

基于局部结构张量的 SSIM 图像质量评价

罗学刚^{1,2}, 吕俊瑞², 王华军¹

(1 成都理工大学 地球探测与信息技术教育部重点实验室, 四川 成都 610059;

2 攀枝花学院 数学与计算机学院, 四川 攀枝花 617000)

摘 要: 针对结构相似度(SSIM)评价模糊图像存在的局限性,提出一种基于局部结构张量的结构相似性图像质量评价方法(ST-SSIM).方法首先利用结构张量提取图像几何结构强度和方向角度信息,然后用局部结构方向相似度比较函数和结构强度相似度比较函数替换 SSIM 中原有的结构比较函数,最后在图像质量映射图上以参考图像的结构强度函数加权计算得出最终的图像质量评价.实验结果表明,由于 LST-SSIM 更加全面考虑到对于人眼感知结构强度和方向的重要性,因此评价结果比 SSIM 更加符合人眼视觉感知特性.

关键词: 图像质量评价;结构信息;结构相似度;局部结构张量;人眼视觉系统

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2014)09-0109-04

Structural Similarity for Image Quality Assessment Based on Local Structure Tensor

LUO Xue-gang^{1,2}, LV Jun-rui², WANG Hua-jun¹

(1 Key Lab of Earth Exploration & Information Techniques of Ministry of Education,

Chengdu University of Technology, Chengdu 610059, China;

2 School of Mathematics and Computer Science, Panzhihua University, Panzhihua 617000, China)

Abstract: Aiming at the drawback of SSIM on blurry images, an improved Structural Similarity method for image quality assessment based on structure tensor is presented. Firstly, geometrical structure magnitude and orientation angle are extracted by local structure tensor. Secondly, a comparison function which is a structural orientation angle function and a structural magnitude function replaces the structure comparison function in the original SSIM. Finally, we use structure magnitude as a weighting function to derive the local quality map to obtain the final image quality evaluation value. The results show that since the importance of structure magnitude and orientation for the human eyes is more comprehensive consideration in perceiving structural information, the ST-SSIM is more accordant with the characteristics of human visual perception than SSIM.

Key words: image quality assessment; structural information; structural similarity; structure tensor; human visual system

1 引言

数字图像在采集、处理、压缩、存储、传输过程中,不可避免地遭受各种类型失真干扰使得其图像质量存在不同程度下降,无法获得图像真实数据,给后期处理与应用带来极大影响.通过对失真图像合

理的质量评价将有利对图像的应用价值做出初步评估,对各种图像处理应用具有重要意义.

为了得到符合人眼视觉的评价方法,研究者们根据人眼视觉系统的不同、视觉信号敏感程度不同(比如亮度、对比度、频谱),提出大量的基于视觉特性的评价方法.其中最具有代表性的是 VSNR(Visual

收稿日期: 2013-10-30; 修回日期: 2013-12-26

基金项目: 国家自然科学基金青年科学基金项目(61202195/F020502)

Signal to Noise Ratio)^[1]、VIF (Visual Information Fidelity)^[2]、SSIM 方法等。虽然这些方法与主观评价有很好的 consistency, 但方法仍不够理想。

本文全面考虑到人眼对于感知结构强度和方向信息的重要性, 提出一种基于局部结构张量的结构相似性图像质量评价方法 (LST-SSIM)。实验表明, 该方法对纹理丰富图像的模糊失真类型和噪声失真类型都能达到较好的图像质量评价结果, 比 SSIM 更加符合人眼视觉感知特性^[3-5]。

2 SSIM 图像质量评价方法

假设 $x = \{x_i | i = 1, 2, \dots, N\}$ 和 $y = \{y_i | i = 1, 2, \dots, N\}$ 分别表示参考图像和失真图像, SSIM 评价系统定义如下:

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (1)$$

式中, $l(x, y)$ 表示亮度比较函数, $c(x, y)$ 表示对比度比较函数, $s(x, y)$ 表示结构相似度比较函数, α, β, γ 分别为亮度比较函数、对比度比较函数以及结构相似度比较函数三个分量的相对重要性。

当 $\alpha = \beta = \gamma = 1$ 时, SSIM 简化为

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + A_1)(2\sigma_{xy} + A_2)}{(\mu_x^2 + \mu_y^2 + A_1)(\sigma_x^2 + \sigma_y^2 + A_2)},$$

式中, $\mu_x, \sigma_x, \sigma_{xy}$ 分别表示 x 的均值、方差以及 x 与 y 的协方差, μ_x, σ_x 分别为亮度和对比度的评估指标, σ_{xy} 为 x 与 y 对应结构变化评估指标, A_1, A_2 为较小的常量正数。

为了简便, SSIM 采用平均加权结构相似度来评估整个图像的质量。

$$\text{MSSIM}(x, y) = \frac{\sum_{\Omega} \text{SSIM}(x_i, y_i)}{M}$$

式中, M 表示图像像素大小。SSIM 考虑了像素间的结构相关性, 评估效果明显优于 PSNR、MSE 等传统评估模型, 取得较好的成功。但也存在一些缺陷, 对模糊失真和噪声失真的区分能力较差与主观图像质量评价方法不一致。例如, 模糊程度较高的图像的 SSIM 值高于噪声程度较低的图像的 SSIM 值, 如图 1 所示。分析其原因, SSIM 在空间域统计计算参考图像与失真图像的像素值的相关系数, 然而在模型中对于图像的几何结构信息比如轮廓、边缘、纹理等没有较好的体现。

3 LST-SSIM 图像质量评价方法

HVS 感知系统对图像结构异常敏感, 图像几



(a) 模糊失真 (SSIM=0.692 8)



(b) 模糊失真 (SSIM=0.688 5)

图 1 比较不同失真类型的 SSIM 值

何结构强度和方向角度信息是图像结构两个重要的属性, 分别体现了视觉显著性和视觉敏感性特征。由于结构方向角度信息更能表达视觉感知信号敏感度, 本文将局部结构方向角度比较函数替换 SSIM 中原有的结构比较函数, 以结构强度函数在图像质量映射图上加权计算得出最终的图像质量评价。

3.1 基于局部结构张量的结构强度与方向角度

边缘与梯度可以较好地描述图像的局部结构信息, 但在噪声干扰下直接计算图像的梯度或边缘并不准确, 局部结构张量可以更精细地刻画图像的边缘结构特征。因此, 本文采用局部结构张量来提取局部结构强度和方向信息。

给定图像 f , 梯度为 $g = \nabla f(a, b)$, 图像中某一像素点 $f(a, b)$ 在 $K \times K$ 领域块 (n) 的局部梯度向量为 $G = (g_1, g_2, \dots, g_n)$, 局部结构张量定义如下:

$$\mathbf{J} = \mathbf{G}\mathbf{G}^T = \begin{bmatrix} \sum_{i \in n} g_{a_i}^2 & \sum_{i \in n} g_{a_i} g_{b_i} \\ \sum_{i \in n} g_{a_i} g_{b_i} & \sum_{i \in n} g_{b_i}^2 \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}.$$

利用协方差矩阵进行 SVD 分解可以得到

$$\mathbf{J} = \mathbf{W} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} [\mathbf{R}_1 \quad \mathbf{R}_2],$$

式中, $\mathbf{W} \perp \mathbf{R}$, 特征向量 $\mathbf{R}_1$ 为图像结构主方向, $\mathbf{R}_2$ 为图像结构次方向, 且 $\mathbf{R}_1 \perp \mathbf{R}_2$, 特征值 λ_1, λ_2 , 分别表示局部结构张量在 $\mathbf{R}_1$ 和 $\mathbf{R}_2$ 方向上能量变化大小. 领域块选取过大不利于计算, 过小无法保证特征值的准确, 经分析在实验中选取 5×5 邻域块比较合适.

据局部结构张量得出结构方向角度和强度信息定义为

$$\begin{cases} \text{结构方向: } \theta = \frac{1}{2} \arctan\left(\frac{2J_{12}}{J_{11} - J_{22}}\right) \\ \text{结构强度: } S = \lambda_1 \frac{(\lambda_1 - \lambda_2)}{(\lambda_1 + \lambda_2)} \end{cases} \quad (2)$$

式中, θ 为相应的单位特征向量 $\mathbf{R} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ 的角度. S 为图像几何结构强度信息, 在图像信息平坦区域, $\lambda_1 \approx \lambda_2 \approx 0$, 说明无图像结构信息, S 值较小. 在图像边缘位置区域, $\lambda_1 > \lambda_2 \approx 0$, S 值较大, 当 S 值越大说明边缘强度能量越大, 与人眼视觉注意选择机制相符合.

3.2 结构强度加权的 LST-SSIM

本文将图像结构方向角度和强度信息与 SSIM 相结合构成一种优于 SSIM 的评价指标 LST-SSIM (Local Structure Tensor-SSIM), 具体的算法结构如图 2 所示.

根据局部结构张量的结构方向角度和强度定义, θ_x, θ_y 分别表示参考图像和失真图像的结构方向角度值, s_x, s_y 分别表示参考图像和失真图像的结构强度值, 结构方向相似度比较函数和结构强度相似度比较函数可以分别表示为

$$a(x, y) = \frac{\sigma_{\theta_x \theta_y} + A_3}{\sigma_{\theta_x} + \sigma_{\theta_y} + A_3},$$

$$g(x, y) = \frac{2S_x S_y + A_4}{S_x^2 + S_y^2 + A_4}.$$

式中, $\sigma_{\theta_x}, \sigma_{\theta_y}$ 分别表示 x 的方向角度方差以及 x 与 y 方向角度的协方差, A_3, A_4 为非常小的正数, 取常量, 是为了避免当分母十分接近零时出现的不稳定性.

LST-SSIM 算法将式(1)中的 $s(x, y)$ 替换成

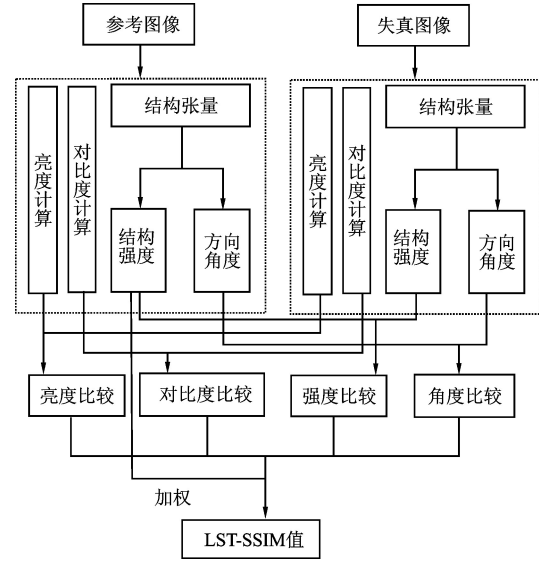


图2 LST-SSIM 评价算法框图

$a(x, y)$ 和 $g(x, y)$, 得到

$$\text{LSTSSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [a(x, y)]^\gamma [g(x, y)]^\delta.$$

一般取 $\alpha = \beta = \gamma = \delta = 1$, 研究者发现, 人眼对于图像的边缘和纹理结构信息最为敏感, 边缘和纹理可能是图像结构信息最重要的部分. LST-SSIM 算法融合了图像结构方向角度和强度相似度, 符合人眼感知图像结构敏感特性. 由于图像质量评价最终决定权在于人眼, 除了图像结构信息敏感外, 视觉显著性也是人眼选择性感知图像信息的重要特征. 受文献[6]中的相位一致值加权图像视觉显著特征的启发, 为了更好地模拟人眼视觉系统, 本文选择参考图像的结构强度信息为计算最终评价值的加权系数, 更加符合人眼感知视觉系统的视觉特性.

因此, ST-SSIM 的最终图像质量评价价值计算公式如下:

$$\text{LST-SSIM} = \frac{\sum_{\Omega} \text{LSTSSIM}(x, y) S_x}{\sum_{\Omega} S_x}$$

参考图像的结构强度值具有更强的显著性特征, 因此本文选用参考图像的结构强度值为最终图像评价指标值的加权系数.

4 实验结果与分析

本文选用了 TID2008^[7] 图像质量评价数据库进行仿真实验. TID2008 包含了 25 幅参考图像, 1 700 幅失真图像, 共 17 种失真类型, 是目前众多图像质量评价数据库中失真类型最全的数据库. 同时给出所有失真图像的描述主观评分 (Mean Opinion

Score, MOS).

为了全面验证算法的有效性,以 SROCC (Spearman Rank Correlation Coefficient)、KRCC (Kendall Rank Correlation Coefficient)、PLCC (Pearson Linear Correlation Coefficient) 和 RMSE (Root Mean Square Error) 四个客观参数作为评估的指标. 参数选取 $A_1 = A_2 = 0.001$, $A_3 = A_4 = 0.003$, $\alpha = \beta = \gamma = \delta = 1$. 采用 Logistic 函数来补偿来拟合客观数据与主观评价分数之间的非线性关系.

选用了四种 SSIM 及改进评价模型 (ESSIM^[4]、MPCSSIM^[5]、MSSSIM^[8]、SSIM) 与本文模型进行比较.

表 1 是五种模型评价对 TID2008 数据库的四个性能指标综合比较, LST-SSIM 的评价结果 SROCC 和 PLCC 性能指标值优于其他四种模型, 其他两种性能指标与最好值较为接近. 表 2 为常见七种失真类型的五种模型评价性能指标 SROCC 值. 从表中可以看出, LST-SSIM 对噪声失真类型和模糊失真类型均优于其他模型. LST-SSIM 使用局部结构张量提取结构方向和结构强度作为结构信息, 既可以防止噪声和模糊的干扰, 又可以较为准确地得到图像真实的结构信息, 获得较好的评价效果.

表 1 五种模型评价对 TID2008 数据库的性能比较

评价模型	性能指标			
	SROCC	KRCC	PLCC	RMSE
ESSIM	0.724 5	0.569 6	0.730 9	0.837 2
MSSSIM	0.852 8	0.654 3	0.842 5	0.837 2
MPCSSIM	0.824 5	0.569 6	0.730 9	0.729 9
SSIM	0.774 9	0.576 8	0.773 2	0.851 1
LST-SSIM	0.863 6	0.651 9	0.858 9	0.792 1

表 2 常见七种失真类型的 SROCC 值

失真类型	评价模型				
	ESSIM	MSSSIM	MPCSSIM	SSIM	LST-SSIM
masked noise	0.801 9	0.815 5	0.817 6	0.779 5	0.818 2
impulse noise	0.692 7	0.686 8	0.691 0	0.673 2	0.701 5
blur	0.971 2	0.960 7	0.972 3	0.954 4	0.980 1
denoising	0.953 8	0.957 1	0.956 0	0.953 0	0.956 8
JPEG 2000	0.972 1	0.973 6	0.972 3	0.962 5	0.973 4
block-distortion	0.761 7	0.851 8	0.838 1	0.846 2	0.847 9
mean shift	0.741 2	0.737 4	0.742 8	0.723 1	0.781 0

从多种失真类型 LST-SSIM 数据点集相对非线性补偿曲线的发散程度看出, 模糊失真和噪声失真以及去噪失真的客观评价结果相对主观评价的合理性, 评价性能有较大提高. 对于 JPEG2000 压缩失真由于失真带来块状结构, 干扰图像原来结构信息, 非线性补偿曲线的数据点集不够集中.

5 结束语

本文从图像的局部结构张量出发在 SSIM 的基础上提出一种新的图像质量评价方法 LST-SSIM, 仿真实验证明 LST-SSIM 模型整体效果较好, 尤其对噪声失真和模糊失真类型的评价质量有较好的提高, 由于增加了视觉显著性加权计算赋予图像结构不同的权重对于其他的失真类型图像评价也得到较好的评价效果.

参考文献:

- [1] Chandler D M, Hemami S S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images[J]. Image Processing, IEEE Transactions on, 2007, 16(9): 2284-2298.
- [2] Sheikh H R, Bovik A C. A visual information fidelity approach to video quality assessment[C] // The First International Workshop on Video Processing and Quality Metrics for Consumer Electronics Auslix, TX, USA: IEEE, 2005: 23-25.
- [3] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity [J]. Image Processing, IEEE Transactions on, 2004, 13(4): 600-612.
- [4] 田浩南, 李素梅. 基于边缘的 SSIM 图像质量客观评价方法[J]. 光子学报, 2013(1), 110-114.
- [5] 苏媛媛, 桑庆兵. 基于相位一致结构相似度的图像质量评价方法[J]. 计算机应用, 2012(8), 2283-2287.
- [6] Zhang L, Zhang L, Mou X, et al. FSIM: a feature similarity index for image quality assessment[J]. Image Processing, IEEE Transactions on, 2011, 20(8): 2378-2386.
- [7] N Ponomarenko, V Lukin, A Zelensky, et al. TID2008-A database for evaluation of full-reference visual quality assessment metrics[J]. Adv Modern Radioelectron, 2009(10): 30-45.
- [8] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment[C] // proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computer. Pacific Grove, CA: IEEE, 2003: 9-12.

作者简介:

罗学刚 男, (1983-), 博士研究生, 讲师. 研究方向为图像处理与计算机视觉.

王华军 男, (1964-), 教授, 博士生导师, 研究方向为计算机视觉、计算机仿真.