyang, et al. Full reference image quality evaluation based on PSO-GRNN pooling strategy[J]. IEEE Transactions on Artificial Intelligence. (已投稿)

二、专利申请情况

- [1] 江本赤, **卞仕磊**, 周焰生, 等. 面向异形物体的机器人抓取手爪[P]. CN202111215368.6, 2021.12.17.
- [2] 江本赤, **卞仕磊**, 周焰生, 等. 小曲率自由曲面仿形打磨装置及其打磨方法[P]. CN202111181091.X, 2021.12.17.
- [3] 江本赤, 周焰生, **卞仕磊**, 等. 双轴头尾式自动焊接变位机及工件焊接方法[P]. CN20211118 1670.4, 2021.12.24.

致谢

时间如白驹过隙稍纵即逝，转眼间三年学习生涯即将结束。这三年，在整个人生中不算很长，但作为一名专业硕士，这三年里学到的知识和道理是未来任何一个三年都无法比拟的。对于如今即将毕业的我来说，2021 年入学的秋天，仿佛就在昨天。三年前，我怀着激动且热切的心情找到了江本赤教授，希望能成为他的学生；三年后，在这论文即将完成之际，我由衷地感谢我的导师江本赤教授和史晨阳老师。

在学术上，江老师学识渊博，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，从入学至今，每周六的 701 办公室集中辅导已经成为了我们课题组的家常便饭。从辅导我们写专利到写小论文再到大论文，一步一个脚印，坦白地说，我不是一个很聪明灵活的人，在写专利或者对学术的理解时，经常会比别人慢半拍，但江老师多次对我进行了专门辅导，耐心为我解惑甚至补充学习本科内容，这些经历时常浮现在记忆中，每每想起，总会忍不住会心一笑。2022 年 8 月，研究生论文开题工作即将展开，导师江本赤教授和史晨阳老师为了让我们对所研究课题有更清晰的认知，亲自带我们到相关企业去考察研究，两位老师结合企业实际不厌其烦地为我们答疑解惑，在学术研究和实践经验上给了我们很大启发。从开题到中期再到如今即将完成，论文的每一部分都少不了江本赤老师的悉心指导，可以说在课题研究和毕业论文写作中给了我们极大帮助。第一次正式见史晨阳老师时，他在看人工智能相关的书籍，这并不是偶然，即使身为导师，也仍坚持每天学习，这种学习精神令我十分敬佩，也激励着我不断前进。在生活中，两位老师勤俭节约，待人真诚，和蔼可亲，在择业、待人处事、生活态度等各个方面，都和我们有过深入交流，为我们树立榜样。

除此之外，我还要感谢我的同门，感谢大师兄詹庆荣，师兄李健康、杜欣和段鹏飞，同届周焰生、朱明生、刘艺、陶文博、杨凯、詹伟健、邹令令和杨怀安等人，感谢他们对我学习及生活中的帮助，也感谢三年中关心或帮助过我的其他老师同学和朋友。

最后，我要感恩我的家人，尤其是我的父母。当我大学临近毕业之际，我向父母提出了考研的想法，他们毫不犹豫就同意了我的想法，并不遗余力的给我提供支持，他们是我坚强的后盾。父母已年近六十，忙碌了半辈子，一年中也只有春节那半个多月的休息和团聚时间，母亲身体不好，不能工作！爸爸在工地上班，他的双手充满了老茧。我曾尝试作为一个普工在工地兼职一个暑假，但每次都是因为扛不动那么多砖块而只能打杂。这一整个暑假我都腰酸背痛，爸爸却说不累，我多么希望他们能别那么辛苦。现在，我即将毕业，以后的日子我来为你们分担！正是你们无条件的支持和鼓励，我的学习生涯才能走到这一步，感谢你们对我这几十年的无私付出！

作者：卞仕磊

2024 年 2 月 22 日

尚德敏学 唯实惟新

