

图像结构失真的编码质量评价

庞建新 张荣 张晖 刘政凯

(中国科学技术大学电子工程和信息科学系,多媒体计算与通信教育部-微软重点实验室,合肥 230027)

摘要 客观图像质量评价方法——自动评价图像质量使其与主观感受一致,对于很多图像和视频处理应用特别是图像编码具有重要意义。提出了一种基于结构失真的图像编码质量客观评价方法。首先根据匹配追逐算法将参考图像在基函数上进行投影分解按照重要性排列获得结构信息,并给出结构重要性度量,然后利用该信息计算编码图像的结构信息,比较编码前后结构信息的差异作为客观质量评价度量,其中在匹配追逐算法中,选取了2维可分离的Gabor基函数集作为字典。实验结果显示,上述方法与主观质量评价有很好的 consistency。

关键词 图像质量评价 匹配追逐 图像结构 图像编码

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2009)08-1560-09

Quality Assessment for Image Coding with Structural Distortion

PANG Jian-xin, ZHANG Rong, ZHANG Hui, LIU Zheng-kai

(MOE-microsoft Key Laboratory of Multimedia Computing and Communication, Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027)

Abstract Objective image quality assessment, which automatically values the image quality consistently with subjective perception, is important to numerous image and video processing applications, especially in image coding. In this paper, a new objective measure for image coding quality based on image structure is proposed. First of all, the structure information is calculated by projecting the reference image onto the base functions using the algorithm of matching pursuit. Secondly, the structure information of the coded image in the order of importance is computed and gained, and the characteristics of the importance of the structure information are given, using the structure information of the reference image. Finally the objective quality score is concluded by comparing the differences of structure information between the reference and the coded images. And the separable 2-D Gabor base functions are selected as the dictionary for the algorithm of matching pursuit. Experimental results show that the proposed approach is consistent with the subjective quality score.

Keywords image quality assessment, matching pursuit, image structure, image coding

1 引言

近年来图像编码技术发展迅速,而图像编码的好坏不仅仅决定于编码的效率,还在于编码的质量。主观评价图像编码质量既耗时又耗力,无法普遍使用,因此,迫切需要一种客观评价图像编码质量的方法。

在过去研究中,提出了很多种图像质量的客观评价方法^[1]。这些方法大体可以分为3类:

(1)基于像素数字统计特征的方法 如均方误差(MSE),以及由它派生出来的信噪比(SNR)、峰值信噪比(PSNR)、均方根误差(RMSE)等。这类方法不能很好地反映人的视觉特性,与人的主观评价脱节^[1-10]。但是由于这些算法计算简单,实现容易,故而被广泛使用。

收稿日期:2007-05-21; 改回日期:2008-04-14

第一作者简介:庞建新(1982~),男,中国科学技术大学电子工程与信息科学系博士研究生。主要研究方向为数字图像处理和理解、图像质量评价。E-mail:waltonpang@ustc.edu

(2) 基于人类视觉系统(HVS)的方法 Mannos 和 Sakrison 在 1974 年的一篇著名的文章中最早提出将人类视觉特性的概念用于图像质量评价^[11],他们通过心理物理实验找出了一些 HVS 的特质,总结出了对比敏感度的闭合表达式,至今仍广泛用在 HVS 模型中。Watson 提出一种基于 DCT(离散余弦变换)的视觉误差衡量标准^[12],其中每一个 DCT 系数的量化误差都按照相应的视觉灵敏度进行伸缩。Karunasekera 和 Kingsbury 研究各种不同的失真类型,按照人眼对不同失真的敏感程度把它们加权平均起来,最后得到失真结果^[13]。Chou 和 Li 提出了一种基于刚辨差(JND, just-noticeable difference)的图像多通道分解模型,来估计一幅灰度图像的 JND/MND(刚辨差/最小辨差)的分布情况,设计出子带分解算法来去除由 JND/MND 模型算出的视觉冗余,从而获得视觉无损或最小损失的效果^[14]。Mayache 等人^[15]比较了以上 3 种著名的基于 HVS 的图像评价方法。李均利等人基于重要性测度和模糊积分,提出了一种基于视觉模型的图像编码质量的评价方法^[16]。这种建立在对人的心理物理实验基础上通过模拟人的 HVS 特性的评价方法尽管获得了普遍的认同,但是由于 HVS 系统的复杂性以及人们认识的有限性,上述方法同样存在一定的局限性,制约着此类方法的进一步发展^[3,9]。

(3) 基于图像结构失真的方法 Wang 等人首先提出结构失真的假设^[3,9],他们认为人眼有很强的提取图像中结构信息的能力,把“结构信息”定义为一种表现物体结构的属性,而与图像的平均亮度和对比度无关,用来比较原始图像与失真图像结构的相似程度。Wang 等人在此基础上提出了 MSSIM(mean structural similarity)算法^[3],并且对结构相似度做出了完整解释和分析;MSSIM 算法将图像信息分成亮度、对比度和结构 3 部分,计算 3 部分的失真获得图像分块的整体失真度量,最后使用求均值的方法获得整幅图像的失真度。Shnayderman 等人从矩阵奇异值分解(SVD)的角度,通过比较参考图像和失真图像矩阵奇异值的差异获得图像的结构失真,提出了 SVD 的评价方法^[17],有较好的效果。此类方法使用了图像结构信息失真的概念,对图像质量进行评价,这种基于结构失真的评价方法是目前研究的热点。上述方法尽管有较好的结果,但没有对什么是结构进行解答,也没有考虑到图像中不同结构具有不同重要性,因此具有一定的局限性。

图像失真类型很多,编码失真是其中最常见也是应用最广的一种,本文仅研究图像编码造成的图像失真。图像通常可以分成纹理区域和结构区域,一般而言,自然图像中结构区域的重要性更大^[3],由于它们各自所表示的内容不同,不同的结构区域往往具有不同的重要性。基于结构失真的思想提出了一种图像结构特征和特征重要性表示、度量方法,并且在此基础上提出一种图像编码质量评价方法。通过选择 2 维可分离的 Gabor 基建立了一个基函数集(字典),采用匹配追逐算法(MP)^[18]将图像在字典上进行投影分解,按照重要性排列获得图像结构信息,并且给出结构信息的重要性度量,通过计算失真图像与参考图像结构信息的差异获得图像的客观质量值(MP_Q)。

2 匹配追逐算法

匹配追逐算法将信号在基函数上进行投影分解,根据能量最大的原则可以分解出信号最重要的部分。根据这个原理,选择合适的基函数来分解出图像中最重要的结构信息,并且给出这种结构信息的重要性度量,通过比较失真前后的结构信息失真得到图像质量评价。

2.1 匹配追逐算法

匹配追逐算法^[18]最早由 Mallat 和 Zhang 于 1993 年提出,这种方法借鉴了传统的投影追逐算法的思想,并将其进一步完善,以作为一种时频分析的手段。简单地说,匹配追逐算法就是将信号在非正交基上以能量最大为准则的投影分解。

匹配追逐算法基本思路如下:

记 $D = \{g_\gamma\}$ 为字典,即基函数集, g_γ 为字典元素,即基函数,满足 $\|g_\gamma\| = 1$, 信号 $f(t)$ 满足:

$$f(t) = \sum_{n=0}^{+\infty} a_n g_{\gamma_n}(t) \quad (1)$$

其中, $g_{\gamma_n}(t)$ 为对于第 n 次匹配时的字典元素,并满足以能量最大的准则下 a_n 的重要性递减,即 $|a_0| \geq |a_1| \geq |a_2| \geq |a_3| \geq \dots \geq |a_n|$ 。

记 Rf 为 f 去除在 g_γ 上投影的残余分量,有如下等式:

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + Rf \quad (2)$$

$$\|f\|^2 = |\langle f, g_{\gamma_0} \rangle|^2 + \|Rf\|^2 \quad (3)$$

记 $R^0 f = f, R^n f$ 为第 n 次的残余分量。 $R^n f$ 可继续分解为

$$R^n f = \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} f \quad (4)$$

显然有:

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^m f \quad (5)$$

$$\|f\|^2 = \sum_{n=0}^{m-1} |\langle R^n f, g_{\gamma_n} \rangle|^2 + \|R^m f\|^2 \quad (6)$$

在 p 次迭代后能量满足下式时停止迭代:

$$\|R^p f\| = \left\| f - \sum_{n=0}^{p-1} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} \right\| \leq \varepsilon \|f\| \quad (7)$$

等价于:

$$\|f\|^2 - \sum_{n=0}^{p-1} |\langle R^n f, g_{\gamma_n} \rangle|^2 \leq \varepsilon^2 \|f\|^2 \quad (8)$$

信号 f 在 g_{γ} 上投影分解出结构信息, 对应不同的结构信息, 可以得到对应的在 f 中匹配位置信息 $[x, y]$ 和在 g_{γ} 上的投影能量, 以及所对应的基函数。搜索匹配字典的过程称为“寻找原子”, 结构原子包括了 3 个参数: 匹配的位置信息 $[x, y]$ 、在 g_{γ} 上的投影能量以及所匹配的基函数。

根据算法流程可以知道, 尽管匹配追逐是一种非线性的算法, 但仍然能保证其展开在能量上的收敛性。

2.2 匹配追逐算法的字典

从 2.1 节对匹配追逐算法理论的介绍可以看出, 匹配追逐首先要有一组基 (字典), 然后才能将信号向基上投影。由于匹配追逐算法的特殊性, 对字典限制非常低, Mallet 和 Zhang 在文献 [20] 中提出了两种字典, 一种是基于小波包的, 另外一种是基于 Gabor 基函数。Gabor 基函数具有较好的时频分析能力, 能较准确地匹配图像中的形状、尺度以及结构信息中的时频特性, 在图像和时频分析中有较好的性能。选取了 2 维可分离的 Gabor 基函数进行字典构造, 其中字典的参数来自于 Neff 和 Zakhor 的工作^[19]。

定义 Gabor 基函数如下:

$$g_{\alpha}^u(i) = C_{\alpha} \cdot g\left(\frac{i - \frac{N}{2} + 1}{s}\right) \cdot \cos\left(\frac{2\pi\xi\left(i - \frac{N}{2} + 1\right)}{16} + \phi\right) \quad (9)$$

$$i \in \{0, 1, \dots, N-1\}$$

其中, N 为字典 α 的字典长度, C_{α} 为字典的缩放比例, ξ 为缩放变量, ϕ 为偏移量。缩放后要求该字典所有元素的平方和为 1, 其中 $g(t) = \sqrt{2}e^{-\pi t^2}$ 为高斯窗函数。

通过 1 维的 Gabor 基函数, 可以在此基础上构造 2 维的 Gabor 字典。2 维 Gabor 基函数定义为

$$g_{\alpha, \beta}(i, j) = g_{\alpha}(i) \cdot g_{\beta}(j) \quad (10)$$

$$i, j \in \{0, 1, \dots, N-1\}$$

由上可知, 通过此方法可获得一个超完备的字典, 字典元素数据庞大。Neff 和 Zakho 采用了训练的方式建立了一个优化的字典。首先建立一个非常大的 2 维 Gabor 字典集, 然后用这些集合对大量的视频进行训练, 那些被使用次数较多的 2 维 Gabor 函数将留在字典中, 得到一个含有 400 个字典元素的 2 维字典^[19], 其中最大尺寸的字典元素为 35×35 , 最小尺寸的是 1×1 。本文参考使用了 Neff 和 Zakho 训练的 2 维字典^[19], 字典函数如图 1 所示。

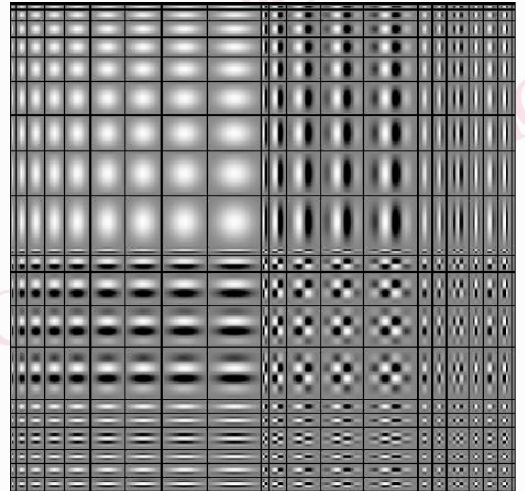


图 1 2 维可分离 Gabor 基函数集

Fig. 1 The separable 2D Gabor basis set

2.3 改进的匹配追逐算法

对匹配追逐算法做出改进以满足质量评价的需要。

式(4)可以改写为

$$R^{n+1} f = R^n f - \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} \quad (11)$$

式(11)是一种可逆的运算, 在图像编解码中是必需的。为了满足图像质量评价的需要, 对该公式进行修改, 定义图像在匹配位置信息 $[x, y]$ 上的图像数据为 df , 其中第 n 阶为 df_n 。用 df_n 代替 $\langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n}$, 则

$$R^{n+1} f = R^n f - df_n \quad (12)$$

其中, $R^0 f = f$ 。当运算阶数 $n = K$ 时, 运算停止, 对于每一个图像都可以获得 K 个不同重要的结构信息。本文定义结构信息 (SC) 为

- g_{γ_n} : 匹配的字典元素
- $[x_n, y_n]$: 匹配位置
- df_n : 匹配位置的图像数据
- $\langle df_n, g_{\gamma_n} \rangle$: 图像数据和字典的内积

其中, g_{γ_n} 表示第 n 个 SC 的结构特征, $\langle df_n, g_{\gamma_n} \rangle$ 为该结构的重要性度量。图 2 表示了式 (12) 的计算流程。

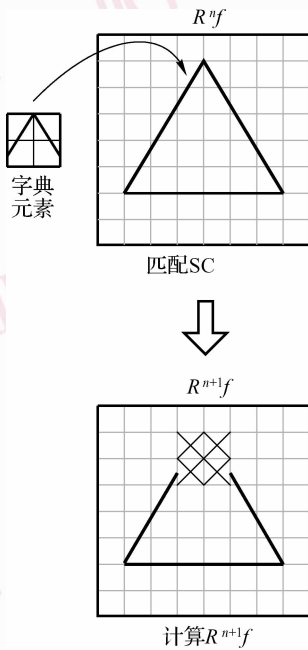


图 2 SC 的匹配过程
Fig. 2 The procedure of matching SC

3 基于结构失真的质量评价

考虑到算法的复杂性以及提高系统性能, 先将图像进行分块处理, 使用匹配追逐的方法计算参考图的结构信息, 再根据参考图像的结构信息计算失真图像的结构信息, 通过比较失真前后图像的结构信息差异, 将这些结构失真信息累计最后得到图像质量的客观度量。

3.1 分块的质量评价

定义参考图像为 f , 大小为 $M \times L$, 将图像分块, 每块记为 $block_i$, 每一分块尺寸为 $m \times l$, 分块总数 $Num = (M \times L) / (m \times l)$, $i = 0, 1, \dots, Num - 1$, 字典为 $\{g_{\gamma}\}$, SC_n^i 为 $block_i$ 中的 SC 信息; 失真图像定义为 \bar{f} , 分块为 \bar{block}_i , 以及失真分块为 \bar{SC}_n^i 。

首先, 对参考图像 f 进行计算, 获取参考图像结构位置和结构特征值。对于其中的任意一个 $block_i$,

将该分块在字典 $\{g_{\gamma}\}$ 上以内积能量最大的原则投影分解, 获得参考图像的结构单元的特征信息:

$$SC_n^i: \{df_n^i, \langle df_n^i, g_{\gamma_n}^i \rangle, g_{\gamma_n}^i, [x_n^i, y_n^i]\} \quad (13)$$

$$n = 0, \dots, K - 1$$

以上操作是单独对参考图像进行, 是独立于失真图像的, 利用该信息计算失真图像的结构信息, 共享其中 2 个参数: $g_{\gamma_n}^i$ 和 $[x_n^i, y_n^i]$:

$$\bar{SC}_n^i: \{\bar{df}_n^i, \langle \bar{df}_n^i, g_{\gamma_n}^i \rangle, g_{\gamma_n}^i, [x_n^i, y_n^i]\} \quad (14)$$

$$n = 0, \dots, K - 1$$

其中, \bar{df}_n^i 是失真图像在参考图像结构信息对应位置 $[x_n^i, y_n^i]$ 所在的图像数据, $g_{\gamma_n}^i$ 是参考图像对应的结构特征, 这里也用来表示失真图像对应的结构特征, $\langle \bar{df}_n^i, g_{\gamma_n}^i \rangle$ 是失真图像数据与 $g_{\gamma_n}^i$ 的内积, 表示该结构信息的重要性度量。失真图像的结构信息不需要再使用匹配追逐算法进行计算, 仅需要利用参考图像的结构信息。

为了表示方便, 定义 $P_n^i = \langle R^n f^i, g_{\gamma_n}^i \rangle$, $\bar{P}_n^i = \langle \bar{df}_n^i, g_{\gamma_n}^i \rangle$ 。同样, 考虑各个 SC 的大小, 也就是 $[x_n^i, y_n^i]$ 所覆盖在图像上面积, 定义为 S_n^i 。

定义 D^i 为失真图像 \bar{f} 中 \bar{block}_i 的结构失真度量:

$$D^i = \sqrt{\sum_{k=0}^{K-1} w_k^i (P_k^i - \bar{P}_k^i)^2}, \sum_{k=0}^{K-1} w_k^i = 1 \quad (15)$$

D^i 表示失真图像中失真分块各结构单元相对于参考图像的失真量, w_k^i 权值用以表示各结构单元的重要程度, 结构单元的重要性有两个因素决定: P_k^i 和结构单元的尺寸 S_k^i , 则

$$w_k^i = \frac{S_k^i |P_k^i|}{\sum_{k=0}^{K-1} S_k^i |P_k^i|} \quad (16)$$

3.2 全局的质量评价

定义失真图像的全局结构失真量为 MD , 用均值进行计算全局失真度量, 则

$$MD = \frac{\sum_{i=0}^{N-1} D^i}{N} \quad (17)$$

基于匹配追逐算法的质量评价方法是对图像亮度信息基于能量最大的原则进行结构失真统计, 在实验中发现全局结构失真量 MD 与主观质量并非线性关系, 而呈一个指数关系, 根据韦伯-费克内准则^[20]:

感觉强度与刺激大小的对数成比例的关系,定义 MP_Q 为匹配追逐算法客观评价度量,则

$$MP_Q = \log_{10}(MD) \quad (18)$$

根据算法知道两幅完全相同的图像进行结构失真计算时,有 $MD = 0$,最后得到:

$$\begin{cases} MP_Q = \log_{10}(MD) & MD \neq 0 \\ MP_Q = \text{NULL} & MD = 0 \end{cases} \quad (19)$$

MP_Q 算法的流程图如图 3 所示。

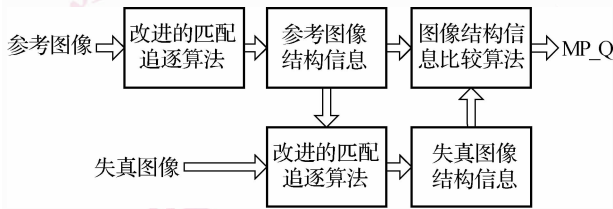


图 3 MP_Q 算法的流程图

Fig. 3 The flowchart of MP_Q

4 实验结果和分析

使用图像数据库为 Sheikh 等人的“LIVE Image Quality Assessment Database Release 2”^[21-23],采用了

最新主观质量 DMOS (difference mean opinion score) 数据^[21-22]。该数据库共有 29 个参考图像,对应每一个参考图像都有 5 种不同参数的失真类型:Fast Fading, Gaussian Blur, White Noise, JPEG 和 JPEG2000, 一共 779 幅失真图像,其中 Fast Fading 是模拟 JPEG2000 压缩图像在无线传输中的误码失真, Gaussian Blur 是对图像进行各种不同参数的高斯模糊处理, White Noise 是在图像中添加不同程度的高斯白噪声。随机选择其中 13 组图像: Bikes, caps, coinsinfountain, flowersonih35, house, lighthouse2, monarch, ocean, paintedhouse, parrots, rapids, sailing1, sailing4, 对应每一个参考图像都经过各种参数的 JPEG 和 JPEG2000 编码,一共 155 幅编码后的图像。参考使用 Video Quality Experts Group (VQEG) Phase I FR-TV test^[24]的性能评价方法。选择了被广泛使用的 PSNR 算法,以及另外两种基于结构失真的评价算法: MSSIM^[3] 和 SVD^[16],与本文匹配追逐算法的实验结果进行了比较,最后对实验结果做了分析。

实验使用的参数为 $m \times l = 32 \times 32, K = 5$,即每个分块提取 5 个 SC。仅对图像亮度进行处理。

4.1 实验结果

图 4 ~ 图 7 中, X 轴分别对应 4 种评价得分, Y

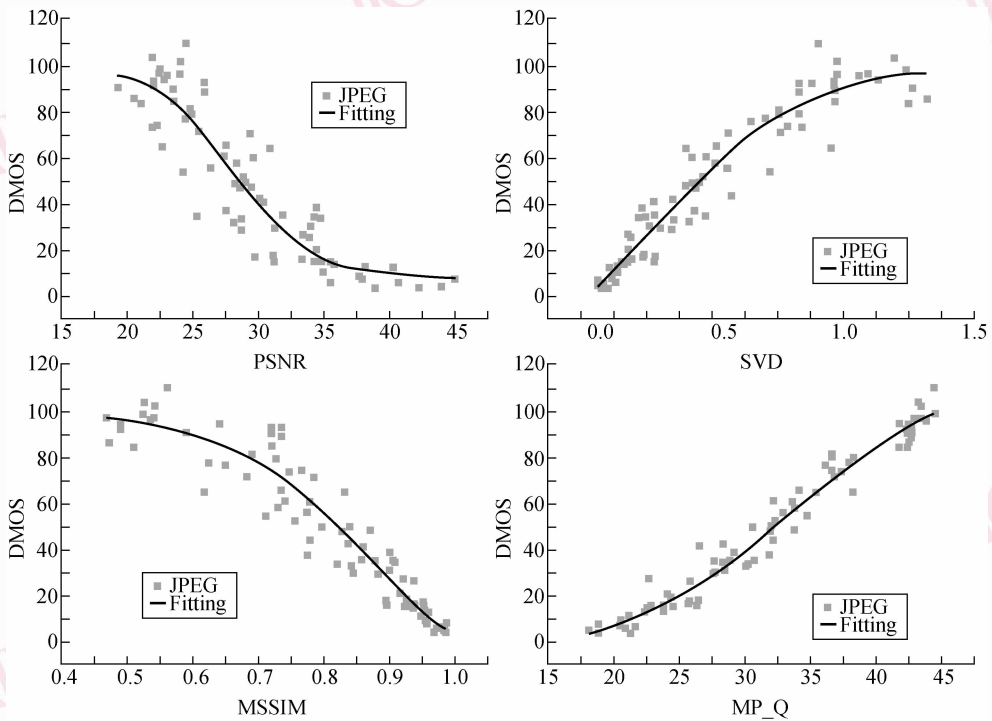


图 4 JPEG 编码下 4 种评价方法与 DMOS 的关系图和拟合曲线

Fig. 4 Scatter plots for each of four metrics for JPEG

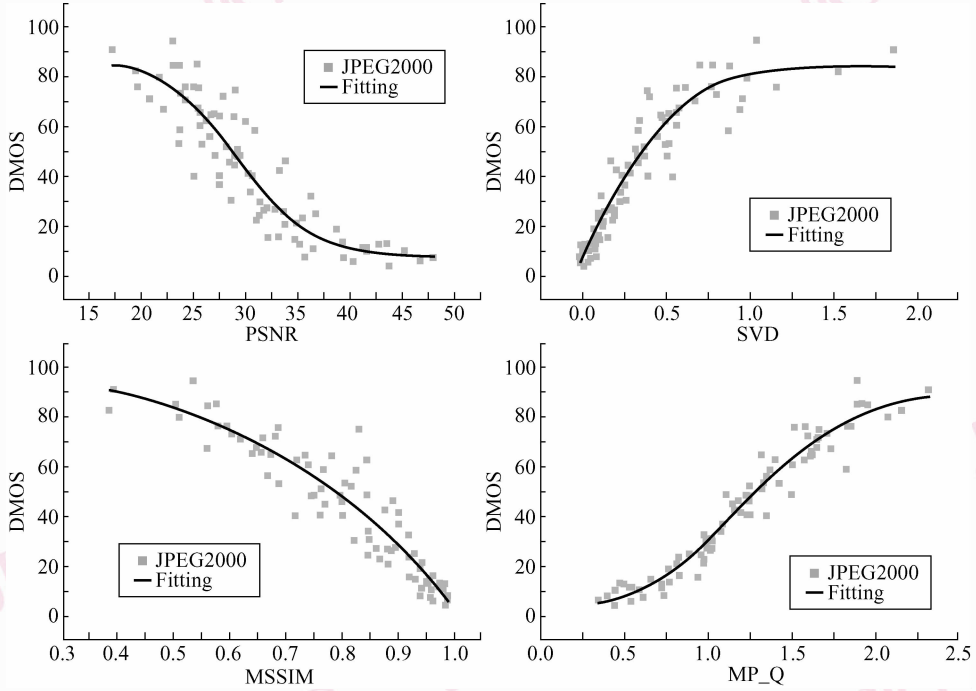


图 5 JPEG2000 编码下 4 种评价方法与 DMOS 的关系图和拟合曲线

Fig. 5 Scatter plots for each of four metrics for JPEG2000

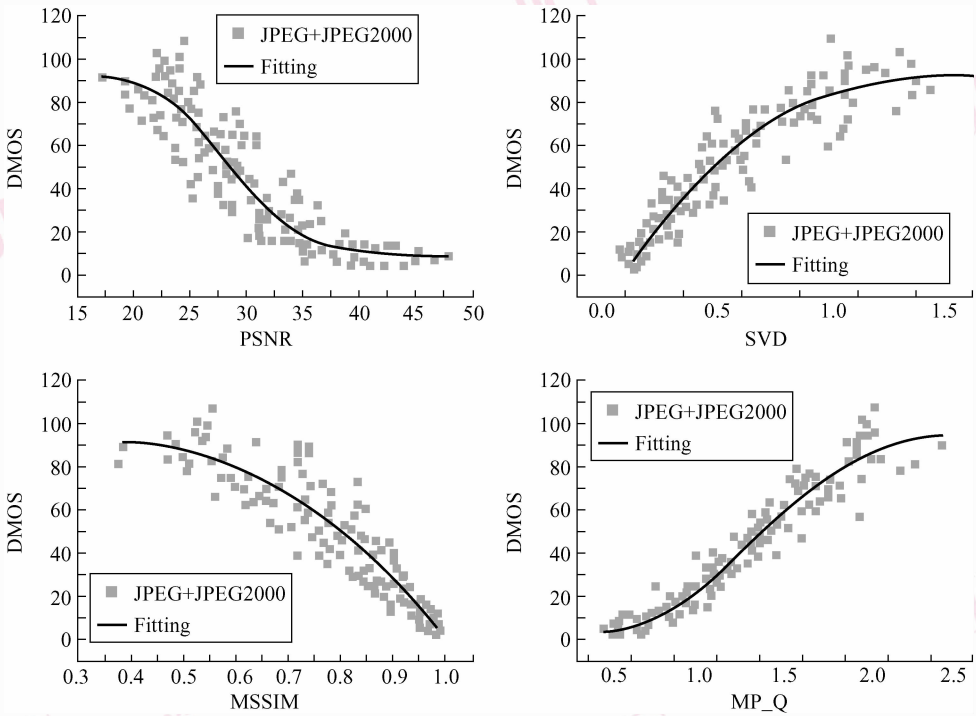


图 6 拟合前 JPEG + JPEG2000 编码下 4 种评价方法与 DMOS 的关系图和拟合曲线

Fig. 6 Scatter plots for each of four metrics for JPEG + JPEG2000

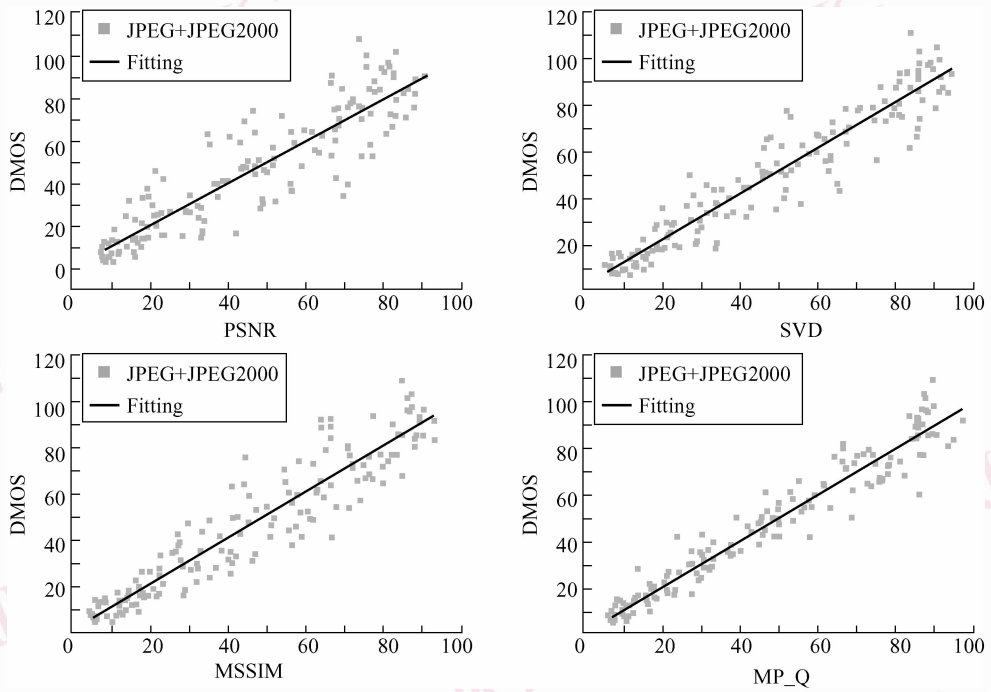


图 7 拟合后 JPEG + JPEG2000 编码下 4 种评价方法与 DMOS 的关系图和拟合曲线
 Fig. 7 Scatter plots for each of four metrics for JPEG + JPEG2000 after a nonlinear fitting

轴对应 DMOS, 图中曲线为非线性拟合曲线, a_1, a_2, a_3, a_4 为拟合参数, 拟合函数为

$$\varphi(x) = a_1 + \frac{a_2}{1 + \exp\left(\frac{x - a_3}{a_4}\right)} \quad (20)$$

可以知道当客观评价与 DMOS 具有越高的相关性时, 对应的客观评价方法就具有越好的性能。

4.2 分析和讨论

从图 4、图 5 中可以看出, 对于 JPEG 和 JPEG2000 而言, SVD 和 MSSIM 具有相当的性能, MP_Q 性能相对更好, 而 PSNR 缺少稳定性。

从图 6、图 7 可以看出, 在拟合前后, MP_Q 的性能相差很小, 而 PSNR、SVD 和 MSSIM 均获得了性能

的提高, 尽管曲线拟合修正在质量评价中是允许的, 但是在实际应用中, 曲线拟合修正无疑会造成一定的限制。在拟合之后, MP_Q 性能依然还是最好, SVD 其次, MSSIM 的性能排在第三。表 1 对比地给出了拟合前后 4 种评价方法与 DMOS 的相关系数和拟合后的 RMSE。在非线性拟合前, MP_Q 相比 PSNR、SVD 和 MSSIM 有较大的优势。非线性拟合之后, SVD 获得最大的性能提高, PSNR 和 MSSIM 也有少量的性能提高。从表 1 中可知, MP_Q 有最好的数据稳定性, SVD 的稳定性其次。

相比于 MSSIM 和 SVD, 本文方法考虑了图像中结构的重要性, 减少了不重要区域对图像质量的影响, 实验结果显示了结构重要性表示和度量的有效性。

表 1 4 种评价方法的相关系数 CC 和均方根差 RMSE
 Tab.1 CC and RMSE – based each of the four metrics

| | 拟合前的 CC | | | 拟合的 CC | | | 拟合后的 RMSE | | |
|-------|---------|----------|---------|---------|----------|---------|-----------|----------|----------|
| | JPEG | JPEG2000 | BOTH | JPEG | JPEG2000 | BOTH | JPEG | JPEG2000 | BOTH |
| PSNR | 0.876 1 | 0.881 5 | 0.864 9 | 0.906 4 | 0.910 9 | 0.901 9 | 13.548 5 | 10.424 8 | 12.623 9 |
| SVD | 0.939 2 | 0.851 2 | 0.901 4 | 0.963 0 | 0.948 8 | 0.951 7 | 8.646 0 | 7.978 0 | 8.968 5 |
| MSSIM | 0.924 6 | 0.914 2 | 0.910 6 | 0.952 6 | 0.935 8 | 0.933 9 | 9.756 9 | 8.906 2 | 10.445 0 |
| MP_Q | 0.981 9 | 0.967 8 | 0.960 6 | 0.984 7 | 0.979 3 | 0.969 2 | 5.592 0 | 5.108 8 | 7.200 3 |

尽管本文算法在实验中仅使用了图像的亮度信息,但实验结果依然显示了很好的性能,如果再考虑加入彩色信息将可以获得更好的结果。本文实验中选择式(15)、式(16)使用两个参数进行加权,考虑到某些特别的应用,优化加权系数可以进一步提高性能。本文算法并不局限在图像质量评价,同样可以用于计算两个一般信号之间的差异,不管是离散或连续,静态图像或者视频信号,都是适用的。

本文算法复杂性绝大部分集中在计算参考图像的结构信息,计算一幅 768×512 的参考图像,计算机平台为 Intel P4 3.0 GHz, 512 M 内存的桌面系统,使用本文实验所选取的参数,需要计算时间为 5 min 左右;计算最后的图像质量值耗时很少,仅需要 0.1 s 左右。参考图像的结构信息可以首先被计算并且保存下来,独立于失真计算。通过优化匹配追逐算法和字典可以提高运算速度,学者对此进行了研究^[25-27],本文下一步工作将集中在研究更适合质量评价需要的字典和高效低复杂度的匹配追逐算法。

5 结 论

本文提出了一种基于结构失真的图像编码质量评价方法,首先利用匹配追逐算法将图像在 2 维可分离的 Gabor 基函数集上投影,按照内积能量最大的原则对图像结构信息进行重要性选择,提取出图像中最重要的结构信息,然后对图像的结构进行特征描述和重要性度量;将图像质量的失真建模为失真前后结构信息的差异,最后计算这种差异得到图像质量的客观评价价值。实验数据显示,MP_Q 比其他 3 种质量评价方法:PSNR, SVD 以及 MSSIM 有更好的性能,与主观评价价值有很好的一致性。

在实际应用中,图像失真类型有很多种,本文仅涉及 JPEG 和 JPEG2000 两种最常见编码失真,下一步的工作将集中在将本文算法推广到更多的失真类型,并且推广到彩色图像和视频中去,更加深入地研究图像结构失真与主观质量的关系。

参考文献 (References)

- Eskicioglu A M. Quality measurement for monochrome compressed images in the past 25 years [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing [C], Istanbul, Turkey, 2000, 4:1907-1910.
- Eskicioglu A M, Fisher P S. A survey of image quality measures for gray scale image compression [A]. In: Proceedings of Space and Earth Science Data Compression Workshop [C], Snowbird, UT, 1993: 49-61.
- Zhou W, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- Eskicioglu A M, Fisher P S. Image quality measures and their performance [J]. IEEE Transactions on Communications, 1995, 43: 2959-2965.
- Eckert M P, Bradley A P. Perceptual quality metrics applied to still image compression [J]. Signal Processing, 1998, 70: 177-200.
- Winkler S. A perceptual distortion metric for digital color video [J]. Proceedings of SPIE, 1999, 3644:175-184.
- Wang Z. Rate Scalable Foveated Image and Video Communications [D]. Austin, Texsx, USA: Deperment of ECE, The University of Texas at Austin, 2001.
- Wang Z, Bovik A C. A universal image quality index [J]. IEEE Signal Processing Letters, 2002, 9: 81-84.
- Wang Z, Bovik A C, Lu L. Why is image quality assessment so difficult? [A]. In: Proceedings of IEEE International Conference on Acoust., Speech, and Signal Processing [C], Orlando, FL, USA, 2002, 4: 3313-3316.
- Teo P C, Heeger D J. Perceptual image distortion [J]. In: Proceedings of SPIE, 1994, 2179:127-141.
- Mannos J L, Sakrison D J, The effects of a visual fidelity criterion on the encoding of images [J]. IEEE Transactions on Information Theory, 1974, 20(4): 525-536.
- Watson A B. DCT quantization matrices visually optimized for individual images[A]. In: Human Vision, Visual Processing, and Digital Display IV [C], Bellingham, WA, USA, 1993, 1913: 202-216.
- Karunasekera S A, Kingsbury N G. A distortion. measure for blocking artifacts in images based on human visual sensitivity [J]. IEEE Transactions on Image Processing, 1995, 4(6):713-724.
- Chou C H, Li Y C. A perceptually tuned subband image coder based on the measure of Just-Noticeable-Distortion profile [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, 5(6): 467-476.
- Mayache A, Eude T, Cherifi H. A comparison of image quality models and metrics based on human visual sensitivity [A]. In: Proceedings of IEEE International Conference on Image Processing[C], Chicago, IL, USA, 1998, 13:409-413.
- Li J L, Chen G, Chi Z R, et al. A new objective assessment method for image coding quality[J]. Journal of Image and Graphics, 2004, 11(9): 1348-1355. [李均利, 陈刚, 池哲儒等. 客观评价图像质量编码新方法 [J]. 中国图象图形学报 2004, 11(9): 1348-1355.]
- Shnayderman A, Gusev A, Eskicioglu A M. A multidimensional image quality measure using singular value decomposition [J]. IEEE Transactions on Image Processing, 2006, 15(2): 422-429.
- Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries

- [J]. IEEE Transactions on Signal Processing, 1993, **41**(12): 3397-3415.
- 19 Neff R, Zakhor A. Very low bit rate video coding based on matching pursuits [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, **7**(1):158-171.
- 20 Levine M W. Fundamentals of Sensation and Perception (3rd ed.) [M]. New York, NY, USA: Oxford University Press, 2000.
- 21 Sheikh H R, Wang Z, Cormack L, Bovik A C. LIVE Image Quality Assessment Database Release 2 [EB/OL]. <http://live.ece.utexas.edu/research/quality>, 2009-02-06.
- 22 Sheikh H R, Sabir M F, Bovik A C. A statistical evaluation of recent full reference image quality assessment algorithms [J]. IEEE Transactions on Image Processing, 2006, **15**(11): 3441-3452.
- 23 Sheikh H R, Bovik A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, **15**(2): 430-444.
- 24 VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment [R/OL]. <http://www.vqeg.org>, 2000-03.
- 25 Neff R, Zakhor A. Matching pursuits video coding I dictionary approximation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, **12**(1):13-26.
- 26 Neff R, Zakhor A. Matching pursuits video coding I operational models for rate and distortion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, **12**(1):27-39.
- 27 Vleeschouwer C D, Macq B. New dictionaries for matching pursuit video coding [A]. In Proceedings of IEEE International Conference on Image Processing [C], Chicago, IL, USA, 1998, **13**:764-768.