

基于流形特征相似度的感知图像质量评价

王朝云¹ 蒋刚毅¹ 郁梅^{1,2} 陈芬¹

摘要 图像质量评价 (Image quality assessment, IQA) 的目标是利用设计的计算模型得到与主观评价一致的结果, 而人类视觉感知特性是感知图像质量评价的关键. 大量研究发现, 认知流形和拓扑连续性是人类感知的基础即人类感知局限在低维流形之上. 基于图像低维流形特征分析, 本文提出了基于流形特征相似度 (Manifold feature similarity, MFS) 的全参考图像质量评价方法. 首先, 利用正交局部保持投影算法来模拟大脑的视觉处理过程获取最佳映射矩阵进而得到图像的低维流形特征, 通过流形特征的相似度来表征两幅图像的结构差异, 从而反映感知质量上的差异. 其次, 考虑亮度失真对人眼视觉感知的影响, 通过图像块均值计算亮度相似度并用于评价图像的亮度失真; 最后, 结合两个相似度得到图像的客观质量评价. 在四个公开图像测试库上的实验结果表明, 所提出方法与现有代表性的图像质量方法相比总体上具有更好的评价结果.

关键词 图像质量评价, 流形特征相似度, 正交局部保持投影, 视觉感知

引用格式 王朝云, 蒋刚毅, 郁梅, 陈芬. 基于流形特征相似度的感知图像质量评价. 自动化学报, 2016, 42(7): 1113–1124

DOI 10.16383/j.aas.2016.c150559

Manifold Feature Similarity Based Perceptual Image Quality Assessment

WANG Chao-Yun¹ JIANG Gang-Yi¹ YU Mei^{1,2} CHEN Fen¹

Abstract Image quality assessment (IQA) aims to use computational models to measure the image quality in consistency with subjective evaluation, and human visual perception characteristics play an important role in the design of IQA metrics. From many researches on human visual perception, it has been found that the cognitive manifolds and the topological continuity can be used to describe the human visual perception, that is, human perception lies on the low-dimensional manifold. With this inspiration and manifold analysis of image, a new IQA metric called manifold feature similarity (MFS) is proposed for full-reference image quality assessment. First, orthogonal locality preserving projection algorithm is used to simulate the brain's visual processing process to obtain the best projection matrix so that low-dimensional manifold features of images are obtained. And the similarity of the manifold features is used to measure the structure differences between the two images so as to reflect differences in perceived quality and get a manifold features-based image quality index. Then, to consider the impact of brightness on human visual perception, the block mean values of the image are used to calculate the distortion of the image's brightness and design a brightness-based image quality index. The final quality score is obtained by incorporating these two indices. Extensive experiments on four large scale benchmark databases demonstrate that the proposed IQA metric works better than all state-of-the-art IQA metrics in terms of prediction accuracy.

Key words Image quality assessment (IQA), manifold feature similarity (MFS), orthogonal locality preserving projections, visual perception

Citation Wang Chao-Yun, Jiang Gang-Yi, Yu Mei, Chen Fen. Manifold feature similarity based perceptual image quality assessment. *Acta Automatica Sinica*, 2016, 42(7): 1113–1124

收稿日期 2015-09-06 录用日期 2015-12-07
Manuscript received September 6, 2015; accepted December 7, 2015

国家高技术研究发展计划 (863 计划) (2015AA015901), 国家自然科学基金 (U1301257, 61271270, 61311140262), 浙江省自然科学基金 (LY15F010005, LY16F010002) 资助

Supported by National High Technology Research and Development Program of China (863 Program) (2015AA015901), National Natural Science Foundation of China (U1301257, 61271270, 61311140262), and Zhejiang Provincial Natural Science Foundation (LY15F010005, LY16F010002)

本文责任编辑 王立威

Recommended by Associate Editor WANG Li-Wei

1. 宁波大学信息科学与工程学院 宁波 315211 2. 南京大学计算机软件新技术国家重点实验室 南京 210093

1. Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211 2. National Key Laboratory of

图像质量评价是图像处理领域中充满挑战性的问题^[1-3]. 由于人是观看图像时的最终接收者, 因此图像质量评价方法应当和人眼的主观感知相一致. 传统的峰值信噪比 (Peak signal-to-noise ratio, PSNR) 等基于保真度的图像质量评价方法虽能较好地评价具有相同内容和失真类型的图像质量, 但面对多幅图像和多种失真时, 其评价结果会出现与主观感知相去甚远的情况. 图像质量评价方法的目的是通过模拟人眼视觉系统的整体感知机制来获取与视觉感知质量具有较高一致性的评价结果. 近年来, 图像质量评价的研究不断深入, 人们提出了很多

Software New Technology, Nanjing University, Nanjing 210093

的评价方法. Wang 等^[4]提出的结构相似度 (Structural similarity, SSIM) 算法与 PSNR 等方法相比性能改进明显, 引起了学者们的关注; 在其后续工作中, Wang 等又提出了多尺度的 SSIM (Multi-scale structural similarity, MS-SSIM), 改进了 SSIM 的性能^[5]. Zhang 等^[6]提出了基于 Riesz 变换的特征相似度 (Riesz transform-based feature similarity, RFSIM) 评价算法, 提取了基于一阶和二阶 Riesz 变换的图像局部结构并利用 Canny 边缘特征用于质量加权. 文献 [7] 认为人眼在对局部图像评分时相位一致性和梯度幅值起着相辅相成的作用, 提出特征结构相似度 (Feature similarity, FSIM). 程光权等^[8]探索自然图像的几何结构特征, 考虑像素点的方向失真、幅度失真和方差失真, 提出了一种基于几何结构失真模型的全参考质量评价方法. 除了基于结构的图像质量评价方法外, 另一些评价方法是从人眼视觉系统的其他特性出发进行设计的. Chandler 等^[9]提出视觉信噪比 (Visual signal-to-noise ratio, VSNR), 该准则先通过视觉阈值确定失真是否可察觉, 再对超过视觉阈值的区域进行失真度量. Larson 等^[10]认为人类视觉系统 (Human visual system, HVS) 在评测高质量图像和低质量图像时采用了不同策略, 提出最明显失真 (Most apparent distortion, MAD) 的质量评价算法. Sheikh 等^[11]将全参考图像质量评价问题看作信息保真度问题, 在信息保真度 (Information fidelity criterion, IFC)^[12]的基础上进行拓展得到视觉信息保真度 (Visual information fidelity, VIF) 评价算法. Zhang 等^[13]发现质量下降会造成图像显著图的变化且与感知质量失真程度密切, 从而提出基于视觉显著性 (Visual saliency induced index, VSI) 的图像质量评价方法. 考虑到结构和对比度变化可以通过图像梯度的变化得到, Liu 等^[14]提出基于梯度相似度 (Gradient similarity metric, GSM) 的全参考质量评价算法.

优异的图像质量评价方法应能很好反映人眼视觉感知特性. 针对视觉感知现象, 有研究表明流形是感知的基础, 大脑中以流形方式对事物进行感知^[15]; 因此, 将图像流形特征应用于视觉质量评价可得到与主观感知一致性较高的评价结果. 流形学习能较好地帮助找到图像在低维流形中的内在几何结构, 反映出事物的非线性流形的本质^[16]. Cai 等^[17]对局部保持投影 (Locality preserving projection, LPP) 算法进行改进, 得到正交局部保持投影算法 (Orthogonal locality preserving projection, OLPP), 该方法可找到数据的流形结构且具有线性特点, 适用于所有邻域空间而并非局限于样本点. Charrier 等^[18]则针对 JPEG 2000 失真的图像提出一种基

于机器学习的图像质量分类方法, 它利用现有的全参考与无参考图像质量方法提取的特征来描述图像, 然后尝试了主成分分析 (Principal components analysis, PCA) 线性降维和拉普拉斯特征映射的流行学习的非线性降维等两种方式的降维方法来对提取到的特征向量进行降维用于训练后续的 SVM 得到分类器; 但从其结果分析中可以看出相较于 PCA 的线性降维而言, 拉普拉斯特征映射的非线性降维并不能有效地帮助其提高质量分类的准确性. 而本文将从人眼视觉感知的角度出发, 利用流形学习的方法直接从图像块中学习得到图像低维流形特征, 得到符合人眼感知特性的图像特征用于图像质量评价.

结合上述分析, 本文提出一种基于流形特征相似度 (Manifold feature similarity, MFS) 的图像质量评价方法 (MFS 准则). 在训练阶段, MFS 准则将利用流形学习 OLPP 算法得到最佳映射矩阵用于提取图像的流形特征; 在质量预测阶段, 在将原始与失真图像划分为图像块后, 去除每个块的均值使所有块向量都具有零均值, 在其基础上得到流形特征相似度; 而所有块均值则用于计算亮度相似度. 其中, 流形特征相似度表征了两幅图的结构差异, 而亮度相似度则度量了失真图像的亮度失真. 最后, 结合两个相似度得到图像的质量评价. 实验结果表明所提出方法的评价结果与人眼主观评价具有很高的一致性.

1 基于流形特征相似度 (MFS) 的图像质量评价准则 (MFS 准则)

流形是感知的基础, 经过自然界长期进化的人脑能够以流形的方法表达对外界对象的感知^[15]. 大脑中神经元群体活动通常可描述为一个神经放电率的集合的结果, 如果一个神经元的触发率对应于一维, 那么图像信息就能够由与像素个数相等的神经元来表示. 研究发现每个神经元在一个神经元群体中的放电率可用一个少数变量的平滑函数表示, 比如人眼转动的角度和头旋转的方向^[19], 这说明神经元群体活动是局限在一个低维流形之上.

基于流形学习理论, 本文定义了流形特征相似度的概念, 进而提出一种基于流形特征相似度的图像质量评价新方法 (MFS 准则); 它使用 OLPP 来模拟神经元群体的视觉感知过程, 并给出显式的最优映射矩阵用于提取测试图像的低维流形特征. 所提出的 MFS 准则框架如图 1 所示, 它分为两个阶段: 训练和相似度计算. 首先, 从训练阶段得到最佳映射矩阵 J , 并将其用于后续参考和失真图像块的流形特征的提取; 然后, 在流形特征的基础上计算图

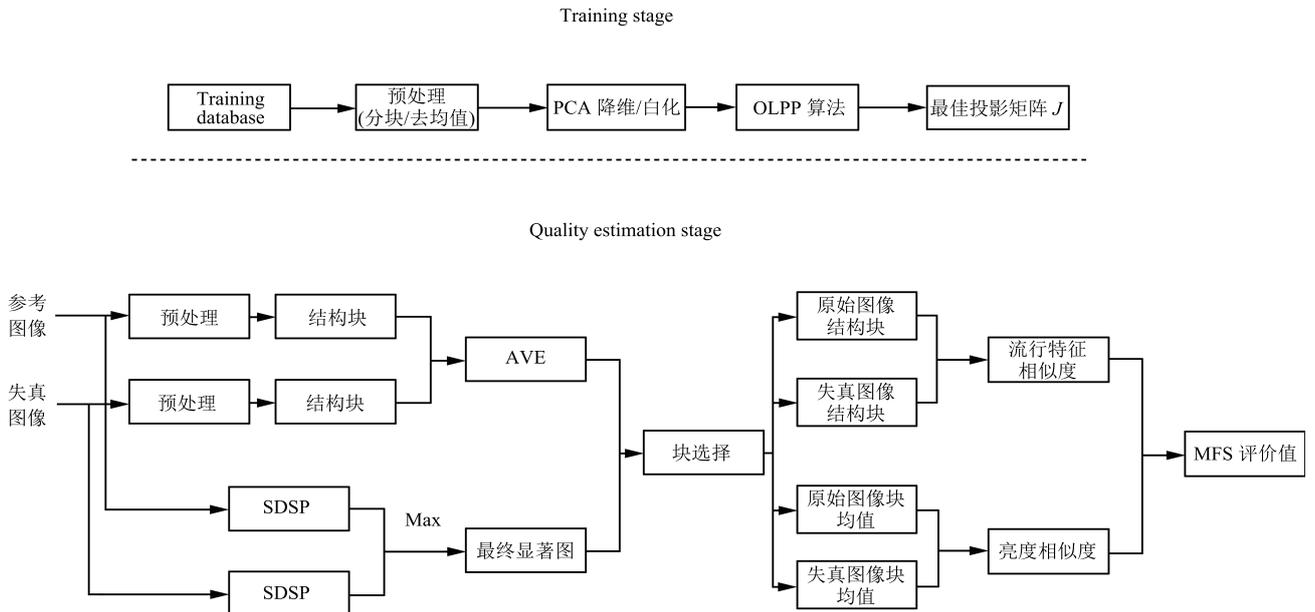


图 1 基于流形特征相似度的图像质量评价准则

Fig. 1 Manifold feature similarity based perceptual image quality index

像质量值.

1.1 训练获取最佳映射矩阵

训练过程先从无失真的 N 幅自然图像中随机选取上万个图像块作为训练样本, 再通过主成分分析对样本向量进行降维以及白化处理; 然后, 通过 OLPP 对白化后的数据 X_w 进行训练得到白化空间中的正交投影矩阵 J_w , 最后将其还原到原始样本空间得到最佳映射矩阵 J .

1) 预处理. 在训练过程的初始阶段, 从无失真的 10 幅自然图像中随机选取 20 000 个 8×8 图像块作为训练样本. 在实际计算时, 需要将每个图像块按逐个通道逐行转换为列向量. 由于彩色图像有三个通道, 由此得到长度为 $8 \times 8 \times 3 = 192$ 的向量. 最后, 每个向量通过减去对应图像块的均值进行中心化, 所有中心化后的样本向量组成矩阵 X .

2) 利用 PCA 进行降维和白化. 有研究表明, 人眼视网膜和外侧膝状体 (Lateral geniculate nucleus, LGN) 会对输入的视觉信号进行白化处理^[20]. 进一步的研究还表明, 视网膜和 LGN 具有很好的去视觉冗余效果; 因此, 这里采用 PCA 进行降维和白化来模拟视网膜和 LGN 的该功能. 一方面, 这样做去除了样本中冗余信息同时减少了计算量; 另一方面, 这样避免了当数据维数大于样本点数时, OLPP 算法中的广义特征值求解很不稳定的问题. 降维和白化处理过程如下:

通过数据协方差矩阵的特征值分解来实现 PCA 过程. 样本数据 X 的协方差矩阵 C 计算

如下:

$$C = \frac{1}{N} (X \times X^T) \quad (1)$$

其中, N 代表样本数. 在进行特征值分解后, 令 $\Psi = \text{diag}\{\Psi_1, \dots, \Psi_M\}$ 和 $E_M = (e_1, \dots, e_M)$ 分别代表从协方差矩阵 C 分解得到的对角特征值矩阵中按特征值从大到小顺序取 M 个最大特征值对应的行向量组成矩阵 Ψ 和对应特征向量矩阵 E 中的 M 个特征向量组成的矩阵 E_M . 在实际应用中, 可只取前 8 个主成分分量用于训练, 也就是说白化后样本向量的维数从 192 维降到了 8 维. 白化矩阵 W 给出如下:

$$W = \Psi^{-\frac{1}{2}} \times E^T \quad (2)$$

其中, $\Psi^{-\frac{1}{2}} = \text{diag}\{1/\sqrt{\Psi_1}, \dots, 1/\sqrt{\Psi_M}\}$. 最终, 通过以下操作将样本数据 X 白化为矩阵 X_w :

$$X_w = W \times X \quad (3)$$

3) 使用 OLPP 算法进行训练. 最佳映射矩阵是从白化样本数据 X_w 中通过流形学习获取的正交投影矩阵还原到原始样本空间得到的. 当高维数据分布于嵌入在子空间的低维流形中时, 正交局部保持投影算法 (OLPP) 通过寻找流形的最佳线性逼近的拉普拉斯 Beltrami 算子进行低维嵌入, 具体流程如下:

步骤 1. 构建邻接图. 令 G 代表有 m 个结点的图, 其中, m 代表样本点个数, 第 a 个结点对应样本 x_{wa} . 当样本 x_{wa} 和 x_{wb} 邻近时则将结点 a 和 b 之

间进行边相连. 通过 k 邻近 (k nearest neighbors) 来判断结点间是否边相连, 即如果结点 a 在结点 b 的 k 邻近的结点中或结点 b 在结点 a 的 k 邻近的结点中时, 结点 a 和 b 之间有边相连 (k 为整数, 取值为 5).

步骤 2. 利用热核构造样本相似矩阵 S . S 为 $m \times m$ 的稀疏对称矩阵, 当结点 a 和 b 之间有边相连给定权值 $S_{ab} = e^{-\|\mathbf{x}_{wa} - \mathbf{x}_{wb}\|^2}$, 否则 $S_{ab} = 0$.

步骤 3. 计算正交基函数. 定义 Φ 为一个对角矩阵, $\Phi_{aa} = \sum_{b=1}^m S_{ab}$; 同时定义拉普拉斯矩阵 $L = \Phi - S$. 令 $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ 为正交基向量组, 定义如下矩阵:

$$P^{(n-1)} = [\mathbf{p}_1, \dots, \mathbf{p}_{n-1}] \quad (4)$$

$$Q^{(n-1)} = [P^{(n-1)}]^T (X\Phi\Phi^T)^{-1} P^{(n-1)} \quad (5)$$

得到上述矩阵后, 正交基向量 $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ 计算如下: \mathbf{p}_1 为 $(X\Phi X^T)^{-1} X L X^T$ 的最小的非零特征值对应的特征向量, \mathbf{p}_n 为 $H^n = \{I - (X\Phi X^T)^{-1} \times P^{(n-1)} [Q^{(n-1)}]^{-1} [P^{(n-1)}]^T\} (X\Phi X^T)^{-1} X L X^T$ 的最小的非零特征值对应的特征向量.

令白化空间中的正交投影矩阵为 $J_w = [\mathbf{p}_1, \dots, \mathbf{p}_l]$, 其中 $l = 8$.

步骤 4. 在进行学习之后, 正交投影矩阵 J_w 应该从白化样本空间转化回到原始样本空间:

$$J = J_w \times W \quad (6)$$

其中, W 为白化矩阵, J_w 表示在白化样本空间中的正交投影矩阵, J 为最终的原始样本空间的最佳映射矩阵. 这里, 将 J 看作大脑以流形方式感知的一个模型, 可以用于提取图像块的流形特征.

1.2 MFS 评价计算

如图 1 所示, MFS 评价值的计算由两部分组成: 特征相似度计算和亮度相似度计算. 为了处理方便, 在计算之前, 将参考图像和失真图像通过 8×8 的滑动窗口分为不重叠的块, 同时对每个图像块进行去均值操作; 由于图像块在去均值后的块包含了对比度和结构等信息, 将其称作结构块. 因此, 可将每个图像块看作由对应的均值块和去均值后的结构块组成. 在全参考图像质量评价中, 参考图像和失真图像采用的是相同的划分方式, 从而可将每个参考图像块和对应的失真图像块作为一对图像对. 由于图像的均值块不包括任何对比度和结构等信息, 因此在计算流形特征相似度时将不使用图像块的均值. 但考虑到图像亮度的变化还是会引起感知失真虽然其影响并不如对比度和结构变化那么大^[14], 因此, 引入亮度相似度来描述图像的亮度失真, 此时则需要利用图像块的均值进行计算. 最后, 通过组合特征相

似度和亮度相似度得到 MFS 评价值.

1.2.1 利用视觉特性选块

在去除每个块的均值后, 原始和失真图像的结构块成为一个零均值的列向量 $\mathbf{x}_i^{\text{ref}}$ 和 $\mathbf{x}_i^{\text{dis}}$ (i 对应图像的第 i 个块), 所有来自参考图像和对应的失真图像的列向量 $\mathbf{x}_i^{\text{ref}}$ 和 $\mathbf{x}_i^{\text{dis}}$ 组成两个矩阵, 即 X^{ref} 和 X^{dis} . 超阈值失真视觉关注中的一个重要影响因素, 在感知质量中也有着举足轻重的作用^[21], 这也是人类视觉系统对图像中的低质量区域比高质量区域更敏感的原因. 正是由于低质量区域对质量评价有更大的影响, 所以通过使用结构差异较大的图像块对来进行质量评价可以提高评价性能^[22]. 为了保持评价方法的执行效率同时提高评价的准确性, 本文使用特定标准来衡量图像块的结构差异, 同时在其基础上设计阈值进行选块进而利用选取的图像块进行相似度计算.

首先, 由于 $\mathbf{x}_i^{\text{ref}}$ 和 $\mathbf{x}_i^{\text{dis}}$ 均是在减均值后的参考和对应的失真图像块向量, 因此, $\sum_{g=1}^h (\mathbf{x}_i^{\text{ref}})_g^2$ 和 $\sum_{g=1}^h (\mathbf{x}_i^{\text{dis}})_g^2$ 代表的均是预处理前原始参考图像和失真图像块的方差. 利用绝对方差差值 (Absolute variance error, AVE) 来衡量图像块的结构差异. 令 $\mathbf{x}_i^{\text{ref}}$ 和 $\mathbf{x}_i^{\text{dis}}$ 分别代表参考和对应的失真图像结构块向量, 则 $\mathbf{x}_i^{\text{ref}}$ 和 $\mathbf{x}_i^{\text{dis}}$ 的 AVE 定义如下:

$$AVE(\mathbf{x}_i^{\text{ref}}, \mathbf{x}_i^{\text{dis}}) = \left| \sum_{g=1}^h (\mathbf{x}_i^{\text{ref}})_g^2 - \sum_{g=1}^h (\mathbf{x}_i^{\text{dis}})_g^2 \right| \quad (7)$$

其中, h 代表每个图像块向量的元素数目. 最终, 所有的参考和失真图像块对的 AVE 值形成一个向量 \mathbf{v} , 向量 \mathbf{v} 中的元素为 $\mathbf{v}_i = AVE(\mathbf{x}_i^{\text{ref}}, \mathbf{x}_i^{\text{dis}})$.

为了选择一组有利于质量评价的图像块对, 在 \mathbf{v} 向量的中值的基础上设计了阈值 TH_x , 利用该阈值来选取参考-失真块对. 如果某一块对的 AVE 值不小于设定的阈值 TH_x , 则保留该块对用于质量评价, 最后所有选取的参考和失真图像块向量记为 $\mathbf{y}_i^{\text{ref}}$ 和 $\mathbf{y}_i^{\text{dis}}$. 最终, 所有保留下来的向量组成两个矩阵, Y^{ref} 和 Y^{dis} , 如式 (8) 所示:

$$(Y^{\text{ref}}, Y^{\text{dis}}) = \{(\mathbf{y}_i^{\text{ref}}, \mathbf{y}_i^{\text{dis}}) | AVE(\mathbf{y}_i^{\text{ref}}, \mathbf{y}_i^{\text{dis}}) > TH_x\}, \quad TH_x = \text{median}(\mathbf{v}) \quad (8)$$

其中, $\text{median}(\cdot)$ 代表选取中值的运算.

给定一幅图像, 利用适当的视觉显著性 (Visual saliency, VS) 模型计算其显著图能反映每个局部区域在人眼视觉系统中的显著程度. VS 和 IQA 有着紧密的联系, 它们都依赖于 HVS 如何感知一幅图像同时在视觉关注中超阈值失真也是一个重要的影响因素^[21], 图像不同区域在 HVS 感知图像质量过程

中有着不同的作用, 显然, 使用与人眼关注特点相关的选块策略可以提高质量评价性能. 然而, 上述的选块只考虑了结构差异大的区域, 这些区域一般对应失真图像质量较低的区域但并不一定是人眼最关注的区域. 因此, 在利用阈值 TH_x 来选取参考-失真块对之后, 使用视觉显著计算模型 (Saliency detection based-on simple priors, SDSP)^[23] 计算原始和失真图像的显著图, 使用 VS 图来表征图像不同区域的视觉重要性, 即给出图像对 f_r 和 f_d 中的每个图像块的 VS 值, 无论 f_r 还是 f_d 中第 i 块图像具有较高 VS 值就说明位置 i 处的图像块在评价图像 f_r 和 f_d 相似度时具有较大的影响. 为此, 对求得的 VS 图按 8×8 不重叠分块后, 分别求取各块的平均显著度得到图像对 f_r 和 f_d 中的每个图像块的 VS 值形成显著图 VS_1 和 VS_2 . 最后, 利用 $VS_m(i) = \max\{VS_1(i), VS_2(i)\}$ 来对图像进行二次选块:

$$(Y^{\text{ref}}, Y^{\text{dis}}) = \left\{ (\mathbf{y}_i^{\text{ref}}, \mathbf{y}_i^{\text{dis}}) \mid AVE(\mathbf{y}_i^{\text{ref}}, \mathbf{y}_i^{\text{dis}}) > TH_x \text{ and } VS_m(i) > TH_{vs} \right\} \quad (9)$$

其中, TH_{vs} 为将所有图像块显著值按降序排序后在前 60% 位置的显著值即选取显著度最高的 60% 的块对 AVE 选块进行二次细选.

1.2.2 流形特征相似度

在图像块对选取结束后, 通过如下操作得到流形特征向量 \mathbf{r}_t 和 \mathbf{d}_t :

$$\mathbf{r}_t = J \times \mathbf{y}_i^{\text{ref}}, \quad \mathbf{d}_t = J \times \mathbf{y}_i^{\text{dis}} \quad (10)$$

由于 J 的大小为 8×192 , 则 \mathbf{r}_t 和 \mathbf{d}_t 的向量长度为 $L = 8$. 为了简洁表示, 使用 $(\mathbf{r}_t, \mathbf{d}_t)$ 向量对来表示参考图像和失真图像块的特征. 而所有流形特征向量 \mathbf{r}_t 和 \mathbf{d}_t 组成两个矩阵 R 和 D .

$$(R, D) = \{(\mathbf{r}_t, \mathbf{d}_t) \mid t = 1, \dots, K\} \quad (11)$$

其中, K 代表在图像中选取的图像块数目, \mathbf{r}_t 和 \mathbf{d}_t 分别为 R 和 D 的列向量.

最后, 定义 MFS 中的流形特征相似度 MFS_f , 计算如下:

$$MFS_f = \frac{1}{K \times M} \sum_{t=1}^K \sum_{j=1}^L \frac{2R_{tj}D_{tj} + C_1}{R_{tj}^2 + D_{tj}^2 + C_1} \quad (12)$$

其中, K 代表一幅图像中选取的图像块数目, 即保留的流形特征向量的数目, R_{tj} 和 D_{tj} 分别表示 R 和 D 的第 t 列和第 j 行的值; C_1 为一个很小的常量, 用于保证结果的稳定性.

1.2.3 亮度相似度

亮度相似度的计算是基于每个图像块的均值进行的, 仅考虑使用在上述选块过程中获取的图像块对对应的均值向量 $(\boldsymbol{\mu}^{\text{ref}}, \boldsymbol{\mu}^{\text{dis}})$ 来定义和计算亮度相似度 MFS_m , 计算过程如式 (13).

$$MFS_m = \frac{\sum_{i=1}^K (\boldsymbol{\mu}_i^{\text{ref}} - \text{mean}(\boldsymbol{\mu}^{\text{ref}})) \times (\boldsymbol{\mu}_i^{\text{dis}} - \text{mean}(\boldsymbol{\mu}^{\text{dis}})) + C_2}{\sqrt{\sum_{i=1}^K (\boldsymbol{\mu}_i^{\text{ref}} - \text{mean}(\boldsymbol{\mu}^{\text{ref}}))^2 + \sum_{i=1}^K (\boldsymbol{\mu}_i^{\text{dis}} - \text{mean}(\boldsymbol{\mu}^{\text{dis}}))^2 + C_2}} \quad (13)$$

其中, $\boldsymbol{\mu}^{\text{ref}}$ 和 $\boldsymbol{\mu}^{\text{dis}}$ 为选块后对应块的均值组成的向量, $\text{mean}(\cdot)$ 代表取向量的均值; C_2 为一个很小的常量, 用于保证结果的稳定性.

1.2.4 MFS 准则的评价值

最后, 对 MFS_f 和 MFS_m 进行线性加权得到 MFS 评价值作为失真图像的质量分数.

$$MFS = \omega \times MFS_m + (1 - \omega) \times MFS_f \quad (14)$$

其中, $0 < \omega < 1$ 用于调节 MFS_f 和 MFS_m 两个分量的线性加权的权值.

2 实验结果与分析

为了验证本文算法的有效性, 在 4 个公开的测试图像库上对本文算法进行了测试和对比; 这 4 个图像库包括 LIVE^[24]、CSIQ^[10]、TID 2008^[25] 和 TID 2013^[26]. 每个图像库包含上千幅失真图像, 同时拥有多种失真类型. 每幅失真图像都会给定一个主观分数例如平均主观分 (Mean opinion score, MOS) 或平均主观分差值 (Differential mean opinion score, DMOS). 各图像库中的参考图像数、失真图像数、失真类型数以及参与主观实验的人数如表 1 所示. 最终的算法性能验证是在比较主观评分与图像质量评价算法客观评价结果基础上进行的.

表 1 应用于图像质量评价算法分析的 4 个测试图像库
Table 1 The four benchmark datasets for evaluating IQA indices

图像库	参考图像数	失真图像数	失真类型数	主观测试人数
TID 2013	25	3 000	25	971
TID 2008	25	1 700	17	838
CSIQ	30	866	6	35
LIVE	29	799	5	161

采用 4 个通用评价指标并根据视频质量评价专

家组 PhaseI/II(VQEG)^[27] 提供的标准验证方法来获取 IQA 的评价性能. 斯皮尔曼秩相关系数 (Spearman rank-order correlation coefficient, SROCC) 和肯德尔秩次相关系数 (Kendall rank-order correlation coefficient, KROCC) 用于评价 IQA 方法的预测单调性的优劣, 这两个指标仅在排序后的数据上进行而忽略数据点之间的相对距离. 为了获取另外两个指标皮尔森线性相关系数 (Pearson linear correlation coefficient, PLCC) 和均方根误差 (Root mean squared error, RMSE), 需要对客观评价价值和主观平均评分 (MOS) 进行非线性映射来去除客观分数的非线性的影响. 采用 5 参数非线性映射函数来进行非线性拟合.

$$Q(q) = \alpha_1 \left(\frac{1}{2} - \frac{1}{1 + \exp(\alpha_2(q - \alpha_3))} \right) + \alpha_4 q + \alpha_5 \quad (15)$$

其中, q 代表原始的客观质量评价分数, Q 代表非线性映射后的分数. 5 个调节参数 $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ 则是由最小化映射后的客观分数与主观评分之间的方差和确定的.

本文提出的 MFS 准则将与具有代表性的 10 个图像质量评价算法进行比较, 包括: SSIM^[4]、MS-SSIM^[5]、IFC^[12]、VIF^[11]、VSNR^[8]、MAD^[10]、GSM^[14]、RFSIM^[6]、FSIMc^[7] 和 VSI^[13].

2.1 参数确定

本文所提出算法在求取特征相似度时, 使用了参数 C_1 来保证结果的稳定性但该参数的改变同时也会在一定程度上影响特征相似度的计算值从而影响最终的评价结果. 为了获取最佳的参数 C_1 , 在固定其他参数 ω 和 C_2 时 ($\omega \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, $C_2 < 10^{-2}$) 进行测试时发现, 当 $C_1 = 0.09$ 时 MFS 在 4 个库上的 SROCC 达到最高, 因此 C_1 取 0.09. 而由于在计算亮度相似过程中式 (13) 中的分子分母数量级均较大, C_2 在 10^{-2} 的数量级以下对亮度相似度的计算结果几乎没有影响, 这里 C_2 取为 0.001. 在最后线性加权特征相似度和亮度相似过程中的参数 ω 则在 $\omega \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ 中选择使得 MFS 在 4 个库上的评价指标 SROCC 达到最高的 0.8 作为最终参数.

2.2 块选择策略

为了验证选块对提高本文算法的评价准确性的作用, 选用了 3 个不同的策略来对图像进行选块, 最后选用 SROCC 来作为验证指标, 每种选块策略得到的评价结果如表 2 所示.

表 2 3 种选块策略对应的 SROCC 值

Table 2 The SROCC of three selection strategies

图像库	不选块	AVE 选块	AVE + VS 选块
TID 2013	0.8655	0.8728	0.8741
TID 2008	0.8763	0.8870	0.8893
CSIQ	0.9508	0.9621	0.9615
LIVE	0.9500	0.9600	0.9578

从表 2 中可知, 利用 AVE 选块能有效地提高算法的评价结果, 这归功于 HVS 对图像中的低质量区域比高质量区域更敏感. 正是由于低质量区域对质量评价具有重大影响, 使用具有巨大结构差异的图像块对来进行质量评价可以提高评价性能. 而在加上 VS 选块后, 在 CSIQ 图像库和 LIVE 图像库上的评价性能较仅使用 AVE 选块策略有下降外在其他两个库上均有提升 (在最大的两个图像库 TID 2008 和 TID 2013 上效果均有提升), 由此可以看出本文选块策略的有效性. 在 CSIQ 和 LIVE 图像库使用 AVE + VS 选块后评价性能虽然相较于 AVE 选块略有下降, 但是相较于不选块来说评价准确性仍有提高, 此处的可能原因是 VS 算法无法完全精确地估计图像显著性, 最终导致了在失真图像数目并不多的 CSIQ 和 LIVE 库的性能略有下降, 相信在失真图像数目越多时 AVE 加 VS 选块性能的优越性将更加明显.

2.3 PCA 白化对评价性能的影响

MFS 利用 OLPP 在白化空间中求取正交基, 然后将其还原到原始样本空间作为最佳映射矩阵来提取图像块的流形特征. 表 3 给出了 MFS 在不同的去冗余后的白化空间中使用 OLPP 寻找正交基, 并将其还原到原始样本空间作为最佳映射矩阵来提取图像块的流形特征得到的 SROCC 值, 可以看出直接使用 OLPP (不降维) 的性能在 4 个图像库上均有很大程度的下降, 因此使用 PCA 白化来模拟人眼的去视觉冗余过程是十分必要的.

2.4 训练库及训练样本数目的影响

由于最佳映射矩阵是通过训练获得的, 因此需要考虑训练样本数目及训练样本库对最终的评价性能的影响, 使用的训练样本库如图 2 和图 3 所示. 图 4 给出了使用在图像样本集 S2 上由不同样本块数目训练获取的最佳映射矩阵来提取流形特征评价图像质量获得的 SROCC 值与图像块样本数目的关系. 其中, 图像样本集 S1 来自于 IVC 图像库的无失真图像而图像样本集 S2 来自 TOY 图像库的无失真图像.



图 2 用于 OLPP 训练的图像集 S1, 其中的图像均来自 IVC 的无失真图像
Fig.2 The set S1 for OLPP, the images in the set were picked from IVC dataset

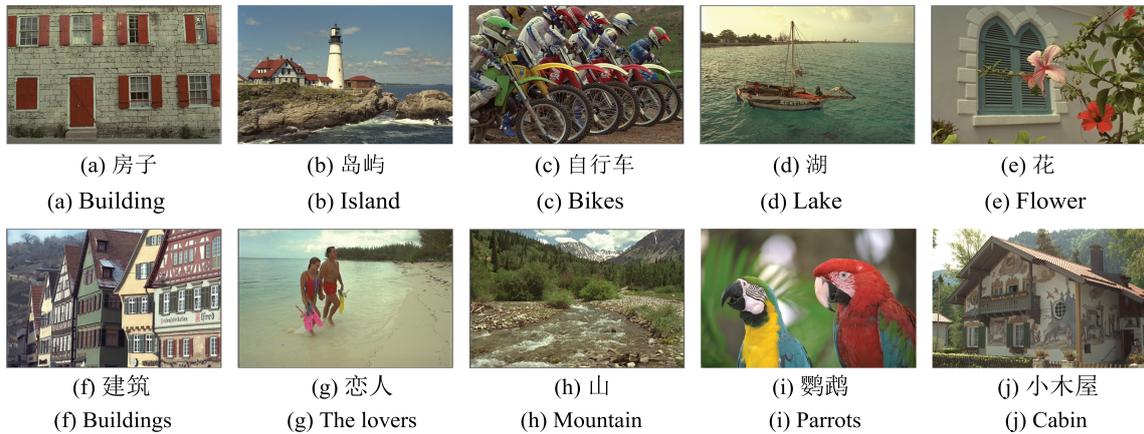


图 3 用于 OLPP 训练的图像集 S2, 其中的图像均来自 TOY 的无失真图像
Fig.3 The set S2 for OLPP, the images in the set were picked from TOY dataset

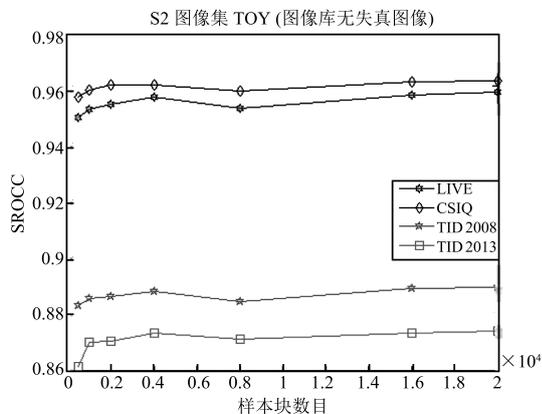


图 4 样本图像块数目与 SROCC 关系
Fig.4 The relationship between sample numbers and SROCC

表 3 不同的 PCA 白化降维维数下, MFS 在 4 个图像库上 SROCC 值

Table 3 The SROCC of MFS at different whitening dimensions on four datasets

PCA 白化后的空间维数	LIVE	CSIQ	TID 2008	TID 2013
8	0.9578	0.9615	0.8893	0.8741
16	0.9509	0.9594	0.8820	0.8585
24	0.9507	0.9587	0.8754	0.8483
32	0.9332	0.9235	0.8205	0.8059
64	0.9283	0.9067	0.7311	0.7444
不降维	0.8163	0.6962	0.2863	0.3864

从图 4 可知, 当样本库图像包含足够多的内容且样本数目足够多时最佳映射矩阵在各个库上的评价性能趋于稳定. 整体趋势是随着样本图像块数目的增加, 评价性能上升并趋于稳定, 本文从测试图像集中随机选择了 20 000 个图像块作为样本进行训练, 取得了不错的效果.

另外, 表 4 给出了在相同参数及样本数目条件下在两个训练集图像库上的 SROCC 值比较, 从中可以发现选用不同的训练库得到的性能结果基本相同, 这说明所提出算法受训练样本的不同的影响微乎其微. 因此本文中的最佳映射矩阵是一个通用的流形特征提取器, 一旦通过 OLPP 训练获取后便可以用于所有图像质量的评价, 而不需要每次评价都进行耗时的训练过程. 值得注意的是, 本文中其他结果均是在 S2 图像集的基础上获得的.

表 4 在两个训练集图像库上的 SROCC 值比较
Table 4 The SROCC of MFS on two training sets

训练集	LIVE	CSIQ	TID 2008	TID 2013	Average
S1	0.9594	0.9615	0.8866	0.8579	0.9168
S2	0.9578	0.9615	0.8893	0.8741	0.9206

2.5 整体性能与比较

表 5 中给出了 ω 取不同值时, 本文的 MFS 准则在 4 个图像库上的 SROCC 值; 表 6 则是给出了

仅使用流形特征进行图像质量评价得到的 SROCC 值. 从表 6 中可知, 在仅考虑流形特征 ($\omega = 0$) 时, MFS 仍然具有较高的评价性能, 在 CSIQ 图像库上, MFS 准则性能最优, 在 LIVE、TID 2013 图像库上 MFS 的评价性能位列第三, 而在 TID 2008 图像库上 MFS 表现较差, 位居第五; 但 MFS 准则的平均性能在所有方法中位居第三. 显然, 所提出 MFS 准则中的流形特征在图像评价中发挥了很大的作用; 而加入亮度分量是对图像评价的一个补充, 这是因为在提取流形特征前去除了图像块的均值, 因此将亮度分量作为评价指标中的一部分加入了最终的评价公式中.

表 7 给出每个 IQA 方法在 4 个数据库上的 4 个预测性能指标 SROCC、KROCC、PLCC 和 RMSE, 表中对所有 IQA 方法中指标性能最优的 2 个 IQA 方法以黑体标出. 从表 7 可知, MFS 准则在所有图像库上的性能都很好. 首先, 在 CSIQ 图像库上, MFS 的性能最优, 优于其他所有 IQA 方法. 其次, 比起其他所有的 IQA 算法, 在最大的两个图像库 TID 2008 和 TID 2013 上的性能较大幅度地优于其他算法, 且与 VSI 算法性能接近. 虽然在 LIVE 库上 MFS 的性能不是最佳的, 但与最佳的 IQA 方法的评价性能相差甚微. 相比之下, MFS 之外的一些方法可能在某些库上效果不错但是在其他库上的效果差强人意. 例如, VIF 和 MAD 在 LIVE 具有很好的评价效果, 但在 TID 2008 和 TID 2013 上的表现却很糟糕. 因此, 整体上来说, 与其他算法相比,

表 5 ω 取不同值时, MFS 在 4 个图像库上的 SROCC 值
Table 5 The SROCC of MFS when ω takes different values

ω	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
LIVE	0.9553	0.9554	0.9557	0.9560	0.9562	0.9565	0.9568	0.9573	0.9578	0.9575	0.9248
CSIQ	0.9516	0.9523	0.9532	0.9542	0.9553	0.9565	0.9579	0.9594	0.9615	0.9635	0.7700
TID 2008	0.8377	0.8439	0.8509	0.8584	0.8661	0.8739	0.8817	0.8879	0.8893	0.8769	0.6598
TID 2013	0.8407	0.8455	0.8505	0.8554	0.8602	0.8650	0.8696	0.8733	0.8741	0.8675	0.7081

表 6 仅考虑流形特征时 MFS 的评价性能 ($\omega = 0$)
Table 6 The performance when just considering the manifold feature ($\omega = 0$)

	SSIM	MS-SSIM	IFC	VIF	VSNR	MAD	GSM	RFSM	FSIMc	VSI	MFS
TID 2013 SROCC	0.7471	0.7859	0.5389	0.6769	0.6812	0.7808	0.7946	0.7744	0.8510	0.8965	0.8407
TID 2008 SROCC	0.7749	0.8542	0.5675	0.7491	0.7046	0.8340	0.8504	0.8680	0.8840	0.8979	0.8377
CSIQ SROCC	0.8756	0.9133	0.7671	0.9195	0.8106	0.9466	0.9108	0.9295	0.9310	0.9423	0.9516
LIVE SROCC	0.9479	0.9513	0.9259	0.9636	0.9274	0.9669	0.9561	0.9401	0.9645	0.9524	0.9553
Ave SROCC	0.8363	0.8761	0.6998	0.8272	0.7809	0.8820	0.8779	0.8780	0.9076	0.9222	0.8963

表 7 11 种方法在 4 个图像库上的整体性能比较 ($\omega = 0.8$)
Table 7 The total performance comparison of 11 IQA indices ($\omega = 0.8$)

		SSIM	MS-SSIM	IFC	VIF	VSNR	MAD	GSM	RFSM	FSIMc	VSI	MFS
TID 2013	SROCC	0.7471	0.7859	0.5389	0.6769	0.6812	0.7808	0.7946	0.7744	0.8510	0.8965	0.8741
	KROCC	0.5588	0.6407	0.3939	0.5147	0.5084	0.6035	0.6255	0.5951	0.6665	0.7183	0.6862
	PLCC	0.7895	0.8329	0.5538	0.7720	0.7402	0.8267	0.8464	0.8333	0.8769	0.9000	0.8856
	RMSE	0.7608	0.6861	1.0322	0.7880	0.8392	0.6975	0.6603	0.6852	0.5959	0.5404	0.5757
TID 2008	SROCC	0.7749	0.8542	0.5675	0.7491	0.7046	0.8340	0.8504	0.8680	0.8840	0.8979	0.8893
	KROCC	0.5768	0.6568	0.4236	0.5860	0.5340	0.6445	0.6596	0.6780	0.6991	0.7123	0.7055
	PLCC	0.7732	0.8451	0.7340	0.8084	0.6820	0.8308	0.8422	0.8645	0.8762	0.8762	0.8865
	RMSE	0.8511	0.7173	0.9113	0.7899	0.9815	0.7468	0.7235	0.6746	0.6468	0.6466	0.6211
CSIQ	SROCC	0.8756	0.9133	0.7671	0.9195	0.8106	0.9466	0.9108	0.9295	0.9310	0.9423	0.9615
	KROCC	0.6907	0.7393	0.5897	0.7537	0.6247	0.7970	0.7374	0.7645	0.7690	0.7857	0.8260
	PLCC	0.8613	0.8991	0.8384	0.9277	0.8002	0.9502	0.8964	0.9179	0.9192	0.9279	0.9614
	RMSE	0.1344	0.1149	0.1431	0.0980	0.1575	0.0818	0.1164	0.1042	0.1034	0.0979	0.0722
LIVE	SROCC	0.9479	0.9513	0.9259	0.9636	0.9274	0.9669	0.9561	0.9401	0.9645	0.9524	0.9578
	KROCC	0.7963	0.8045	0.7579	0.8282	0.7616	0.8421	0.8150	0.7816	0.8363	0.8058	0.8199
	PLCC	0.9449	0.9489	0.9268	0.9604	0.9231	0.9675	0.9512	0.9354	0.9613	0.9482	0.9543
	RMSE	8.9455	8.6188	10.264	7.6137	10.506	6.9073	8.4327	9.6642	7.5296	8.6816	8.1691

所提出的 MFS 算法的质量预测结果与主观评价更加接近. 另一方面, 表 7 给出的是考虑图像亮度的 MFS 方法与其他算法的整体性能比较, 从中可以看出亮度分量确实对图像质量评价有一定的辅助作用但不是决定性作用.

2.6 特定失真上的性能比较

为了更加综合地评判 IQA 方法预测特定失真引起的图像质量降质的能力, 将本文 MFS 准则与其他对比算法在特定失真下的评价性能进行了测试. 选择 SROCC 作为性能指标, 因为 SROCC 适用于数据点较少的情况而且不会受到非线性映射的影响, 当然使用其他的性能指标例如 KROCC, PLCC 和 RMSE 也可以得到类似的结论. 表 8 给出了 4 个图像库中 52 组特定失真子库的评价结果.

表 8 中用粗体标识出每个图像库中每种失真类型下的 SROCC 值前三的 IQA 方法. 可以看出 VSI 的 IQA 方法共 31 次位于前三, 而 MFS 准则共 25 次位于前三, 其次是 FSIMc 和 GSM. 因此, 可以得出如下结论: 总的来说, 在特定失真类型下, VSI 的表现最优, 而 MFS 紧随其后, 其次是 FSIMc 和 GSM. 最重要的是, VSI, MFS, FSIMc 和 GSM 均大大优于其他方法. 另外, 在最大的两个库 TID 2008 和 TID 2013 上, 本文的 MFS 准则对 AGN、SCN、MN、HFN、IN、JP2K、J2TE 等失真的评价性能较其他算法更加优异, 而在 LIVE

和 CSIQ 图像库上则是 AGWN、GB 两种失真的评价效果最优. 针对 TID 2008 和 TID 2013 中的 Block、MS 和 CTC 失真, 本文的 MFS 准则的评价性能有待提高.

2.7 算法时间复杂度

表 9 给出了各个 IQA 方法处理一对 384×512 (取自 TID 2013 图像库) 的彩色图像需要的运行时间. 实验是在 lenovo 台式机上进行的, 其中处理器为 Intel(R) core(TM) i5-4590, CPU 为 3.3 GHz, 内存为 8 GB, 软件平台为 Matlab R2014b. 从表 9 可知, MFS 具有一个折中的时间复杂度. 特别地, 它比 IFC、VIF、MAD、FSIMc 等运行速度更快, 但却得到了与其接近甚至更好的评价效果.

3 结论

从人眼视觉感知的流形描述的角度出发, 本文提出了一种新颖的全参考图像质量评价算法, 即流形特征相似度 (Manifold feature similarity, MFS) 准则. MFS 准则的计算过程分为两部分: 训练和保真度计算. 首先, 利用正交局部保持投影算法从自然图像上获取样本块进行训练获得一个通用的最佳映射矩阵. 接着计算评价值的两个组成成分: 特征相似度和亮度相似度. 特征相似度是在结构块的基础上提取流形特征进行计算得到, 而亮度相似度则是基于图像块的均值来求取的. 最后, 将特征相似度和

表 8 11 种方法在特定失真上的 SROCC 评价值
Table 8 SROCC values of 11 IQA indices for each type of distortions

	Type	SSIM	MS-SSIM	IFC	VIF	VSNR	MAD	GSM	RFSM	FSIMc	VSI	MFS	
TID 2013	AGN	0.8671	0.8646	0.6612	0.8994	0.8271	0.8843	0.9064	0.8878	0.9101	0.9460	0.9153	
	ANC	0.7726	0.7730	0.5352	0.8299	0.7305	0.8019	0.8175	0.8476	0.8537	0.8705	0.8273	
	SCN	0.8515	0.8544	0.6601	0.8835	0.8013	0.8911	0.9158	0.8825	0.8900	0.9367	0.9001	
	MN	0.7767	0.8073	0.6932	0.8450	0.7072	0.7380	0.7293	0.8368	0.8094	0.7697	0.8186	
	HFN	0.8634	0.8604	0.7406	0.8972	0.8455	0.8876	0.8869	0.9145	0.9040	0.9200	0.9063	
	IN	0.7503	0.7629	0.6408	0.8537	0.7363	0.2769	0.7965	0.9062	0.8251	0.8741	0.8313	
	QN	0.8657	0.8706	0.6282	0.7854	0.8357	0.8514	0.8841	0.8968	0.8807	0.8748	0.8421	
	GB	0.9668	0.9673	0.8907	0.9650	0.9470	0.9319	0.9689	0.9698	0.9551	0.9612	0.9553	
	DEN	0.9254	0.9268	0.7779	0.8911	0.9081	0.9252	0.9432	0.9359	0.9330	0.9484	0.9178	
	JPEG	0.9200	0.9265	0.8357	0.9192	0.9008	0.9217	0.9284	0.9398	0.9339	0.9541	0.9377	
	JP2K	0.9468	0.9504	0.9078	0.9516	0.9273	0.9511	0.9602	0.9518	0.9589	0.9706	0.9633	
	JGTE	0.8493	0.8475	0.7425	0.8409	0.7908	0.8283	0.8512	0.8312	0.8610	0.9216	0.8885	
	J2TE	0.8828	0.8889	0.7769	0.8761	0.8407	0.8788	0.9182	0.9061	0.8919	0.9228	0.9081	
	NEPN	0.7821	0.7968	0.5737	0.7720	0.6653	0.8315	0.8130	0.7705	0.7937	0.8060	0.7727	
	Block	0.5720	0.4801	0.2414	0.5306	0.1771	0.2812	0.6418	0.0339	0.5532	0.1713	0.1755	
	MS	0.7752	0.7906	0.5522	0.6276	0.4871	0.6450	0.7875	0.5547	0.7487	0.7700	0.6285	
	CTC	0.3775	0.4634	0.1798	0.8386	0.3320	0.1972	0.4857	0.3989	0.4679	0.4754	0.4598	
	CCS	0.4141	0.4099	0.4029	0.3099	0.3677	0.0575	0.3578	0.0204	0.8359	0.8100	0.8102	
	MGN	0.7803	0.7786	0.6143	0.8468	0.7644	0.8409	0.8348	0.8464	0.8569	0.9117	0.8630	
	CN	0.8566	0.8528	0.8160	0.8946	0.8683	0.9064	0.9124	0.8917	0.9135	0.9243	0.9052	
LCNI	0.9057	0.9068	0.8180	0.9204	0.8821	0.9443	0.9563	0.9010	0.9485	0.9564	0.9290		
ICQD	0.8542	0.8555	0.6006	0.8414	0.8667	0.8745	0.8973	0.8959	0.8815	0.8839	0.9072		
CHA	0.8775	0.8784	0.8210	0.8848	0.8645	0.8310	0.8823	0.8990	0.8925	0.8906	0.8798		
SSR	0.9461	0.9483	0.8885	0.9353	0.9339	0.9567	0.9668	0.9326	0.9576	0.9628	0.9478		
TID 2008	AGN	0.8107	0.8086	0.5806	0.8797	0.7728	0.8386	0.8606	0.8415	0.8758	0.9229	0.8887	
	ANC	0.8029	0.8054	0.5460	0.8757	0.7793	0.8255	0.8091	0.8613	0.8931	0.9118	0.8789	
	SCN	0.8144	0.8209	0.5958	0.8698	0.7665	0.8678	0.8941	0.8468	0.8711	0.9296	0.8951	
	MN	0.7795	0.8107	0.6732	0.8683	0.7295	0.7336	0.7452	0.8534	0.8264	0.7734	0.8375	
	HFN	0.8729	0.8694	0.7318	0.9075	0.8811	0.8864	0.8945	0.9182	0.9156	0.9253	0.9225	
	IN	0.6732	0.6907	0.5345	0.8327	0.6471	0.0650	0.7235	0.8806	0.7719	0.8298	0.7919	
	QN	0.8531	0.8589	0.5857	0.7970	0.8270	0.8160	0.8800	0.8880	0.8726	0.8731	0.8500	
	GB	0.9544	0.9563	0.8559	0.9540	0.9330	0.9196	0.9600	0.9409	0.9472	0.9529	0.9501	
	DEN	0.9530	0.9582	0.7973	0.9161	0.9286	0.9433	0.9725	0.9400	0.9618	0.9693	0.9488	
	JPEG	0.9252	0.9322	0.8180	0.9168	0.9174	0.9275	0.9393	0.9385	0.9294	0.9616	0.9416	
	JP2K	0.9625	0.9700	0.9437	0.9709	0.9515	0.9707	0.9758	0.9488	0.9780	0.9848	0.9825	
	JGTE	0.8678	0.8681	0.7909	0.8585	0.8055	0.8661	0.8790	0.8503	0.8756	0.9160	0.8766	
	J2TE	0.8577	0.8606	0.7301	0.8501	0.7909	0.8394	0.8936	0.8592	0.8555	0.8942	0.8947	
	NEPN	0.7107	0.7377	0.8418	0.7619	0.5716	0.8287	0.7386	0.7274	0.7514	0.7699	0.7094	
	Block	0.8462	0.7546	0.6770	0.8324	0.1926	0.7970	0.8862	0.6258	0.8464	0.6295	0.4698	
	MS	0.7231	0.7336	0.4250	0.5096	0.3715	0.5163	0.7190	0.4178	0.6554	0.6714	0.4810	
	CTC	0.5246	0.6381	0.1713	0.8188	0.4239	0.2723	0.6691	0.5823	0.6510	0.6557	0.6348	
	CSIQ	AGWN	0.8974	0.9471	0.8431	0.9575	0.9241	0.9541	0.9440	0.9441	0.9359	0.9636	0.9647
		JPEG	0.9546	0.9634	0.9412	0.9705	0.9036	0.9615	0.9632	0.9502	0.9664	0.9618	0.9548
		JP2K	0.9606	0.9683	0.9252	0.9672	0.9480	0.9752	0.9648	0.9643	0.9704	0.9694	0.9750
AGPN		0.8922	0.9331	0.8261	0.9511	0.9084	0.9570	0.9387	0.9357	0.9370	0.9638	0.9607	
GB		0.9609	0.9711	0.9527	0.9745	0.9446	0.9602	0.9589	0.9634	0.9729	0.9679	0.9758	
GCD		0.7922	0.9526	0.4873	0.9345	0.8700	0.9207	0.9354	0.9527	0.9438	0.9504	0.9485	
LIVE	JP2K	0.9614	0.9627	0.9113	0.9696	0.9551	0.9676	0.9700	0.9323	0.9724	0.9604	0.9645	
	JPEG	0.9764	0.9815	0.9468	0.9846	0.9657	0.9764	0.9778	0.9584	0.9840	0.9761	0.9759	
	AGWN	0.9694	0.9733	0.9382	0.9858	0.9785	0.9844	0.9774	0.9799	0.9716	0.9835	0.9868	
	GB	0.9517	0.9542	0.9584	0.9728	0.9413	0.9465	0.9518	0.9066	0.9708	0.9527	0.9622	
	FF	0.9556	0.9471	0.9629	0.9650	0.9027	0.9569	0.9402	0.9237	0.9519	0.9430	0.9418	

亮度相似度组合获得 MFS 评价值. 为了提高评价的准确性和稳定性, 采用了视觉显著和视觉阈值两个策略来去除对于视觉感知不重要的图像块. 更重要的是, MFS 不仅考虑了流形结构失真同时也考虑了图像亮度变化对图像质量的影响, 这使得 MFS 具有更高的评价准确性也扩大了其对各类失真的评价能力. 与 VSI 利用先验知识对图像进行处理从而构建显著图来作为图像降质的评价依据不同, 本文的 MFS 从图像数据本身出发通过流形学习寻找数据的本质流形特征来进行图像质量评价. 从在四个公开的图像库上的实验结果和对比实验表明 MFS 可以获得比当前权威的评价算法更好的评价性能, 评价结果与主观评分具有更高的一致性.

表 9 11 种质量评价方法的时间复杂度
Table 9 Time cost of 11 IQA indices

IQA index	Time cost (ms)
SSIM	17.3
MS-SSIM	71.2
IFC	538.0
VIF	546.4
VSNR	23.9
MAD	702.3
GSM	17.7
RFSM	49.8
FSIMc	142.5
VSI	105.2
MFS	140.7

进一步的研究将考虑使用更符合人眼视觉注意机制的选块策略来提高图像质量评价的准确性, 以及在训练阶段使用更好的更有效率的算法来代替正交局部保持投影算法获取最佳映射矩阵.

References

- Ma L, Deng C W, Ngan K N, Lin W S. Recent advances and challenges of visual signal quality assessment. *China Communications*, 2013, **10**(5): 62–278
- Saha A, Wu Q M J. Utilizing image scales towards totally training free blind image quality assessment. *IEEE Transactions on Image Processing*, 2015, **24**(6): 1879–1892
- Wang Zhi-Ming. Review of no-reference image quality assessment. *Acta Automatica Sinica*, 2015, **41**(6): 1062–1079 (王志明. 无参考图像质量评价综述. *自动化学报*, 2015, **41**(6): 1062–1079)
- Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612
- Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment. In: *Proceedings of the 37th Conference Record of Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, CA, USA: IEEE, 2003. 1398–1402
- Zhang L, Zhang D, Mou X Q. RFSIM: a feature based image quality assessment metric using Riesz transforms. In: *Proceedings of the 17th International Conference on Image Process*. Hong Kong, China: IEEE, 2010. 321–324
- Zhang L, Zhang D, Mou X Q, Zhang D. FSIM: a feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 2011, **20**(8): 2378–2386
- Cheng Guang-Quan, Zhang Ji-Dong, Cheng Li-Zhi, Huang Jin-Cai, Liu Zhong. Image quality assessment based on geometric structural distortion model. *Acta Automatica Sinica*, 2011, **37**(7): 811–819 (程光权, 张继东, 成礼智, 黄金才, 刘忠. 基于几何结构失真模型的图像质量评价研究. *自动化学报*, 2011, **37**(7): 811–819)
- Chandler D M, Hemami S S. VSNR: a wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 2007, **16**(9): 2284–2298
- Larson E C, Chandler D M. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging*, 2010, **19**(1): 011006
- Sheikh H R, Bovik A C. Image information and visual quality. *IEEE Transactions on Image Processing*, 2006, **15**(2): 430–444
- Sheikh H R, Bovik A C, De Veciana G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, 2005, **14**(12): 2117–2128
- Zhang L, Shen Y, Li H Y. VSI: a visual saliency-induced index for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 2014, **23**(10): 4270–4281
- Liu A M, Lin W S, Narwaria M. Image quality assessment based on gradient similarity. *IEEE Transactions on Image Processing*, 2012, **21**(4): 1500–1512
- Seung H S, Lee D D. The manifold ways of perception. *Science*, 2000, **290**(5500): 2268–2269
- de Silva V, Tenenbaum J B. Global versus local methods in nonlinear dimensionality reduction. In: *Proceedings of the 2002 Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2002. 705–712
- Cai D, He X F, Han J W, Zhang H J. Orthogonal Laplacianfaces for face recognition. *IEEE Transactions on Image Processing*, 2006, **15**(11): 3608–3614
- Charrier C, Lebrun G, Lezoray O. Image quality assessment with manifold and machine learning. In: *Proceedings of the 2009 SPIE 7242, Image Quality and System Performance VI*. San Jose, CA: SPIE, 2009.

- 19 Taube J S. Head direction cells and the neurophysiological basis for a sense of direction. *Progress in Neurobiology*, 1998, **55**(3): 225–256
- 20 Simoncelli E P, Olshausen B A. Natural image statistics and neural representation. *Annual Review of Neuroscience*, 2001, **24**: 1193–1216
- 21 Engelke U, Kaprykowsky H, Zepernick H, Ndjiki-Nya P. Visual attention in quality assessment. *IEEE Signal Processing Magazine*, 2011, **28**(6): 50–59
- 22 Moorthy A K, Bovik A C. Visual importance pooling for image quality assessment. *IEEE Journal of Selected Topics in Signal Processing*, 2009, **3**(2): 193–201
- 23 Zhang L, Gu Z Y, Li H Y. SDSP: a novel saliency detection method by combining simple priors. In: Proceedings of the 20th IEEE International Conference on Image Processing. Melbourne, VIC: IEEE, 2013. 171–175
- 24 Sheikh H R, Sabir M F, Bovik A C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 2006, **15**(11): 3440–3451
- 25 Ponomarenko N N, Lukin V V, Zelensky A, Egiazarian K, Carli M, Battisti F. TID2008 — a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 2009, **10**: 30–45
- 26 Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F, Kuo C C J. Image database TID2013: peculiarities, results and perspectives. *Signal Processing: Image Communication*, 2015, **30**: 57–77
- 27 VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment [Online], available: <http://www.vqeg.org/>, November 3, 2015



王朝云 宁波大学硕士研究生. 主要研究方向为图像及视频质量评价.

E-mail: wcy_417@126.com

(**WANG Chao-Yun** Master student at Ningbo University. His research interest covers image and video quality assessment.)



蒋刚毅 宁波大学教授. 主要研究方向为计算机图像处理, 图像与视频信号编码与传输. 本文通信作者.

E-mail: jianggangyi@126.com

(**JIANG Gang-Yi** Professor at Ningbo University. His research interest covers computer image processing, image and video signal encoding and transmission. Corresponding author of this paper.)



郁梅 宁波大学教授. 主要研究方向为计算机图像处理, 图像与视频信号编码与传输. E-mail: yumei2@126.com

(**YU Mei** Professor at Ningbo University. Her research interest covers computer image processing, image and video signal encoding and transmission.)



陈芬 宁波大学副教授. 主要研究方向为光通信技术, 数字信号处理技术.

E-mail: chenfen@126.com

(**CHEN Fen** Associate professor at Ningbo University. Her research interest covers optical communication technology, and digital signal processing technology.)