

基于分形维数的图像质量客观评价方法研究

庞全 王振华 耿丽硕 范影乐

(杭州电子科技大学生物医学工程与仪器研究所, 杭州 310018)

摘要 建立图像质量客观评价模型对于图像编码、增强、重建以及分析等领域具有重要的现实意义。鉴于传统的图像质量评价方法的评价结果与主观感知存在较大的误差等缺陷,为此从分形角度考虑,并兼顾人眼视觉特性,首先提取了分形维数作为图像质量的评价指标;然后从非线性角度来表征引起人眼视觉敏感变化的图像亮度以及纹理信息,并将能准确反映图像质量变化的空隙度参数作为有效补偿;最后采用线性回归分析直接对图像进行建模,并将分形维数差值和空隙度差值两分量表示在统一的模型中。实验证明,相对于传统的 PSNR 和 SSIM 评价指标而言,该评价模型不仅对于不同类型的失真、相同失真类型的不同失真级别的图像能够准确进行评估,而且与主观评价值(MOS)具有更好的关联性,即与人眼视觉感受具有较高的吻合性,同时能够实现对图像质量进行全面、科学的评价。

关键词 质量评价 分形维数 空隙度 结构相似法 峰值信噪比

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2009)04-0657-06

A Criterion of Objectively Assessing Image Quality Based on Fractal Dimension

PANG Quan, WANG Zhen-hua, GENG Li-suo, FAN Ying-le

(Biomedical Engineering & Instrument Institute, Hangzhou Dianzi University, Hangzhou 310018)

Abstract It is very urgent to build an evaluation criterion in image compression, enhancement, restoration and analysis fields. In this paper, fractal dimension is utilized to be an image quality assessment index and describe image luminance and texture characteristics from non-linear aspects. Lacunarity changes can also describe image quality. Therefore fractal dimension difference and lacunarity difference are integrated into one single mathematics model by linear regression analysis. Experiment results show that against traditional model such as PSNR, SSIM and so on, the proposed approach can not only assess different distortion types, same distortion type with different distortion levers accurately, but also has stronger correlation with MOS, more agreement with the perceptual of human beings, and can assess image quality accurately and effectively.

Keywords quality assessment, fractal dimension, lacunarity, SSIM, PSNR

1 引言

当前,图像压缩技术在多媒体和计算机网络等需求的推动下被广泛应用于广播电视、视频会议和数据传真等方面,但由于大数据量、高压缩比的要求

使图像编码过程不可避免地丢弃一定数量的图像信息,从而导致恢复图像的失真。此外,图像存储、传输等过程同样可导致图像降级。因此,进行图像质量评价方法的研究,对于图像编码算法的选择以及图像系统性能的评定具有重要的现实意义^[1],已成为近期图像处理领域新的研究热点。

基金项目:国家自然科学基金项目(60872090);浙江省重大科技专项(2008C01015-2);浙江省新苗人才计划项目(2008R40G2040120)

收稿日期:2007-08-06;改回日期:2007-11-13

第一作者简介:庞全(1951~),男,教授,博士生导师。1986年于西安交通大学获控制科学与工程专业硕士学位。主要研究方向为图像处理、模式识别。E-mail:pangquan@hdu.edu.cn

图像质量评价分为主观质量评价和客观质量评价^[2]两类。主观质量评价方法由于受观察者环境、情绪等主观因素的影响,稳定性差,同时,由于此方法具有劳动强度大以及费时等缺点,因此实际可操作性较差。传统的客观质量评价方法主要采用最小均方误差(MSE)和峰值信噪比(PSNR)等质量评估指标,由于此方法未考虑图像结构内容信息,因此与人的主观感受不符;目前基于人类视觉系统(HVS)的图像质量评价方法的研究较多^[3],但其均基于特定的假设条件,另外,由于对人类高度复杂和非线性特性的认知系统的机理研究非常有限,而且HVS算法复杂,因此并不适合于实际应用;基于结构相似性(SSIM)的图像质量评价方法^[4]是利用亮度、对比度和结构信息的幂乘积作为测度,但忽略了3者的非线性映射关系,易产生误差;基于分形编码的方法是采用加权平均拼贴误差来进行度量,其对于块噪声的失真图像的评价结果并不比PSNR好。因此,迫切需要有效的客观质量评价方法,使评价过程既简单易行,又符合人眼的主观感受。

人眼视觉系统的一个基本特征是局部对比的敏感性,视觉只对视场中亮度或纹理发生显著变化的区域感兴趣,尤其是对图像边缘、轮廓信息的失真以及中高亮度背景中的纹理细节的变化较为敏感^[5]。由于自然界绝大多数自然景物具有分形特征,因此分形维数具有从非线性角度表征图像纹理的粗糙程度以及模式复杂性信息的特点;而对于人造景物,将提取的人造图像的分形特征应用于质量评价同样是可行和有效的。但同时,由于分形维数不能描述图像纹理变化的快慢,因此很多视觉差别很大的图像具有相似的分形维数。而空隙度作为高阶的分形特征,由于其可以描述相同分形维数下图像质量变化的快慢,因此可将空隙度作为分形维数的有效补偿。考虑到分形维数和空隙度的变化可以精确地表征图像边缘信息的失真,特别是图像纹理细节的变化,因此本文采用基于分形维数理论的方法,并将分形维数和空隙度作为评价指标来建立图像质量的评价模型。

2 基本原理

2.1 分形维数

1975年Mandelbrot首次提出了分形及分形维数的概念,分形维数作为分形的重要特征和度量,可

把图像的空间信息和灰度信息简单而又有机地结合起来。针对不同的研究对象,可采用不同的描述方法计算其分形维数。其中计盒维(differential box-counting, DBC)可在保持良好的精确性和适用性的同时,还能满足动态特性和计算效率方面的要求,因此本文选用计盒(box)维数来进行图像表面粗糙程度的描述^[6]。

设 F 是 \mathbf{R}^n 中任一非空有界子集,记 $N_\delta(F)$ 表示边长为 δ ,且能覆盖集合 F 的最少分块数,若存在

$$D_{\text{box}} = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{\log(1/\delta)} \quad (1)$$

则称这个极限值为集合 F 的分块维数。

对于大小为 $M \times M$ 的图像 S ,其灰度集合为 $I = \{I(i, j), 1 \leq i, j \leq M\}$,本文将3维空间 (x, y, z) 引入 I 中,其中 (x, y) 为图像像素点的平面坐标, z 即代表图像在点 (x, y) 处的灰度值 $I(x, y)$ 。 $N_\delta(F)$ 的计算方法如下:先将大小为 $M \times M$ 的图像分割成大小为 $r \times r$ 的子块($1 < r \leq M/2, r$ 为整数),然后 xy 平面被分割成多个 $r \times r$ 大小的网格。在每个网格上是一列 $r \times r \times r$ 的盒子柱。若图像中坐标为 (i, j) 的小格所包括的灰度最大值和最小值分别落在盒子上面编号为 l 和 k 的盒子内部,则

$$n_r(i, j) = l - k + 1 \quad (2)$$

式中, $n_r(i, j)$ 为覆盖坐标为 (i, j) 的尺度为 r 的网格中的图像所需的盒子数,而覆盖整个图像所需的盒子数为

$$N_r = \sum_{i, j} n_r(i, j) \quad (3)$$

若对 r 取不同的值,并计算相应的 N_r ,则盒维数的计算公式如下:

$$D = \log N_r / \log(1/r) \quad (4)$$

2.2 空隙度

当两个表面或结构完全不同的分形集有相同或相似的分形维数时,则分形维已无法进行区分。分形维数只反映物体表面的不规则程度,却无法反映表面变化的快慢,而空隙度(lacunarity)则能够描述物体表面质量变化的快慢,当图像光滑时,其值较小,当图像粗糙时,其值较大。空隙度定义为

$$\Lambda = E \left[\frac{Q}{E(Q)} - 1 \right]^2 \quad (5)$$

式中, Q 为分形集的“质量”, $E(Q)$ 是“质量”期望值。空隙度是一个二阶统计量,而分形集的密度则与测量尺度 L 有关,且服从 $\rho(L) = KL^D$,其中 ρ 是分形集的密度, K, D 为常数。

为了简化空隙度的计算过程,本文采用 Keller 等人提出的改进空隙度计算的方法^[7]进行了计算。对一个数字图像 S ,若用尺度为 L 的盒子进行覆盖,则空隙度为

$$C(L) = \frac{M(L) - N(L)}{M(L) + N(L)} \quad (6)$$

式中, $N(L)$ 等于覆盖图像 S 所需的盒子数除以图像 S 中所含像素点的总数,即每个像素点平均所占的盒子数。设图像 S 中所含像素点的总数为 P ,则

$$M(L) = \frac{P}{N_r}, N(L) = \frac{N_r}{P}, \text{代入式(6)可得}$$

$$C(L) = \frac{P^2 - N_r^2}{P^2 + N_r^2} \quad (7)$$

3 图像质量评价方案及实验结果分析

本文采用美国 UTEXAS 大学的 LIVE (laboratory for image & video engineering) 标准图像质量评价数据库 (quality assessment database)^[8],该库对多种不同类型失真的图像进行了主观打分,即先由 29 名观察者根据主观感受对每一幅图像给出分数值 (范围:1 ~ 100),然后取平均后作为主观打分值,并分为 5 个等级 (差,较差,一般,好,非常好)。

3.1 单一分形维数评价

单一分形维数评价步骤如下:

(1) 人眼对激励的反应依赖于相对光强,即亮度的相对变化,本文将提取的亮度信息作为图像空间区域内的像素灰度值 $I(x, y)$ 。

(2) 由于 UTEXAS 大学的质量评价数据库中不同类型图像的像素点总数不同,并非所有图像均为 512×512 或 256×256 标准大小,因此在实际计算分形维数时,可以 64×64 pixel 为一个单元进行模块划分。对于图像最右方和最底部的不足 64×64 大小的模块,可将其分别并入到其相邻左方、相邻上方的模块当中,以 $r = 2, 4, 8, 16, 32$ 等尺度变化先计算盒数,再根据式(4)得出分形维值,以此即可求得整幅图像的分形维数矩阵。

(3) 计算原始图像和降质图像的分形维数矩阵差值,并通过对该差值矩阵的所有元素求和来得到 d_{fractal} 。

图 1 为原始 stream 图像及其 3 幅不同级别失真的图像,其中图 1(a) 是原始 stream 图像,剩余 3 幅为不同级别的失真图像,从主观感受可以判断,

图 1(b) ~ 图 1(d) 质量依次变差。表 1 为用单一的分形维数差值 d_{fractal} 对图 1 中的 4 幅图像质量进行评价的结果,由于图像是 2 维几何体,因此分形维数值介于 2 和 3 之间;主观评价 (MOS) 值依次增大,反映了图像质量依次变差;单一的分形维数差值同样依次变大,其虽能够反映图像质量依次变差,但仍有不足之处。例如图 1(c) 和图 1(d) 相比,图 1(d) 的主观评价质量较差,且与图 1(c) 相差悬殊,分形维数差值 d_{fractal} 虽能够表征图 1(d) 较图 1(c) 质量差,但两者的分形维数差值 d_{fractal} 较相近,不能有效地进行分辨。究其原因在于不同的图像纹理可能会具有相似的分形维数,不同的纹理方向也可能具有相似的分形维数,单一的分形维数无法区分单一分形集和多重分形集,因此,在这种情况下,急需引进另一参数用来改善评估结果,以提高可分辨能力。

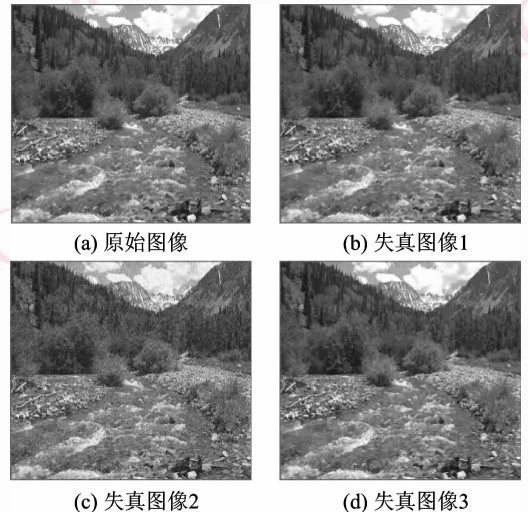


图 1 Stream 原始图像及其失真图像
Fig. 1 Stream original and distortion images

表 1 单一分形维数评价结果

Tab. 1 The results of single fractal dimension assessing

评估指标	图 1(a)	图 1(b)	图 1(c)	图 1(d)
分形维数	2.318 3	2.203 4	2.638 4	2.621 4
分形维数差值 (d_{fractal})	0	9.613 7	18.674 2	18.751 8
MOS 值	0	44.397 1	48.862 3	63.649 0

3.2 FL 模型评价

在单一分形维数评估方法的基础上,引进了空隙度这一特征量,评价步骤如下:

(1) 计算图像的空隙度。根据实验,先选取测

量尺度 $L=4$,再由式(7)计算空隙度 $C(L)$ 。

(2) 建立图像质量综合评价模型——FL(fractal and lacunarity)模型。该模型取两个参数的加权累加和作为最后评价结果,即用 FL 值表示图像质量失真程度, d_{fractal} 表示分形维数差值, $d_{\text{lacunarity}}$ 为空隙度差值, ω_1 、 ω_2 分别为加权系数。

$$FL = \omega_1 d_{\text{fractal}} + \omega_2 d_{\text{lacunarity}} \quad (8)$$

由式(8)可知,FL 值越大,图像质量越差,当降质图像与原始图像完全相同时,FL 取值为 0。

图 2 为原始 Buildings 图像及其待评价图像,若分别运用不同的评价指标对其进行评价,则其评价结果如表 2 所示。

根据评估指标的定义不同,PSNR 值越大,信噪比越高,说明图像质量越好,其对参考图像的评价值为 INF(infinity);用 SSIM 指标评价图像时,SSIM 值越小,说明待评估图像和参考图像越接近,当待评估图像为参考图像时,值为 0;若 d_{fractal} 值小,则降质图像与原图像的分形维数差值越小;若本文方法的 FL 值越小,则待评估图像的质量越好,而参考图像的 FL 评估值为 0;MOS 值为图像库中的主观打分值,参考图像的 MOS 值为 0。

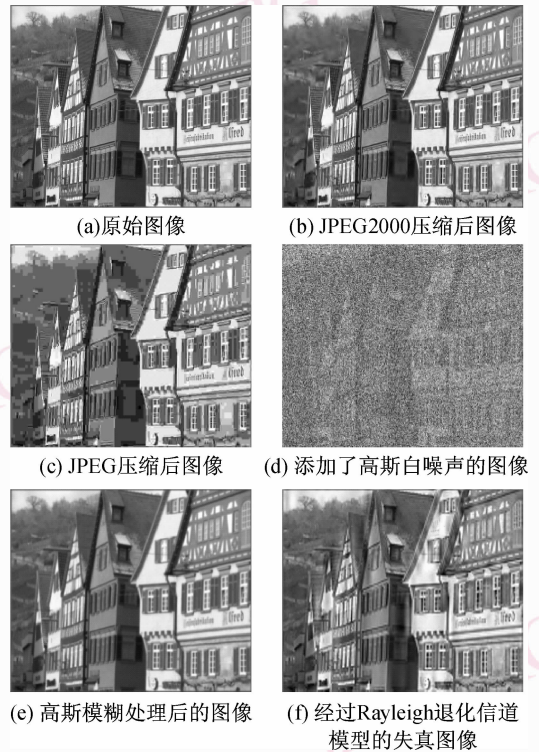


图 2 Buildings 原始图像及不同类型降质图像

Fig. 2 Buildings original and different distortion types images

表 2 本文评价指标与其他评价指标对图像不同失真类型进行评价的结果对照

Tab. 2 Comparison between FL and other measures on different distortion types images

评估指标	不同失真类型图像					
	原图像	JPEG2000 压缩图像	JPEG 压缩图像	添加高斯白噪声 的图像	高斯模糊处理 后的图像	Rayleigh 快速退化 的图像
PSNR 值 (dB)	INF	31.895 2	29.863 3	27.431 7	29.863 3	29.398 8
SSIM 值	1	0.631 1	0.490 5	0.030 3	0.515 0	0.440 7
d_{fractal}	0	0.523 9	0.531 7	0.891 2	0.570 3	0.571 1
FL 值	0	0.480 4	0.596 2	0.880 6	0.558 5	0.561 6
MOS 值	0	49.581 1	59.013 1	72.062 5	56.175 4	56.515 1

表 2 给出了 FL 模型和其他评价指标对不同失真类型的 Buildings 图像进行评价的结果。其中 ω_1 决定了 d_{fractal} 值在 FL 模型的评价结果中所占的比例, ω_1 取值越大, d_{fractal} 值的作用越明显。但同时由于 d_{fractal} 值变化较快,急需通过 $d_{\text{lacunarity}}$ 来调整,因此 ω_2 值同 FL 模型联系密切,实验结果表明,参数 ω_1 和 ω_2 的取值范围一般分别介于 $0 \sim 3.5$ 和 $-1 \sim 0$ 之间。在实际应用中,需兼顾上述因素,本文基于线性回归方法对大量样本数据进行了实验,并设定预测模型参数 ω_1 和 ω_2 取值分别为 1.834 3 和

$-0.509 7$ 。通过对图 2 的主观判断可见,图 2(c) 的图像质量较图 2(f) 差,这与 MOS 值的评价结果比较相符(图 2(c) 的 MOS 值为 59.013 1,图 2(f) 的 MOS 值为 56.515 1),而 PSNR 值却呈现相反的评价结果,图 2(c) 的 PSNR 值为 29.863 3 dB,图 2(f) 为 29.398 8 dB,这表明图 2(c) 图像质量优于图 2(f),这与主观判断结果相背离。图 2(c) 和图 2(f) 的 SSIM 值分别为 0.490 5 和 0.440 7,即图 2(c) 的 SSIM 值更接近 1,这表明图 2(c) 的图像质量优于图 2(f),这与 MOS 值不相符。 d_{fractal} 虽然可以大体描述

图像的失真情况,但未能将图 2(b) 和图 2(c) 的失真级别明显区分开来 ($d_{fractal}$ 值分别为 0.523 9 和 0.531 7),这与 MOS 值有一定差距。而且从表 2 中还可以看出,通过空隙度参数的引入,FL 模型评价方法不仅可以解决不同失真情况但具有相同分形维数差值的问题,而且能够对相同维数不同程度的失真图像进行正确评价,同时,与 PSNR 和 SSIM 评估指标相比而言,本文 FL 模型评价方法不仅可以对不同类型失真的图像质量进行正确评估,且评估结果与 MOS 值具有非常高的吻合度。

为了验证 FL 评价指标的有效性,本文针对同一幅图像的不同失真级别,采用不同评价指标进行了评估对比。图 3 为 womanhat 5 级失真图像,失真程度从左至右越来越严重,失真类型为 Rayleigh 快速退化失真。表 3 为分别采用 PSNR 值、SSIM 值、

FL 值和 MOS 值评价指标对 Womanhat 5 级失真图像进行质量评测的结果。

表 3 的实验结果中,PSNR 评价指标对 II 级失真图像的评测结果未能与 MOS 值相符,即对于失真度较低的图像,利用 PSNR 评价指标无法正确反映图像的信息;同样,SSIM 指标对 V 级失真的评测结果也与 MOS 值有偏差,即 SSIM 评价指标对失真度较高的图像存在评估误差;而从表 3 评价结果的分析可看出,本文的 FL 值评价指标对同一图像的不同级别失真的评估非常有效,FL 值反映的失真程度与 MOS 值基本相符。因此,对于同一图像不同级别失真的评估,相比 PSNR 值与 SSIM 值评价指标,本文的 FL 模型能够更加直观有效地描述图像的失真程度。

表 3 本文质量评价指标与其他评价指标对不同失真级别图像进行评价的结果对照

Tab. 3 Comparison between FL and other measures on same distortion type with different distortion levels

评估指标	不同失真级别图像				
	I	II	III	IV	V
PSNR 值	40.894 9	42.344 7	37.582 2	36.312 6	35.732 6
SSIM 值	0.758 5	0.796 6	0.822 3	0.935 2	0.897 0
FL 值	0.247 7	0.250 1	0.338 8	0.373 6	0.437 4
MOS 值	27.496 5	30.551 6	40.450 0	44.673 2	47.997 0

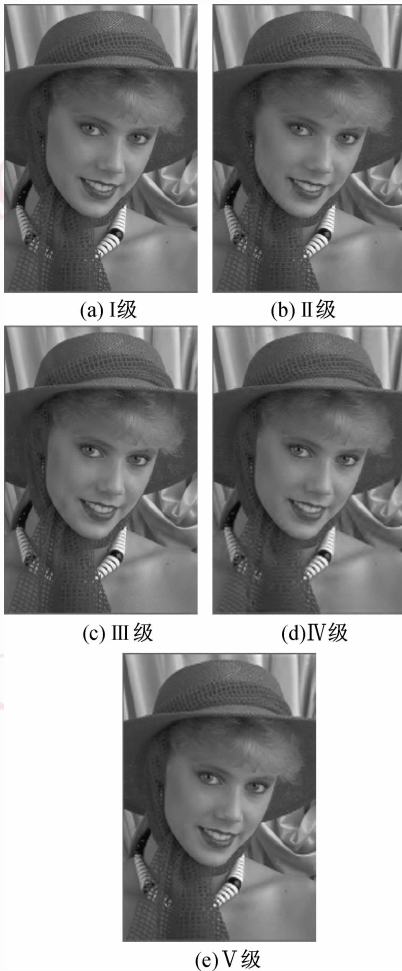


图 3 Womanhat 5 级失真图像

Fig. 3 Womanhat distortion images with 5 levels

图 4 给出了 PSNR、SSIM 和 FL 评价指标分别与主观评价(MOS)值的对应关系图,横坐标为归一化的客观评价价值,纵坐标为 MOS 值。由于归一化后的 PSNR 值 ($PSNR_{uni}$) 和归一化后的 SSIM 值 ($SSIM_{uni}$) 的取值越大, MOS 值越小,因此为方便观察,本文将 $1 - PSNR_{uni}$ 和 $1 - SSIM_{uni}$ 值作为横坐标。由图 4 可以看出,PSNR 评价指标的计算结果与 MOS 值有较大偏差,不仅“野点”较多,而且与主观感受差距较大,评价性能低劣;尽管 SSIM 评价指标相比 PSNR 评价指标更接近 MOS 值,但仍有一些与 MOS 值不符的数据点存在;而本文提出的 FL 模型计算的数据斜率更接近于 1,与 PSNR 和 SSIM 指标的评价结果相比,同 MOS 值的关联性有较大提高,不仅“野点”明显得到校正,而且更能有效地描述图像质量,与人的主观感受一致。

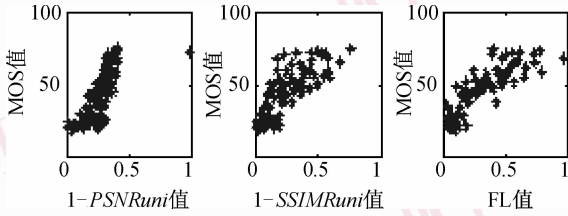


图 4 多种客观评价指标值与主观评价值的对应关系图

Fig. 4 The relative fig of objectively assessing results and MOS

4 结 论

通过建立质量评价模型来对视频图像进行正确评估已经成为图像处理领域一个极具挑战性的课题。传统的 PSNR 和 MSE 指标的评测结果与主观感受相背离;HVS 评价指标不仅计算复杂,且要基于诸多假设条件;SSIM 指标也不能取得较好的评价效果;本文从人眼视敏感度的角度出发,将分形维数和缝隙度作为图像的质量评价指标,通过实验与 PSNR 和 SSIM 等评价指标进行比较,结果表明,无论对于不同类型的图像失真,还是对同一幅图像的不同级别的失真,与 PSNR 和 SSIM 评价指标相比,本文所采用的 FL 质量评价模型均能够直观而准确地进行评价,并可得到与人眼完全一致的评价结果,同时通过 3 种方法与

MOS 值的对应关系的对照,显而易见,本文 FL 模型方法的评价结果与 MOS 值的关联度更高,更能符合人眼的主观感受。

参考文献 (References)

- 1 Zhou Wang, Bovik A C, Lu Li-gang. Why is image quality assessment so difficult? [J]. *Acoustics, Speech and Signal Processing*, 2002, **4**(8): 3313-3316.
- 2 Lin D C, Chau P M. Objective human visual system based video quality assessment metric for low bit-rate video communication systems [A]. In: *Proceedings of IEEE 8th Workshop on Multimedia Signal Processing*[C], Victoria Canada, 2006:320-323.
- 3 Sheikh H R, Bovik A C. Image information and visual quality [J]. *IEEE Transactions on Image Processing*, 2006, **15**(2): 430-444.
- 4 Zhou Wang, 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.
- 5 Winkler S. *Vision Models and Quality Metrics for Image Processing Applications* [EB/OL]. <http://stefan.winkler.net/publications.html>, 2004-09-06.
- 6 Hyun K J, Chang K S, Jin K T. Fractal Dimension Co-occurrence Matrix Method for Texture Classification [A]. In: *Proceedings of IEEE 10th Region Conference on Communications*[C], Hong Kong, China, 2006:1-4.
- 7 Keller J M, Chen S. Texture description and segmentation through fractal geometry [J]. *Computer Vision, Graphics and Image Processing*, 1989, **45**(2): 150-166.
- 8 LIVE- release2 _ database [DB/OL]. <http://live.ece.utexas.edu/index.htm>, 2007-07-06.