

边缘加权的结构相似性测度

吕 鹏, 张建秋

(复旦大学电子工程系, 上海 200433)

摘 要: 针对结构相似性测度(SSIM)不能较好地客观评价图像模糊与强高斯噪声失真问题, 提出一种边缘加权的结构相似性测度(EWSSIM), 以符合人眼视觉系统(HVS)特性。EWSSIM 将原始图像和失真图像的整体轮廓信息与局部纹理细节信息加权, 更充分地描述图像的结构相似度。通过 LIVE 图库的仿真结果表明, 与 SSIME 相比, WSSIM 能够更好地评价图像模糊与强高斯噪声失真, 且在各类失真图像的评价一致性上优于 SSIM。

关键词: 图像质量评估; 结构相似性测度; 边缘检测; 边缘加权结构相似性测度

Edge-weighted Structural Similarity Index

LV Peng, ZHANG Jian-qiu

(Dept. of Electronic Engineering, Fudan University, Shanghai 200433, China)

【Abstract】 This paper proposes an Edge-weighted Structural Similarity Index(EWSSIM), which can match well with Human Vision System(HVS). Structural Similarity Image quality assessment(SSIM) does not evaluate highly blurred and Gaussian white noise distorted images well. EWSSIM assigns different weights to contour correlation and local texture correlation of the original image and distorted image, which can represent structural similarity better than SSIM. Experimental results of LIVE image database indicate that the proposed index outperforms SSIM in blurred and Gaussian white noise distorted images and also gives a better coherent evaluation for all kinds of distortions in LIVE database.

【Key words】 image quality assessment; Structural Similarity index(SSIM); edge detection; Edge-weighted Structural Similarity Index(EWSSIM)

DOI: 10.3969/j.issn.1000-3428.2011.14.076

1 概述

数字图像在获取、传输、压缩和存储等过程中受到不同类型和程度的失真影响。人眼作为图像接收终端, 采用 MOS(Mean Opinion Score)或 DMOS(Differential MOS)的主观评价方法是图像质量评估最合理的方法。主观的评价方法由于其不易操作性, 需要大量的人力与时间。因此, 有必要研究客观图像的质量指标, 即研究一套量化的方法, 用于自动评估压缩、传输等方法对图像质量的影响, 并为进一步改善这些方法提供理论的指导。

客观的图像质量评估根据有无原始的无失真图像分为有参考图像质量评估、无参考图像质量评估和部分参考图像质量评估, 本文主要研究有参考图像的质量评估。

在过去几十年, 基于误差感知的最小均方误差(Mean Square Error, MSE)或峰值信噪比(Peak Signal to Noise Ratio, PSNR)的图像质量评估方法, 由于其简单的计算方法和清晰的物理意义已经被广泛使用, 但研究表明这些基于误差感知的方法并不符合人眼视觉系统(HVS)的特性^[1]。因此, 文献[2]提出了结构相似性测度(Structural Similarity index, SSIM)。SSIM 基于 HVS 中结构信息的假设, 从而比 PSNR 更好地符合 HVS 特性。

此外, SSIM 的计算方法简单, 并且可将其应用到图像的变换域^[3], 取得了很好的评估结果。然而, 随着对 SSIM 的深入研究, 发现 SSIM 算法存在一些问题, 如它不能较好地评估模糊失真类与强高斯噪声图像的质量^[4]。

本文提出边缘加权的结构相似性, 利用图像边缘这一最重要的结构信息改进 SSIM 测度, 从而更好地评价图像模糊和强高斯噪声失真。

2 结构相似性测度(SSIM)及其存在的问题

2.1 SSIM 简介

SSIM 指标包含了 2 幅图像之间的亮度差异 $l(x, y)$, 对比度差异 $c(x, y)$ 和结构相似性差异 $s(x, y)$, 有:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (1)$$

3 个部分相乘得到式(2):

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (2)$$

其中, $\alpha = \beta = \gamma = 1$; $C_3 = \frac{C_2}{2}$ 。最后得到式(3):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

全局图像质量 MSSIM 是由计算 SSIM 在所有窗下的平均值得到。对整幅图像的评分, 通过平均各个子块的 SSIM 得到:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{i=1}^M SSIM(x_j, y_j) \quad (4)$$

MSSIM 的分值越高, 表示原图像和降质图像越相似, 降质图像的质量越高。

虽然 SSIM 计算方法简单, 同时性能也优于 MSE 等算法, 但是随着对 SSIM 的深入研究, 发现 SSIM 算法存在一些问题。

2.2 SSIM 算法中存在的问题

许多研究者发现, 人眼视觉系统对于图像的边缘和轮廓

基金项目: 国家自然科学基金资助项目(60872059)

作者简介: 吕 鹏(1986—), 男, 硕士研究生, 主研方向: 图像处理, 质量评估; 张建秋, 教授、博士生导师、IEEE 高级会员

收稿日期: 2010-12-30 **E-mail:** 082021029@fudan.edu.cn

结构信息是最为敏感和重视的, 边缘和轮廓的结构信息很有可能是图像结构信息最重要的部分。但在 SSIM 中结构部分的评价中, $s(x, y)$ 实质上求取的仅仅是原图像子块和降质图像子块之间像素值的相关系数, 只能反映图像子块纹理的相关性, 缺乏对边缘结构的相关性, 而对于轮廓结构信息的评价也没有在图像子块的评价中体现出来。对于图像模糊失真, SSIM 不能很好地衡量出原图像和降质图像之间边缘与轮廓结构信息的差异。当模糊程度较高时, 图像的边缘与轮廓信息丢失很多, 但是图像子块的像素依然有很高的相关性, SSIM 结构评价部分仍有较高的值, 因此, SSIM 对模糊图像的评估较高。

对于高斯噪声失真图像, 不同位置的噪声点对于图像质量的影响是不同的, 而这一点在 SSIM 的评价中并没有体现。人眼视觉系统对于区域的强噪声点比较敏感, 而对于边缘位置的强噪声点不敏感。在强噪声失真的图像中, 子块像素的相关性大大下降, 于是 SSIM 对其的评价很低。但是很多时候高斯噪声失真图像的轮廓信息依然可以被人眼视觉系统很好地提取。

SSIM 在对高斯模糊失真图像的评价的一致性上也存在问题, 模糊失真程度高的图像其 SSIM 评价却较高, 因此, SSIM 不能很好地评价图像的模糊失真。可以发现当高斯噪声大到一定程度时, SSIM 的数值会极低, 分析 SSIM 算法, 从式(1)的协方差公式可以看出: 当 $(x_i - \mu_x)$ 和 $(y_i - \mu_y)$ 异号时, σ_{xy} 会迅速变小, 甚至为负数, 而强高斯噪声失真图像在噪声点处其 $(x_i - \mu_x)(y_i - \mu_y)$ 多为负值, 也导致了 SSIM 对此类图像评分的整体过低。

针对这一问题, 文献[4]提出了一种提取每一个图像子块的边缘统计直方图, 然后计算无失真图像与失真图像对应子块的边缘统计直方图相似度。利用这一相似测度还代替 SSIM 中结构部分的测度评价。但这一方法边缘检测算法的阈值不能设置过低, 高度模糊失真图像通过边缘检测无法检测到边缘, 评价结果很低, 但是此时的图像纹理结构还是有较高的相似度的。

3 边缘加权的结构相似性测度(EWSSIM)

图像的轮廓信息和边缘信息是最重要的结构信息。而图像的轮廓信息应该是从图像的整体考虑的, 这样可以区别图像的轮廓与图像中纹理细节。

基于这种想法, 提出一种改进 SSIM 的算法——边缘加权的相似性测度(Edge-weighted Structural Similarity Index, EWSSIM)。在 SSIM 中评价结构相似度的 $s(x, y)$ 部分实际评价的是每个图像子块的纹理结构的相似度, 没有考虑图像的整体轮廓的相似程度。将 2 幅图像的整体轮廓信息的相似度与局部纹理结构相似度加权, 可以更好地捕获图像的整体轮廓与局部的纹理信息, 从而获得更优的结构评价系数 $w(x, y)$ 。

对 2 幅图像整体轮廓的提取采用边缘检测算法中性能较好的 canny 算子, 获得原始图像的边缘图 $E1$ 与失真图像的边缘图像 $E2$ 。通过计算 2 幅边缘图像的相关系数 $\text{corr}(E1, E2)$ 定义 2 幅图像的整体轮廓相关指数。

而对于图像的局部纹理结构的相似度仍采用 SSIM 中的 $s(x, y)$ 计算 2 个子块的相关度。最后获得的结构评价因子为:

$$w(x, y) = \alpha \cdot \text{corr}(E1, E2) + \beta \cdot s(x, y) \quad (5)$$

其中, $\alpha + \beta = 1$, α 、 β 为 2 个部分结构评价因子的加权系

数, 本文采用 $\alpha = 2\beta$ 。

$w(x, y)$ 可以更好地对图像模糊与强高斯噪声失真做出评估。因为在高度高斯模糊失真的情况下, 降质图像的整体结构信息损失很大, 体现在 $\text{corr}(E1, E2)$ 部分的值很小, 而对于降质图像与原始图像的纹理细节的相关度体现在 $s(x, y)$ 部分。在高度高斯噪声失真的情况下, 降质图像与原始图像的纹理细节受到很大影响, 于是 $s(x, y)$ 的值很小, 但是降质图像仍保持了一些整体轮廓信息, 包含在 $\text{corr}(E1, E2)$ 中, 而且不同高斯噪声点所在的位置带来的影响也是不一样的, 人眼视觉系统对于平坦区域的强噪声点比较敏感, 而对于边缘位置的强噪声点不敏感, 通过边缘检测, 将平坦区域的噪声点提取至边缘图像中。

于是定义加权边缘整体相似性测度(EWSSIM)为:

$$\text{EWSSIM}(X, Y) = \frac{1}{M} \sum_{i=1}^M l(x_i, y_i) c(x_i, y_i) w(x_i, y_i) \quad (6)$$

其中, $l(x_i, y_i)$ 、 $c(x_i, y_i)$ 与 SSIM 中定义相同; $w(x_i, y_i)$ 定义如式(6)。

4 仿真结果

大部分图像处理系统的最终用户是人类视觉系统(HVS), 所以评价一种图像质量评价算法的标准是依据图像质量评价算法的评价结果与人类主观视觉评分结果的非线性拟合程度。目前, 国际通用人类主观视觉评分标准是 DMOS, 它是人类视觉主观评分均值(MOS)和满分(100)的差值, DMOS 越小代表图像质量越好, DMOS 的取值范围是[0,100]。本文的实验测试图像选用美国 TEXAS 大学图像和视频工程实验室提供的图像质量估计数据库 databaserelease2^[5], 每个图库中有不同质量的失真图像及并附有每幅图像对应的 DMOS。

表 1 是对评价结果做非线性回归下相关系数, 这一指标评价了预测的精确性。大的相关系数表示更好的精确程度。从表 1 中可以看出, EWSSIM 在对于不同失真类型图像的评价上均有提高, 尤其对于高斯模糊失真图像与高斯白噪声失真图像, 因为图像整体轮廓信息的加入更好地评价了 2 幅图像结构信息的相似程度。EWSSIM 的优点在于不同失真类型评价的一致性的大幅提高, 使得 EWSSIM 的通用性增强。

表 1 EWSSIM 和 SSIM 的相关系数性能比较

评估方法	失真类型				总体相关系数	
	JPEG2000	JPEG	White noise	Gaussian Blur		
EWSSIM	0.964	0.965	0.964	0.969	0.961	0.951
SSIM	0.965	0.964	0.942	0.929	0.954	0.936

表 2 是 Spearman 等级相关系数(Spearman Rank Order Correlation Coefficients, SROCC), 表征了评价方法的单调性。从表 2 可以看出, EWSSIM 在评价单调性方面也是优于 SSIM 的。

表 2 EWSSIM 与 SSIM SROCC 评价单调性性能比较

评估方法	失真类型				总体 SROCC	
	JPEG2000	JPEG	White noise	Gaussian Blur		
EWSSIM	0.972	0.959	0.979	0.978	0.976	0.945
SSIM	0.971	0.958	0.978	0.938	0.961	0.925

图 1 是 SSIM 和 EWSSIM 评价图像的 DMOS 散点图。从图 1、表 1 和表 2 的结果可以看出, 无论是数据样本的分散度还是从与主客观评分的相关性来分析, EWSSIM 均优于 SSIM。

(下转第 230 页)