

基于结构信息的图像质量评价模型^①

顾丽旻, 王晓红

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: 针对客观数字图像质量评价方法结构相似度(SSIM)算法对严重失真类型的评价与主观评价(MOS)值之间存在着明显差异的问题, 对几种将 SSIM 算法与人眼视觉系统(HVS)特性相结合的算法基于图像边缘的 ESSIM 算法、基于空间域频域敏感的 HSSIM 算法, 以及基于信息内容权值的 IWSSIM 算法进行了比较。从实验数据可知, ESSIM, HSSIM 算法的准确性有不同程度的提高, 而 IWSSIM 模型在比较主客观值相关的单调性方面更为稳定。实验结果为今后结构相似度算法的改进提供了依据, 指明研究方向。

关键词: 结构信息; 人眼视觉特征; 图像质量评价; 对比度敏感度特性; 互信息量

Structural Information-Based Image Quality Assessment Models

GU Li-Min, WANG Xiao-Hong

(Dept. of Photoelectric Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The objective assessment algorithm Structural Similarity(SSIM) of image quality was not consistent with the Mean Opinion Score (MOS) in particular types of distortion, such as seriously blurred type, etc. In this paper, three models Edeg-based SSIM, HVS-based SSIM, and IW-SSIM which had considered the characteristics both of HVS and structural information were compared in evaluating the seriously Gaussian blurred images. The experimental datas showed that, the IW-SSIM's monotony performance was more stable than other models in the relationship between the subjective and objective values, as well as the growing accuracy of the ESSIM and HSSIM. These experimental results are benefit for improving the SSIM approach.

Key words: structural information; human visual system; image quality assessment; contrast sensitivity; mutual information

图像在传输和处理时, 经过编码压缩, 采样, 传输, 解压等过程, 不可避免地会产生失真现象, 影响了图像所传递的信息的真实性^[1]。为了得到较好的图像质量, 可以通过对传输后的图像进行质量评价, 来优化传输系统, 进而得到理想的效果。目前图像的质量评价 IQA(Image Quality Assessment)可分为主观质量评价与客观质量评价两种。

主观质量评价是在一个受控环境(指观测环境等)下, 让观察者按实验人员设计好的观看方式对图像质量进行打分, 再对所得的分数进行统计计算, 得到平均主观得分 MOS(Mean Opinion Score), 通常使用 5 分

制来表示图像质量的等级^[1]。由于主观质量评价的测试环境需要满足特定条件, 观察者由非经验人员和专家组成, 图像出现需按特定顺序, 需要耗费大量时间与人力, 实施起来非常复杂。客观质量评价是通过计算机模型给出图像质量的量化标准, 根据过程中是否使用原始图像进行比较, 客观质量评价又可分为全参考评价、部分参考评价和无参考评价三种方法^[2]。

早期的客观方法有均方误差 MSE (Mean Square Error)和峰值信噪比 PSNR(Peak Signal to Noise Ratio)等, 这些方法是利用两幅图像对应像素间的差异来评价图像质量, 其程序实现简单, 操作方便, 能够应用

① 基金项目:上海市科委基金(09220502700)

收稿时间:2011-05-23;收到修改稿时间:2011-06-23

于大部分的图像失真评价，但有两点不足：其一由于只比较了像素点间的绝对误差，没有考虑像素点间的相关性，对于特定的失真类型的评价与主观评价结果存在着较明显差异^[1]；其二，对于不同原图像的不同失真类型的评价无法进行横向比较^[3]。因此研究基于人眼视觉系统 HVS 特征的评价模型并给出一个量化图像质量的标准显得尤为重要。近年来，对图像质量的评价受到了越来越多学者和相关组织的重视与研究，其中，视频质量专家组 VQEG(Video Quality Experts Group)专门从事视频质量评价方法的研究，目前已发布了电视图像主观质量标准 BT-500-11 和客观评价模型的最终报告^[2,4]。

从研究现状可知，客观质量评价的模型可分为两类，基于感知的评价方法和基于结构信息的评价方法^[5]。基于感知的评价方法是根据 HVS 特征对图像进行评价的，它认为只有参考图像的失真影响到大部分人的主观评价才在评价中予以考虑；基于结构信息的评价方法建立在图像失真后其结构信息也随之发生质降的基础之上。而图像质量评价的研究趋势则是将以上两者相结合，以期得到更符合主观质量评价的客观量化标准^[5]。本文所介绍的方法都是针对灰度图像进行的研究。

1 传统SSIM算法

在 2002 年，Wang Zhou 等人提出了 SSIM 算法，这一算法建立在假设 HVS 对于从视野中获取结构信息具有高度适应性的基础上，所以测量结构信息的改变可以得到人眼感知到的图像失真的近似值^[3]。该算法首先计算两幅图像的亮度比较函数 $l(x,y)$ ，对比度比较函数 $c(x,y)$ ，以及结构比较函数 $s(x,y)$ ，然后将三部分函数合并(Pooling)，最后得到一个全局相似性测量 SSIM 指数。具体过程如下：

假设 X, Y 分别代表原参考图像和待测图像， $X=\{x_i|i=1,2,\dots,N\}, Y=\{y_i|i=1,2,\dots,N\}$ 。X 的平均亮度和标准方差，以及 X 和 Y 的协方差分别表示为 $\mu_x, \sigma_x, \sigma_{xy}$ ，表示，则亮度对比函数，对比度比较函数，结构比较函数分别表示为：

$$l(x, y) = (2\mu_x\mu_y + C_1) / (\mu_x^2 + \mu_y^2 + C_1) \tag{1}$$

$$c(x, y) = (2\sigma_x\sigma_y + C_2) / (\sigma_x^2 + \sigma_y^2 + C_2) \tag{2}$$

$$s(x, y) = (\sigma_{xy} + C_3) / (\sigma_x\sigma_y + C_3) \tag{3}$$

其中，系数 $C_1=(K_1L)^2, C_2=(K_2L)^2, K_1, K_2$ 都是非常

小的常数，L 是图像像素的最高值。X 和 Y 的结构相似度指数可定义为：

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \tag{4}$$

其中 α, β, γ 都大于零，用于调节三个对比度函数之间的相关性。在实际应用中，用平均 SSIM 来评价图像的全局质量：

$$MSSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \tag{5}$$

x_j, y_j 是第 j 个局部窗口中的图像内容；M 是使用滑动窗口后的局部窗口的数量。

2 几种基于HVS特征的SSIM改进算法

自 Zhou Wang 等提出 SSIM 算法以来，由于该算法复杂度低，应用性较强，且评价结果与主观质量评价具有很好的一致性，因此得到推广与应用，但 SSIM 算法对于模糊等失真类型的评价准确性很差，一些学者在此基础上进行了改进。

2.1 基于边缘的结构相似度算法^[6]

许多研究人员发现，HVS 对图像的边缘和内容信息非常敏感，这可以理解为边缘与内容信息是人眼所获得的图像信息中最重要的部分，基于这一思想，提出了基于边缘的结构相似度算法，将基于边缘的结构对比函数 $e(x,y)$ 代替原算法中的结构比较函数 $s(x,y)$ 。

使用索贝尔算子(Sobel Operator)进行滤波提取图像边缘，图像中每个像素 pi,j 的边缘向量可定义 $D=\{dx_{i,j}, dy_{i,j}\}$ ， $dx_{i,j}, dy_{i,j}$ 分别表示滤波后的水平方向向量与垂直方向向量。边缘向量的幅值，像素的边缘方向角度可定义为：

$$Amp_{i,j} = |dx_{i,j}| + |dy_{i,j}| \tag{6}$$

$$Ang_{i,j} = \frac{180^\circ}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right) \tag{7}$$

一幅图像中所有像素的边缘向量构成了图像的边缘示意图。边缘方向直方图可用于比较失真图像块与原参考图像块中的边缘信息。将连续方向()分成 8 个离散方向，如图 1 所示：

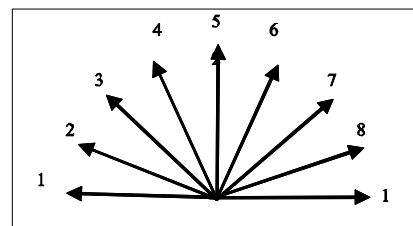


图 1 8 个离散方向

对于每个图像块的边缘方向直方图可通过下列方式获得:

- 1) 计算每个像素的边缘幅值和方向;
- 2) 将每个像素的方向量化成 8 个离散方向的其中一个;
- 3) 将一个宏块中统一方向上的所有像素的边缘幅值进行相加。

设 D_x 和 D_y 分别表示原参考图像和失真图像的块边缘方向向量, 通过计算它们的相关系数获得边缘对比函数:

$$e(x, y) = (\sigma_{xy}' + C_3) / (\sigma_x' \sigma_y' + C_3) \tag{8}$$

其中, σ_x' 和 σ_y' 分别表示向量 D_x 和 D_y 的标准偏差, σ_{xy}' 是 D_x 和 D_y 的协方差, C_3 为一个很小的常数。基于边缘的结构相似度和图像全局结构相似度分别表示为:

$$ESSIM(x, y) = l(x, y) \cdot c(x, y) \cdot e(x, y) \tag{9}$$

$$MESSIM(x, y) = \frac{1}{M} \sum_{j=1}^M ESSIM(x_j, y_j) \tag{10}$$

2.2 基于时域频域敏感的结构相似度算法^[7]

通过对 HVS 心理学的研究, 人们发现人眼对空间高频的敏感性较差, 因此视觉的空间频域响应具有低通特性, 即满足对比度敏感度函数 CSF(Contrast Sensitivity Function)^[8]。

算法考虑了这一重要特征, 提出了基于频域敏感度和空间敏感度因子。频域部分能有效反应图像的纹理特征, 因此可以对图像的不同频域部分分配不同权值。对图像采用离散余弦变换 DCT(Discrete Cosine Transformation)来获得频域信息。

定义频域结构对比函数:

$$freq(x, y) = (2\sigma_{Fxy} + C_3) / (\sigma_{F_x}^2 + \sigma_{F_y}^2 + C_3) \tag{11}$$

其中, σ_{F_x} 和 σ_{F_y} 分别是 x, y 在频域的考虑权重的标准偏差, σ_{Fxy} 是 x, y 的权重协方差。

研究还发现, 光敏感细胞主要集中在视网膜的黄斑区, 因此中间黄斑区的分辨率比视网膜其它区域的高。当观察一幅图像时, 首先注意到它的中心部分, 然后才逐渐关注四周, 这意味着图像中信息的重要性是从中间到边缘递减的, 同时也符合我们观察图像的习惯。对每个宏块加入空间影响权值, 设一个宏块的中心位置为 (x, y) , 影响权值被定义为:

$$w_{spatial(x, y)} = 1 - C \times \sqrt{(x - x_c)^2 + (y - y_c)^2} / d_{max} \tag{12}$$

其中, x_c, y_c 分别代表图像中心位置, d_{max} 是从每个

像素到中心位置的最大距离, $C \in (0, 1)$, 是由空间关注敏感度所决定的参数。将空间影响权值加入到每个宏块中, $h(x, y)$ 表示基于 HVS 的结构比较函数:

$$h(x, y) = w_{spatial(x, y)} \cdot freq(x, y) \tag{13}$$

所以, 基于 HVS 的结构相似性指数和整幅图像的质量可表示为:

$$HSSIM = f(l(x, y), c(x, y), h(x, y)) \tag{14}$$

$$MHSSIM(x, y) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N HSSIM(x_i, y_j) \tag{15}$$

2.3 基于信息内容权值的结构相似度算法^[9]

Zhou Wang 等发现, 现有的许多 IQA 算法都具有类似的两层结构: 1) 局部质量的测量, 2) 将(1)中所得值合并。很多算法对 1) 部分的研究都很深入, 但对于 2) 部分的合并, 很多算法都采用了通用的方法(如 Minkowski 合并等)进行处理, 缺乏一定理论依据与可靠的计算原则。该算法提出了合并的最佳感知权值应与局部质量内容成比例的设想, 将局部信息内容 IW 的测量应用到 IQA 权值中, 其中局部质量内容可以用先进的自然图像统计模型以比特位单位进行测量。信息内容的获取如图 2 所示。

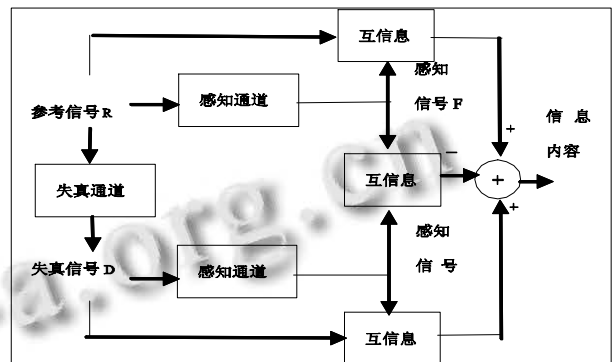


图 2 信息内容计算流程图

图像信息内容的获取需要性能优秀的统计模型, Wang 采用了高斯尺度混合(GSM)模型。通过对信息理论化 IQA 方法的进一步研究发现, 图像中所包含的信息量并不等同于 HVS 所感知到的信息量, 而经视觉感知信道前后的互信息量恰恰有效的测量这一点。 IW 可定义为:

$$w = I(R; E) + I(D; F) - I(E; F) \tag{16}$$

$I(R; E)$, $I(D; E)$ 分别表示视觉系统感知到的原参考图像和失真图像的互信息量; $I(E; F)$ 为原参考与失真图像经感知信道后的互信息量。在实际应用中, 首先对

原参考和失真图像采用一个 5-尺度的拉普拉斯金字塔分解, 然后再对每个尺度使用滑动窗口计算其 IW, 得到每个尺度的 IW map 图。将 IW 应用到 MSSIM 算法中, 设 $x_{j,i}, y_{j,i}$ 表示第 j 个尺度下第 i 宏块的像素, N_j 是该尺度下测量窗口的数量, 第 j 尺度下 SSIM 和 IW-SSIM 分别为:

$$SSIM_j = \frac{1}{N_j} \sum_i l(x_{j,i}, y_{j,i})c(x_{j,i}, y_{j,i})s(x_{j,i}, y_{j,i}) \quad (17)$$

$$IW_SSIM_j = \sum_i \omega_{j,i} c(x_{j,i}, y_{j,i})s(x_{j,i}, y_{j,i}) / \sum_i \omega_{j,i} \quad (18)$$

当 $j=M$ 时, 全局的 IWSSIM 可定义为:

$$IW_SSIM = \prod_{j=1}^M (IW_SSIM_j)^{\beta_j} \quad (19)$$

3 结果及分析

本文中讨论的三种模型的性能测试数据均采用美国 TEXAS 大学图像视频工程实验室所提供的 Live IQA 数据库(第二版)^[10]。该库是由 29 幅原始图像由 5 种失真类型经不同程度处理所组成: jpeg2000 压缩, jpeg 压缩, 高斯模糊, 白噪声, 瑞利衰落。

根据最终报告^[5], VQEG 提出主客观数据间存在非线性关系, 运用非线性逻辑回归函数对客观值进行处理, 选用公式(20), 式中系数经回归拟合后获得。

$$b1/(1+\exp(-b2 \times (x-b3))) \quad (20)$$

基于 VQEG 提出的客观模型评价准则^[1], 本文选用 4 个参数来比较算法的性能: 相关系数(CC)表征主客观数值的相关性, 值越大相关性越好; 均方误差(RMSE), 平均绝对误差(MAE), 这两个值越小说明模型误差越小; Spearman 秩相关系数, 衡量模型单调性, 值越大越好。

将文中的三个模型与 PSNR, SSIM 进行比较, 表 1 给出了模型对高斯模糊失真类型的表现, 另外, 对 Live release2 的所有图像(除原参考图像)进行了测试比较, 5 个模型的表现表 2 中给出。图 3 给出各个模型对所有图像的客观评价与 DMOS 值的对比散点图。从表 1 可以发现, 改进后的 ESSIM, HSSIM 模型由于考虑了 HVS 特征在评价模糊失真类型图像时准确性都比 PSNR, SSIM 有所提高, 但二者单调性却不如 SSIM, IWSSIM; IW-SSIM 另辟蹊径, 从信息论角度出发, 在合并处理时加入信息权值进行处理, 具有可靠的理论依据, 单调性比其它模型都更稳定, 对模糊类型图像的评价的准确性较高, 但对于全体图像的

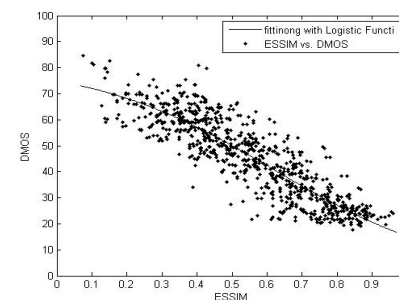
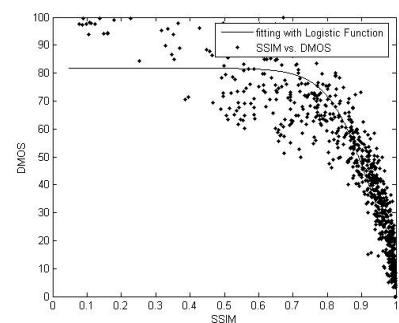
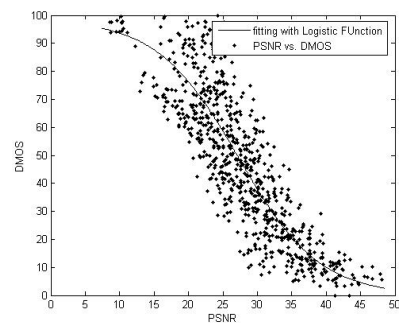
评价质量却比 HSSIM 稍差, 表 2 系数的表述与图 3 散点图所观察的情况一致。

表 1 高斯模糊类型的 IQA 模型性能比较

| Model | CC | MAE | RMSE | SROCC |
|--------|--------|--------|---------|--------|
| PSNR | 0.7834 | 8.1096 | 10.4793 | 0.7823 |
| SSIM | 0.9366 | 5.6274 | 6.8491 | 0.9317 |
| ESSIM | 0.9409 | 4.7726 | 6.0988 | 0.9184 |
| HSSIM | 0.9381 | 5.7483 | 6.8364 | 0.9279 |
| IWSSIM | 0.9468 | 4.9486 | 6.397 | 0.9485 |

表 2 LIVE Release2 图像的 IQA 模型性能比较

| Model | CC | MAE | RMSE | SROCC |
|--------|--------|--------|---------|--------|
| PSNR | 0.8712 | 10.185 | 12.4931 | 0.8213 |
| SSIM | 0.9226 | 7.0943 | 9.3524 | 0.9206 |
| ESSIM | 0.9042 | 6.2776 | 8.1278 | 0.9032 |
| HSSIM | 0.9291 | 4.8451 | 6.674 | 0.9296 |
| IWSSIM | 0.9257 | 4.8955 | 6.0322 | 0.9385 |



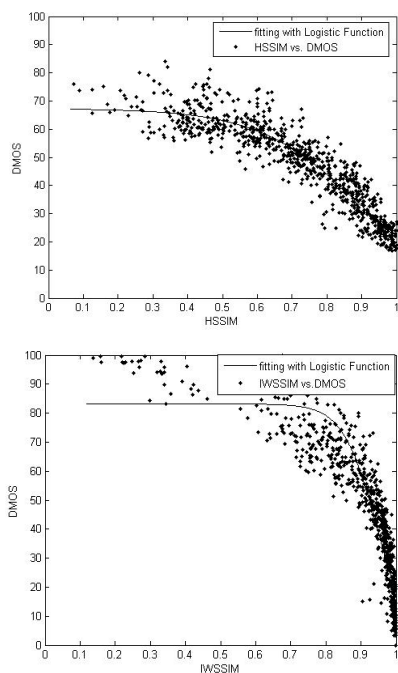


图 3 模型与 DMOS 对比散点图

4 结语

本文在深入研究 SSIM 算法基础上, 比较了基于视觉特征的 SSIM 改进模型。经比较发现, 将视觉特征引入到 SSIM 评价模型中, 对于特定失真类型的表现有所提高; 将信息理论结合到 SSIM 合并过程, 模型的单调性性能更加稳定。由此可见, 在今后的研究中, 将采用 HVS 特征对局部质量测量与基于信息内容的合并方法相结合, 可以得到更加符合 HVS 特征的客观质量评价模型。

参考文献

- 1 Sheikh HR, Bovik AC. Image information and visual quality. Acoustics, Speech, and Signal Processing, IEEE international conference, 2004, 3(3): 1-26.
- 2 VQEG. Final Report from the video quality experts group on the validation of objective models of video quality assessment. VQEG, 2008. <http://www.vqeg.org/>.
- 3 Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. IEEE Transaction on Image Processing magazine, 2004, 13(4): 600-612.
- 4 ITU-R. Recommendation BT. 500-11 Methodology for the subjective assessment of the quality of television pictures. ITU-R. 2002.
- 5 蒋刚毅, 黄大江, 王旭, 等. 图像质量评价方法研究进展. 电子与信息学报, 2010, 32(1): 219-226.
- 6 Chen GH, Yang CL, et al. Edge-based structural similarity for image quality assessment. ICASSP. IEEE international conference, 2006, 2(5): 933-936.
- 7 Wang B, Wang ZB, et al. HVS-based structural similarity for image quality assessment. IEEE Signal Processing Magazine, 2008, 4(9): 1194-1197.
- 8 Wang Z, Sheikh HR. Objective image quality assessment. Furht B, Marqure O. The handbook of Video Database: Design and Applications. CRC Press, 2003: 1041-1078.
- 9 Wang Z, Li Q. Information content weighting for perceptual image quality assessment. IEEE Transaction on Image Processing magazine. 2010, 20(5): 1185-1197.
- 10 Sheikh HR, Wang Z, Bovik AC, et al. Image and video quality assessment research at LIVE Database Release 2. (2006-05-10)/2007(2007-06-30). <http://live.ece.utexas.edu/research/quality>