

基于正则化方向失真的压缩图像质量评价

程光权 成礼智

(国防科学技术大学理学院 长沙 410073)

摘要: 图像质量评价是图像和视频处理系统中的基本算法。该文在分析人眼视觉系统对图像感知特点的基础上,结合二维图像的几何本征特征研究,提出了一种新型的基于图像正则化方向失真的完全参考质量评价算法。该方法计算简单,物理意义明确。仿真试验结果表明,该方法符合人眼视觉系统特性,所得到的图像质量客观评价结果与主观评价结果之间具有更好的一致性。

关键词: 图像质量评价; 正则化方向失真; 图像压缩; 人眼视觉系统

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2010)06-1316-05

DOI: 10.3724/SP.J.1146.2009.00867

Regular Directional Distortion Based Compressed Image Quality Assessment

Cheng Guang-quan Cheng Li-zhi

(School of Science, National University of Defense Technology, Changsha 410073, China)

Abstract: Image quality assessment is fundamental importance for image or video processing system. Following the knowledge of properties of human visual system and features of natural image, a novel full reference image quality assessment based on the regular directional distortion is proposed in this paper. The proposed method is low complexity and has clear physical structure. The experimental results upon image database show that the proposed method is in accordance with characteristic of human visual system and has a good consistency with the subjective assessment of human beings.

Key words: Image quality assessment; Regular directional distortion; Image compression; Human visual system

1 引言

图像质量评价在图像与视频处理系统中占据十分重要的地位,直接说明了图像处理系统的优劣和算法的有效性,同时图像质量评价可以指导算法的设计与优化,进一步提升图像处理系统的效率。近年来,越来越多的研究者关注这个领域的研究,提出了一系列超越传统算法的新的图像质量评价标准。

图像质量评价方法可以分为主观质量评价和客观质量评价两大类。由于对图像的评价最终取决于观察者的感受,因此直观上主观评价方法最可靠。主观评价方法就是让多名观察者按照已规定好的评价准则对被测图像进行主观打分,然后对所有观察者给出的评分进行加权平均,也称为平均评定得分法(MOS)。主观评价方法结果符合人眼视觉系统特性,但是这种方法计算不便、速度慢、成本高,受到诸多因素限制,且无法用数学模型进行描述,无

法应用于图像处理算法的设计和优化过程中,因而在实际应用中无法得到广泛推广。相对于主观质量评价,客观质量评价方法具有操作简单、成本低、易于实现等特点,成为图像质量评价研究的重点。图像质量的客观度量有效弥补了主观评价方法的不足,并被广泛应用到图像处理系统的优化中。近年来客观评价方法受到广泛重视,国内外众多的研究机构和人员都对此展开了深入的研究,如 VQEG (Video Quality Evaluation Group) 和 LIVE (Laboratory for Image and Video Engineering) 等研究小组。根据接收端利用原始图像信息的程度,客观评价方法又可分为完全参考、部分参考和无参考 3 种类型^[1]。本文主要研究完全参考图像质量评价,即在接收端失真图像和原始图像信息都完全可利用。传统的完全参考客观评价方法主要是均方误差(MSE)和峰值信噪比(PSNR),这两种方法由于操作简单、计算量小而得到广泛应用,也是目前主要使用的质量评价方法。但 MSE 和 PSNR 只是对失真图像和原始图像纯误差的简单数学统计,难以反映人眼的视觉误差,有时甚至会导致和主观评价方法相矛盾的结论^[2]。近年研究者在研究人眼视觉系统特

2009-06-12 收到, 2009-10-19 改回

国家自然科学基金(60573027, 10601068)和全国优秀博士学位论文作者专项资金(200543)资助课题

通信作者: 程光权 cgg299@163.com

性的基础上, 提出了一系列新型的有效的图像质量评价方法, 以弥补传统方法的不足。其中, 在文献[3]中作者考虑失真图像的结构信息损失, 认为人眼视觉系统的主要功能是提取图像和视频中的结构信息, 因此提出基于结构失真模型的图像质量评价标准(SSIM), 仿真试验结果证明了该方法优于经典方法, 较好地符合人眼视觉系统特性, 吸引了众多研究者的关注^[4,5]。本文也是关注基于结构信息失真的图像质量评价研究, 符合人眼视觉系统特性的结构信息是一个开放性的问题, 文献[3]中的方法考虑的结构信息主要是信号的一些统计信息特征, 而图像作为一类特殊的 2 维信号具有其自身独有的特点, 因此, 本文从另外一个方面考虑影响失真图像质量评价的结构信息。

本文在研究人眼视觉系统视觉感知的特点以及 2 维自然图像本征特征的过程中发现, 自然图像区别于 1 维信号, 具有实际的物理含义, 其中最重要的一点就是自然图像具有几何正则性的结构特性, 这些特征构成了 2 维图像信号中人眼可感知的内容, 人眼在感知图像失真的过程中, 更多的关注这些几何正则性结构特征的失真情况。基于此, 本文考虑用几何正则化方向失真来刻画这种几何结构信息损失, 提出了一种新的基于正则化方向失真模型的质量评价方法, 并在图像数据库中验证本文方法的有效性。

2 正则化方向失真模型

自然图像只是 2 维信号中的一小部分, 区别于 2 维随机信号, 之所以成为人眼观测有物理意义的自然图像, 其中十分重要的一点就是其具有几何结构特征, 这些几何特征构成了图像中符合人类认知的各种自然界的具体目标。根据神经科学以及心理物理学等领域的研究可知, 人眼对一幅图像进行质量评价的时候, 主要观测这幅图像中富含的符合自然界事物本身具有的本征几何特征改变的程度。自然图像特征丰富, Ran 提出了符合人眼视觉特性的图像 3 构件模型^[6], 即图像由边缘、纹理(弱边缘)和平坦区域 3 部分构成, 其中边缘对视觉观察起决定性作用, 纹理次之, 平坦区域只能起陪衬作用。这就是说, 边缘(包括弱边缘)是图像中影响视觉效果的主要因素。因此在图像质量评价体系中, 对边缘和纹理部分内容应该予以较多的关注, 这些几何特征构成了图像中各种具有物理意义的目标, 是图像中最令人眼视觉系统感兴趣的部分。因此在图像质量评价系统中, 比较原图像和失真图像的这些几何结构的变化情况, 将很好地反应图像失真对人眼视觉系统的影响。

如何定量地考虑失真图像中边缘和纹理这些几何特征的改变, 决定了质量评价模型的建立和评价算法的有效性。我们注意到 2 维图像中的像素点具有一定的方向奇异性, 沿平行于边缘或纹理的方向, 图像像素点灰度值变化是正则的或者说是缓慢的, 而垂直于边缘或纹理的方向, 图像灰度值变化剧烈。这种性质称之为图像的几何正则性^[7], 因此图像中像素点的这种方向特性也称为正则化方向。这种几何正则性特征是图像几何特征中最根本的一类特征, 图像质量的变化对图像正则性十分敏感, 因此首先要考虑失真图像和原图像中几何特征的正则化方向的比较。这也符合图像的实际应用背景, 图像在实际应用过程中特别关注局部目标方向特征的情况, 例如在图像压缩的军事应用中, 特别关注压缩后图像中是否改变了船只航行的尾迹方向, 以及车辆等行驶的痕迹方向等。这些特征的丢失或改变都将对压缩图像的后续应用造成很大的影响。

准确地计算 2 维图像中像素点的方向是比较困难的, 我们注意到, 在包含一段边缘的区域内, 像素的正则化方向平行于经过该点边缘曲线的切线, 而梯度方向垂直于切线, 如图 1 所示, $v(i, j)$ 为该点切线向量, $d(i, j)$ 为该点的梯度向量, C 为一段边缘。基于这个性质, 我们利用该点的梯度方向得到该点的正则化方向。

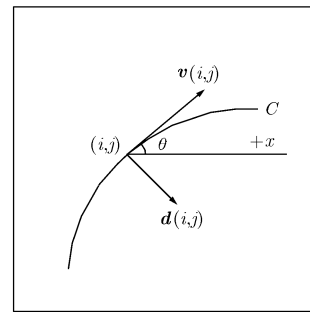


图 1 方向图

$$\theta(i, j) = \tan^{-1} \left(\frac{\partial f_y(i, j)}{\partial f_x(i, j)} \right) + \pi/2 \quad (1)$$

其中 $\theta(i, j)$ 表示该点的正则化方向和 $+x$ 数轴的夹角, $\partial f_x(i, j)$ 和 $\partial f_y(i, j)$ 分别表示图像在 (i, j) 点水平方向和垂直方向梯度, 在计算中, 方向梯度采用 Sobel 算子获得。这个公式可以推广到非边缘区域像素点方向的定义, 从而得到整幅图像的方向图。显然, 几何正则方向的改变由失真图像和原图像在该点正则化方向的差决定的:

$$\Delta\theta(i, j) = |\theta_f(i, j) - \theta_g(i, j)| \quad (2)$$

其中 $\theta_f(i, j)$, $\theta_g(i, j)$ 分别表示原图像 f 和失真图像 g

在 (i, j) 点正则方向和 $+x$ 数轴的夹角。显然, 当 $\Delta\theta(i, j) = \pi/2$ 时, 表示正则方向改变最严重, 而 $\Delta\theta(i, j) = 0$ 或 π 则表示方向没有改变, 如图 2 所示。

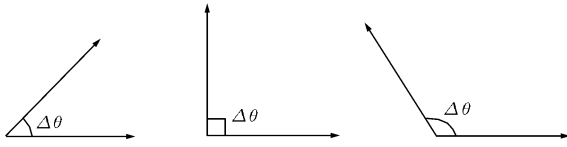


图 2 方向差比较

这个性质和余弦函数的性质相似, 因此这里定义失真图像和原图像的方向失真为

$$GD(i, j) = |\cos(\Delta\theta)| = |\cos(\theta_f(i, j) - \theta_g(i, j))| \quad (3)$$

从式(1)和式(3)可以推得

$$GD(i, j) = \frac{|\partial f_x(i, j) \cdot \partial g_x(i, j) + \partial f_y(i, j) \cdot \partial g_y(i, j)| + C_1}{\sqrt{\partial f_x^2(i, j) + \partial f_y^2(i, j)} \cdot \sqrt{\partial g_x^2(i, j) + \partial g_y^2(i, j)} + C_1} \quad (4)$$

这里包含常数 C_1 是为了避免分母等于或近似等于 0 时计算不稳定的情况。上述定义的方向失真是对称的, 且其数值范围是 $[0, 1]$, 当且仅当 $\theta_f(i, j) = \theta_g(i, j)$ 时取得最大值 1, 此时意味着原图像和失真图像在该点的正则化方向保持一致。注意到, 参考文献[5]中也将图像梯度信息应用于图像质量评价中, 但是作者仅仅将梯度信息作为一个参考因子, 利用与 SSIM 方法类似的梯度的局部统计信息比较来衡量图像质量, 并没有意识到梯度信息中所能包含的更进一步的几何正则方向信息的失真, 但正因为梯度信息也包含了一部分的几何正则信息, 因此文献[5]中的方法也获得了一定提高。而本文方法是直接挖掘图像本质的几何正则化结构信息, 具有直观的物理意义, 只是利用梯度信息来定义图像的这种正则化方向失真。

其次, 方向的强度可以反映边缘和纹理这些几何结构的强度失真, 也是几何正则结构变化的重要一部分。前面计算像素点的几何正则方向时候, 利用了图像的方向梯度信息, 而梯度的幅值可以反应边缘和纹理的强度, 因此梯度的幅值失真是比较几何正则结构失真的另外一个分量。

$$m(i, j) = \sqrt{\partial f_x^2(i, j) + \partial f_y^2(i, j)} \quad (5)$$

为了保持与方向失真数值标准一致, 定义幅值失真如下:

$$GM(i, j) = \frac{2m_f(i, j) \cdot m_g(i, j) + C_2}{m_f^2(i, j) + m_g^2(i, j) + C_2} \quad (6)$$

这里, $m_f(i, j)$ 和 $m_g(i, j)$ 分别表示原图像 f 和失真图像 g 在 (i, j) 点梯度幅值, 同样地, 常数 C_2 是避免分

母出现不稳定现象。这个定义式同样是对称的, 数值范围为 $[0, 1]$, 当且仅当 $m_f(i, j) = m_g(i, j)$ 时取得最大值 1。联合方向失真和幅值失真, 定义本文的几何正则化方向失真, 这里选用如下函数:

$$GDM(i, j) = GD(i, j)^\alpha \cdot GM(i, j)^\beta \quad (7)$$

在本文中仅简单地选择 $\alpha = \beta = 1$ 的情形。可以简单地验证, 这个定义式是对称的, 取值范围是 $[0, 1]$, 数值越大表明图像质量越好。

前面说明了, 人眼视觉系统在感知图像质量过程中, 对图像不同部分的内容敏感度不同。直观上, 人眼对图像中的物体目标区域即灰度值变化较剧烈区域较为重视, 对背景等变化平缓区域关注度较低。而图像的失真往往也发生在图像灰度值变化剧烈的区域, 如在图像压缩中, 经常出现的边缘模糊等现象。因此, 本文仅选择图像中影响人眼视觉系统评价图像质量的像素点作为参考像素集合。图像梯度也是反映图像局部变化的一个参量, 梯度值大的表明图像在该区域变化剧烈, 正是影响视觉系统评价的区域。选择一个阈值 T , 对图像(包括原图像和失真图像)中梯度幅值大于该阈值的像素点作为图像质量评价的参考点, 设为 Ω , 即

$$\Omega = \{(i, j) | m_f(i, j) > T \text{ or } m_g(i, j) > T\} \quad (8)$$

其中阈值 T 的设定根据失真图像和原图像梯度幅值的均值 \bar{m}_f 和 \bar{m}_g 决定。

$$T = \lambda \cdot (\bar{m}_f + \bar{m}_g) \quad (9)$$

其中 λ 为参数。最后, 整个失真图像的质量评价定义为所有参考点上的几何正则化方向失真的平均 (MGDM)。

$$MGDM(f, g) = \frac{1}{M} \sum_{(i, j) \in \Omega} GDM(i, j) \quad (10)$$

其中 M 是参考点集合 Ω 中像素点样本数。

3 试验结果

这里通过一些数值试验来验证本文质量评价方法的性能。其中, 参数设定为 $C_1 = 0.1$, $C_2 = 0.2$, $\lambda = 0.05$ 。试验数据选用美国 TEXAS 大学 Laboratory for Image and Video Engineering (LIVE) 的图像数据库^[8]进行仿真试验, 该库包含 29 幅高分辨率的 RGB 图像和通过 JPEG 和 JPEG2000 压缩得到的一系列失真图像, 给出了参考图像、压缩后的图像以及压缩后图像即失真图像的主观打分值。这里采用 PSNR, 文献[3]的 SSIM 方法和本文方法 MGDM 分别对图像序列进行质量预测, 试验方法应用于灰度图像, 彩色图像与灰度图像通过下式转换。

$$I = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B \quad (11)$$

其中, R, G, B 分别表示 RGB 图像的红色、绿色和蓝色分量。

图 3 显示了 PSNR, SSIM 和 MGDM 3 种客观评价结果和图像主观评价价值 MOS 的关联曲线图, 可以看出, 本文模型相比 PSNR 和 SSIM 方法, 主客观质量评价关联图中数据相对更集中, 关联图中“野点”减少, 直观上表明本文方法和主客观观测数值具有更好的一致性, 更加符合人眼视觉系统特性。为了进一步客观量化比较本文方法和其他方法的优劣, 根据 VQEG 提出的准则^[9], 在客观评价和主观评价价值之间存在一定的非线性关系, 这里采用下面的函数建立这种非线性映射关系:

$$f(x) = \frac{a_1 - a_2}{1 + \exp\left(-\frac{x - a_3}{a_4}\right)} + a_2 \quad (12)$$

其中 x 表示客观评价价值, a_1, a_2, a_3 和 a_4 为模型参数, 通过 Matlab 中非线性优化函数求得。在建立非线性映射之后, 选用如下 3 个度量指标来比较各种方法的优劣。(1)相关系数(CC), 客观评价和主观评价价值之间的相关系数反映预测的精确性; (2)

Spearman 秩相关系数(SROCC), 反映预测的单调性, 即模型预测值与主观值相对幅度的一致程度; (3)离出率(OR), 反映客观评价的稳定性。CC 和 SROCC 值越大, OR 值越小则表明模型预测越好。表 1 列出了几种测试模型试验结果的性能指标参数值。可以看出, 本文方法相比较 PSNR 和 SSIM 模型具有更高的预测精度(CC), 预测单调性(SROCC)和更低的离出率(OR): 例如对 JPEG2000 失真模型试验数据结果, CC 值从 PSNR 结果的 0.9331 和 SSIM 结果的 0.9492 增加到本文方法的 0.9596; SROCC 值从 PSNR 结果的 0.9041 和 SSIM 结果的 0.9368 增加到本文方法的 0.9476; 而 OR 值从 PSNR 结果的 0.0947 和 SSIM 结果的 0.0651 降低为本文方法的 0.0473。

通过试验结果的主、客观数据的关联图以及不同模型性能量化参数的比较, 可以发现, 本文方法都优于 PSNR 和 SSIM 模型, 考虑了几何正则方向信息比较, 结合人眼对不同区域的观测影响, 能更好地判断失真图像的特征信息损失, 符合人眼的视觉系统模型, 结果与主观评价方法更加一致, 取得

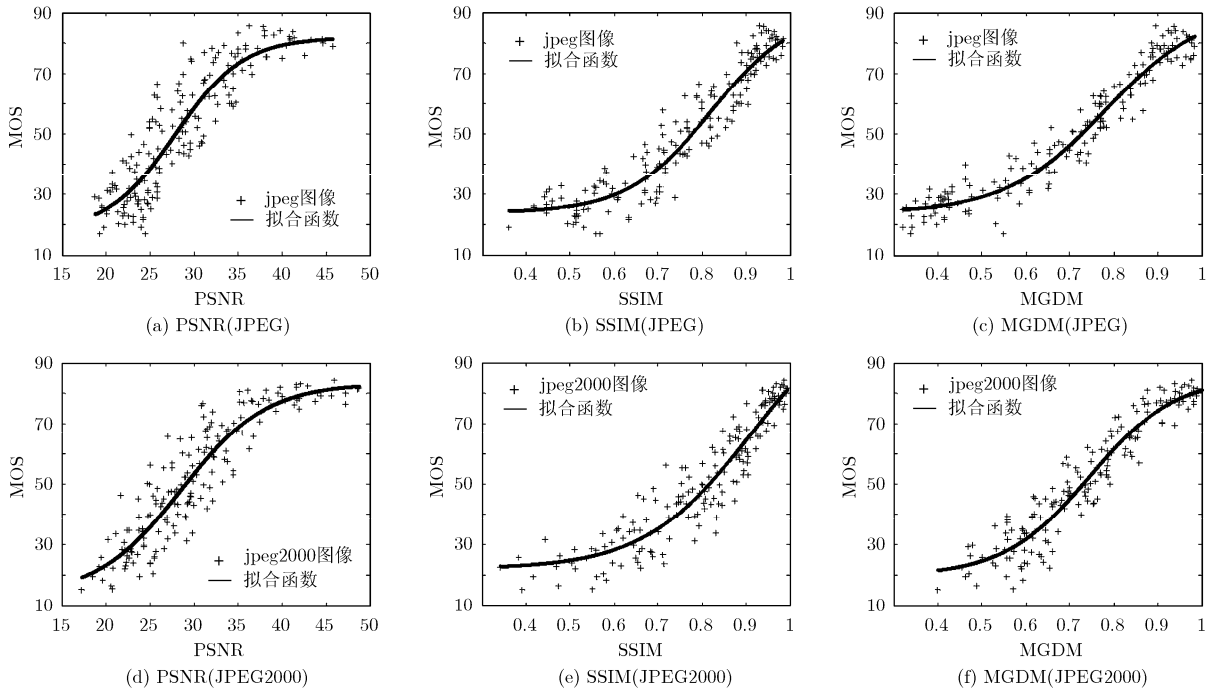


图 3 客观图像质量评价方法与主观 MOS 值的关联图

表 1 客观评价模型性能指标比较

	JPEG			JPEG2000		
	CC	SROCC	OR	CC	SROCC	OR
PSNR	0.9230	0.8899	0.1886	0.9331	0.9041	0.0947
SSIM	0.9674	0.9485	0.0414	0.9492	0.9368	0.0651
MGDM	0.9742	0.9566	0.0343	0.9596	0.9476	0.0473

很好的预测效果。因此,从试验结果可以得出结论:本文提出的方法计算简单,物理意义明确,试验结果相比较传统 PSNR 方法和最近流行的 SSIM 方法有较大改进,与主观评价具有更好的一致性。

4 结论

建立通用的可被广泛接受的和主观评价结果相一致的图像质量评价方法是图像处理算法中亟需解决的问题之一,本文讨论了静止图像质量评价方法,提出了一种新型的完全参考图像质量评价方法。根据 2 维图像信号具有的本征特征,以及人眼视觉系统在感知图像过程的特点,在新型图像质量评价过程中,主要关注图像的正则方向以及强度的失真情况,并且选择在影响图像质量的感兴趣参考点集合上度量这种失真。算法简单较易实现,物理背景清晰。仿真试验结果表明,与其他方法相比,该模型的野点数目相对减少,与人眼的主观评价具有更好的关联性,更加符合人眼视觉系统判断结果,能有效评价静止图像质量。

参考文献

- [1] Sheikh H R, Sabir M F, and Bovik A C. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transaction on Image Processing*, 2006, 15 (11): 3441-3453.
- [2] Wang Z and Bovik A C. Mean squared error: Love it or leave it? — A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 2009, 26(1): 98-117.
- [3] Wang Z, Bovik A C, and Sheikh H R, *et al.*. Image quality assessment: from error visibility to structural similarity. *IEEE Transaction on Image Processing*, 2004, 13(4): 600-612.
- [4] Brooks A C, Zhao X, and Pappas T N. Structural similarity quality metrics in a coding context: Exploring the space of realistic distortions. *IEEE Transaction on Image Processing*, 2008, 17(8): 1261-1273.
- [5] Chen G H, Yang C L, and Xie S L. Gradient-based structural similarity for image quality assessment. Proc. IEEE International Conference on Image Processing, Atlanta, USA, Oct. 2006: 2929-2932.
- [6] Ran X and Farvardin N. A perceptually motivated three-component image model-part I: Description of model. *IEEE Transaction on Image Processing*, 1995, 4(4): 401-405.
- [7] Mallat S and Peyre G. Orthogonal bandlet bases for geometric images approximation. *Communications on Pure and Applied Mathematics*, 2008, 61(9): 1173-1212.
- [8] Sheikh H R, Wang Z, and Bovik A C, *et al.*. Live image quality assessment database release, <http://live.ece.utexas.edu/research/quality>.
- [9] VQEG, Final report from the video quality experts group on the validation of objective models of video quality assessment, <http://www.vqeg.org>.

程光权: 男, 1982 年生, 博士生, 研究方向为新型小波理论、图像处理、计算数学。

成礼智: 男, 1962 年生, 教授, 博士生导师, 研究领域为信息科学中新型算法与软件、小波变换与图像处理、计算数学。