

基于人眼视觉特性的彩色图像质量评价

付 伟, 顾晓东, 汪源源

(复旦大学 电子工程系, 上海 200433)

摘 要: 图像处理系统的性能优劣的评判往往需要一个合理迅速的图像质量评价算法作为支撑. 传统的图像质量评价算法由于没有充分考虑人眼的视觉特性, 使得质量评价结果与实际图像的人眼感知质量不符. 根据人眼对图像边缘信息非常敏感这一人眼视觉特性, 提出一种综合图像边缘和背景相似度的算法(ÉBS)来评价彩色图像质量, 即通过比较失真彩色图像与原始参考图像的边缘以及除边缘之外的背景相似程度最终确定失真图像的质量. 应用于由 779 幅包含五种类型失真的图像质量评价库的实验结果表明, 该算法的评价结果相比 PSNR, MSSIM, IFC 以及基于像素域的 VIF 等算法与图像的主观评价结果(由 DMOS 值表示——将背景不同的一组观察者对失真图像的评分进行统计平均后所得到的评价结果)更一致, 也即该算法的评价结果更接近图像的实际视觉感知质量.

关键词: 图像质量评价; 人眼视觉特性; 图像边缘相似度

中图分类号: TN911.73

文献标识码: A

文章编号: 1000-7180(2010)02-0059-05

Color Image Quality Assessment Based on HVS

FU Wei, GU Xiao-dong, WANG Yuan-yuan

(Department of Electronic Engineering, Fudan University, Shanghai 200433, China)

Abstract: Measurement of visual quality is of fundamental importance to some image processing applications, so an efficient image quality assessment algorithm is greatly needed. And the evaluation results of the conventional algorithms do not match to the ones of human vision because those algorithms ignore the characteristic of the Human Visual System (HVS). According to the HVS that human have strong sensitivity to the image edge information, this paper propose a novel algorithm for image quality assessment based on the edge and background similarity between the distorted image and the reference (perfect) image. We test the validation of the proposed algorithm on the database of 779 images with five different distortion types, the results show our method has better performance than PSNR, MSSIM, IFC and VIF, because the evaluation results are closer to the ones of human vision.

Key words: image quality assessment; HVS; image edge similarity

1 引言

图像质量评价是图像信息处理中十分重要的研究内容, 在有关图像处理的几乎所有领域中, 诸如图像压缩编码、信息隐藏、图像增强、图像传输和图像恢复等均会使图像遭受一定程度的失真, 所以需要有一个合理的图像质量评价标准来评估各图像处理算法的性能. 由于人眼是图像的最终观测者, 所以主观图像质量评价方法无疑是唯一最准确的方法, 也

即组织一批不同知识背景的观察者, 对失真图像按照失真程度进行打分, 最后对得分进行平均, 获得评价结果. 然而这种方法非常耗时耗力所以通常不被采用.

传统的客观评价方法(如 PSNR)由于物理意义明显, 计算非常简单而被广泛应用, 然而它完全没有考虑人眼的视觉特性, 从而与实际的图像感知质量不是很匹配. 近年来, 许多研究学者们对如何迅速准确的评价失真图像的质量进行了大量的研究, 目标

收稿日期: 2009-01-15; 修回日期: 2009-03-05

基金项目: 国家自然科学基金项目(60671062, 60571052); 国家重点基础研究发展计划项目(2005CB724303)

就是设计新的图像质量测度算法使其评价结果最大限度的趋近人眼的主观图像质量评价结果。

相比传统的 PSNR 评价方法,利用人眼视觉特性(HVS)来提高图像质量的评价性能显然是更有效的,前人在这方面已有大量研究^[1-2]。这类算法通常将图像通过信道分解为各子带,在各子带中利用各种已有的人眼视觉模型(如频率选择性,对比灵敏度函数 CSF)加权平均原始参考图像与失真参考图像的信号差值得到失真图像的质量测度。然而已有的各种人眼视觉模型均由简单实验确定,实际的图像则具有非常复杂的纹理结构,所以该类算法所得到的图像质量测度与实际的图像视觉质量还未能很好的吻合,并且该类算法计算相对比较复杂。

文献[3]中基于结构相似度的概念,提出了一种新的图像质量评价算法 MSSIM,取得了比较好的图像质量评价效果,且算法比较简单^[3]。该算法将两幅图像的相似度分为三个部分:亮度相似度、对比度相似度及结构相似度。根据失真图像与原始参考图像(完美图像)的相似程度来评定图像质量,与参考图像相似度越高,则认为失真图像质量越好。然而该算法对模糊失真类图像及严重失真图像的质量评价则不是很有效。根据自然图像的信息统计特性,文献[4-5]等通过计算失真图像与参考图像的互信息相关性相继提出了 IFC 和 VIF 的图像质量评价方法^[4-5]。该方法评价性能很好,但由于涉及到小波分解计算,非常复杂费时。

文中基于彩色图像的边缘及背景相似度,提出一种新的图像质量测度(Edge and Background Similarity),根据 EBS 值(0—1)的大小来评价图像质量。新方法在性能上要优于 PSNR, MSSIM, IFC 以及基于像素域的 VIF 方法,而时间消费则要远低于基于小波域的 IFC 方法。

2 原理

人类识别物体需依赖于物体的边缘。对于图像许多研究表明边缘是图像认知的重要特征,人眼对于图像的边缘纹理信息最为敏感和重视,也即边缘纹理信息很有可能是图像结构信息的最重要的部分。因而失真图像边缘的失真程度将直接影响失真图像的人眼视觉感知质量。

如图 1 所示,失真的 Peppers 图像(图 1(b))与原始图像(图 1(a))具有相同的边缘梯度,显然两图像存在很高的相似度。然而失真的 Peppers 图像(图 1(b))的 PSNR, MSSIM, IFC, VIF 值(值越小说明

与原始图像相差越大,从而说明图像质量越差)却显示其严重降质或者根本就是一个毫无意义的值如 MSSIM, VIF 等。可见这些图像质量评价算法忽视了图像边缘失真对图像视觉感知质量的影响,有其局限性。基于这一特性,文中提出通过图像边缘综合图像背景相似度来评价图像质量的方法。

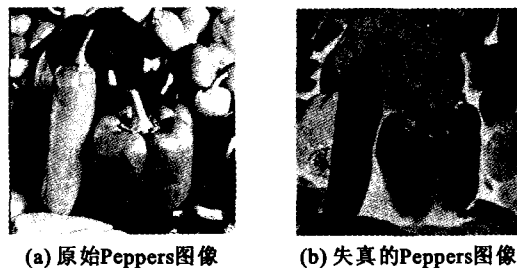


图 1 失真图像质量测度

PSNR=6.79, MSSIM=-0.07, IFC, VIF=0

2.1 彩色图像的边缘提取

要计算边缘相似度,则首先要提取图像的边缘。由于梯度能较好地反应图像的边缘纹理信息且计算简单,所以计算梯度是在图像中检测边缘的最常用方法之一。图 2 是用来计算梯度图像的 Sobel 掩模算子,梯度图像则通过 Sobel 掩模算子滤波后的图像的绝对值求和获得。对于 RGB 彩色图像,通常是利用 Sobel 掩模算子计算每个彩色分量图像的梯度,然后合并该结果得到彩色图像边缘。为了避免由于某一色彩分量通道所得较大梯度边缘对其他通道所得较小梯度边缘的掩盖造成边缘检测精度不高的问题,文中采用下面的方法直接在 RGB 向量空间中提取彩色图像边缘。

-1	-2	-1
0	0	0
1	2	1

(a) 垂直方向

-1	0	1
-2	0	2
-1	0	1

(b) 水平方向

图 2 计算梯度图像的 Sobel 掩模算子

令 $c(x, y)$ 代表坐标 (x, y) 处的 RGB 彩色空间中的向量:

$$c(x, y) = \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix} \quad (1)$$

其中对一个大小为 $M \times N$ 的图像来说,有 MN 个这样的向量 $c(x, y)$ 。

且令 r, g 和 b 是 RGB 彩色空间沿 R, G 和 B 轴

的单位向量,并定义向量

$$\mathbf{u} = \frac{\partial R}{\partial x} \mathbf{r} + \frac{\partial G}{\partial x} \mathbf{g} + \frac{\partial B}{\partial x} \mathbf{b} \quad (2)$$

及

$$\mathbf{v} = \frac{\partial R}{\partial y} \mathbf{r} + \frac{\partial G}{\partial y} \mathbf{g} + \frac{\partial B}{\partial y} \mathbf{b} \quad (3)$$

则这些向量的点积 g_{xx} , g_{yy} 和 g_{xy} 可如下表示:

$$g_{xx} = \mathbf{u} \cdot \mathbf{u} = \mathbf{u}^T \mathbf{u} = \left| \frac{\partial R}{\partial x} \right|^2 + \left| \frac{\partial G}{\partial x} \right|^2 + \left| \frac{\partial B}{\partial x} \right|^2 \quad (4)$$

$$g_{yy} = \mathbf{v} \cdot \mathbf{v} = \mathbf{v}^T \mathbf{v} = \left| \frac{\partial R}{\partial y} \right|^2 + \left| \frac{\partial G}{\partial y} \right|^2 + \left| \frac{\partial B}{\partial y} \right|^2 \quad (5)$$

$$g_{xy} = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \frac{\partial R}{\partial x} \frac{\partial R}{\partial y} + \frac{\partial G}{\partial x} \frac{\partial G}{\partial y} + \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \quad (6)$$

则在 (x, y) 处的最大变化率角度为

$$\theta(x, y) = \frac{1}{2} \arctan \left[\frac{2g_{xy}}{(g_{xx} - g_{yy})} \right] \quad (7)$$

且在 $\theta(x, y)$ 方向上变化率的值(即梯度)计算如下:

$$F_{\theta}(x, y) = \left\{ \frac{1}{2} [(g_{xx} + g_{yy}) + (g_{xx} - g_{yy}) \cos 2\theta + 2g_{xy} \sin 2\theta] \right\}^{1/2} \quad (8)$$

需要注意的是由式(9)计算出的角度 $\theta(x, y)$ 在半开区间 $[0, \pi)$ 上有两个相隔 90° 的值,也即沿着这两个方向之一 F 最大,而沿着另一个方向 F 最小,所以最终结果是由选择每个点上的最大值产生,具体的推导过程见文献[6].实际中对于很小的梯度边缘人眼很难观测到,故设定阈值 T ,对于小于 T 的梯度值令其为 0,其他情况下,梯度值不变.如图 3 所示,其中 T 设为 0.2.由图可见对于计算每个 RGB 分量梯度,然后合并得到的色彩图像边缘中由于大的垂直梯度边缘对小的水平梯度边缘的掩盖使得水平梯度边缘消失,而直接在 RGB 向量空间中提取边缘的方法则能检测到水平边缘.所以文中采用了直接在 RGB 向量空间中提取彩色图像边缘的边缘检测方法.



(a) 原始彩色图像 (b) 计算每个RGB分量梯度,然后合并得到色彩图像边缘 (c) 直接在RGB向量空间中提取的色彩图像边缘

图3 图像示例

2.2 边缘相似度

对于提取的参考图像和失真图像的梯度边缘 F_1 及 F_2 ,文中采用相关系数来计算边缘相似度:

$$ES(F_1, F_2) = \frac{\sigma_{xy} + C}{\sigma_x \sigma_y + C} \quad (9)$$

式中, σ_x 和 σ_y 分别为梯度边缘 F_1 及 F_2 的标准差, σ_{xy} 为 F_1 及 F_2 的协方差如下所示(为了避免出现 σ_x 或 σ_y 为 0 时,使得边缘相似度为无穷大的情况,分子分母均加上一个很小的系数 $C = 10^{-10}$).

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2} \quad (10)$$

$$\sigma_y = \left(\frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2 \right)^{1/2} \quad (11)$$

$$\sigma_{xy} = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \right) \quad (12)$$

如文献[3]所述,由于人眼视觉系统的视网膜的中央凹的局域特性,某一瞬间人眼只能清楚的观测到图像的某一局部区域,所以实际中将两梯度边缘 F_1 及 F_2 以 16×16 窗口大小逐个像素滑动的分成若干块,再在各块内求边缘相似度.考虑人眼的频率选择特性,人眼对于尖锐的边缘纹理要比平滑的边缘纹理更敏感,所以对各块内边缘尖锐平滑程度进行加权求平均(块内边缘越尖锐则加权值越大)得到整个失真图像与参考图像的边缘相似度(MES):

$$MES = \frac{\sum_{j=1}^M w_j ES_j}{\sum_{j=1}^M w_j} \quad (13)$$

式中, ES_j 为第 j 块的边缘相似度, w_j 为第 j 块的边缘尖锐度,用参考图像第 j 块的平均边缘强度来表示:

$$w_j = \frac{1}{N} \sum_{i=1}^N x_{ji} \quad (14)$$

2.3 图像背景相似度

实验表明仅用边缘相似度来评价图像质量已经能够达到很好的评价效果,说明图像边缘纹理失真程度对于图像的视觉质量具有很重要的影响.然而对于边缘失真很小但背景失真比较大,如色彩失真,对比度下降等也可能使图像质量存在很大程度的失真,为此文中综合边缘和背景相似度来评价图像质量.为了简便本文仅用参考彩色图像与失真彩色图像的灰度信号 G_1 和 G_2 的相关系数来计算图像背景相似度:

$$BS(G_1, G_2) = \frac{\sigma_{xy} + C_1}{\sigma_x \sigma_y + C_1} \quad (15)$$

为了避免分母为 0 使得相似度为无穷大的情况,分子分母均加上一个系数 $C_1 = 12.5$. 实际中同样将两灰度图像 G_1 及 G_2 以 16×16 窗口大小逐个像素滑动的分成若干块,再在各块内求背景相似度,最后平均得到整个图像的背景相似度(MBS):

$$MBS = \frac{1}{M} \sum_{j=1}^M BS_j \quad (16)$$

2.4 基于边缘背景相似度的质量测度(EBS)

由前面两部分内容综合得失真图像的质量测度 EBS 为

$$EBS = (MES)^\alpha \cdot (MBS)^\beta \quad (17)$$

式中, α, β 为大于 0 的实数权重,分别用来表示边缘相似度和背景相似度对图像质量测度的影响程度,本文令 $\alpha = \beta = 1$. 由此可用 EBS 值(0-1)来评价失真图像质量, EBS 值越大,则失真图像与参考图像相似度越高,即图像质量越好.

2.5 具体算法实现

文中提出的 EBS 算法包含四个步骤,具体算法步骤如下:

(1) 直接在 RGB 向量空间中利用上述梯度法提取参考彩色图像与失真彩色图像边缘;

(2) 分块计算边缘相似度,最后加权平均各分块得到失真图像的边缘相似度 MES;

(3) 提取两彩色图像的灰度信号分量 G_1 和 G_2 ,通过分块计算两灰度图像的相关系数得到整个图像的背景相似度 MBS;

(4) 综合边缘及背景相似度得到失真图像与原始参考图像的相似度 EBS 值并以此评价失真图像质量.

3 实验结果与分析

3.1 各算法应用于几幅彩色图像的评价结果比较

为直观地比较各客观评价算法,给出几幅彩色失真图像的质量评价结果,如图 4 所示,为便于观测,所列出图像为实际图像裁剪出的大小为 250×180 的部分),其中 DMOS 值(一组背景不同的观察者对失真图像的评分进行统计平均后所得到的评价结果)表示各失真图像主观视觉质量,DMOS 值越小表示图像质量越好,DMOS 的取值范围为 $[0, 100]$. 各算法的评价结果如表 1 所示. 由表 1 可以看出,文中提出的 EBS 评价结果与 DMOS 值一致,也即与图像的实际视觉感知质量一致,而其他的评价

算法则不能很好的区分上面三幅不同失真程度的彩色图像质量.



图 4 几幅原始及其失真图像

表 1 各评价方法对三幅彩色失真图像质量的评价结果

评价方法 失真图像	PSNR	MSSIM	IFC	VIF	EBS	DMOS
Light house	19.4	0.82	1.57	0.20	0.50	46.28
house	27.0	0.88	0.75	0.20	0.25	56.44
woman	22.2	0.84	0.73	0.12	0.10	60.19

3.2 文中算法与各类算法基于图像质量评价库的整体性能比较

为了全面地比较文中算法与其他算法的性能,采用美国 TEXAS 大学图像和视频工程实验室提供的图像质量评价库来检验各类算法的性能^[7]. 该图像质量评价库由 29 幅高分辨率原始参考彩色图像经过五种不同类型的不同程度失真的 779 幅降质图像组成,包括不同程度的 JPEG2000 压缩失真图像 169 幅, JPEG 压缩失真图像 175 幅,加入白噪声的失真图像 145 幅,高斯模糊失真图像 145 幅及传输信道失真(Fastfading)图像 145 幅,同样其主观视觉质量由 DMOS 值确定. 图 5 为各类客观图像质量评价算法的散点图.

为了定量的分析各图像质量评价模型,文中根据视频质量专家组(VQEG)Phase I test 报告^[8],采用文献[4]中介绍的 Logistic 函数将各客观评价算法评价结果非线性回归到主观评价结果(DMOS)空间,然后采用下面五个标准来衡量各图像评价算法的性能:

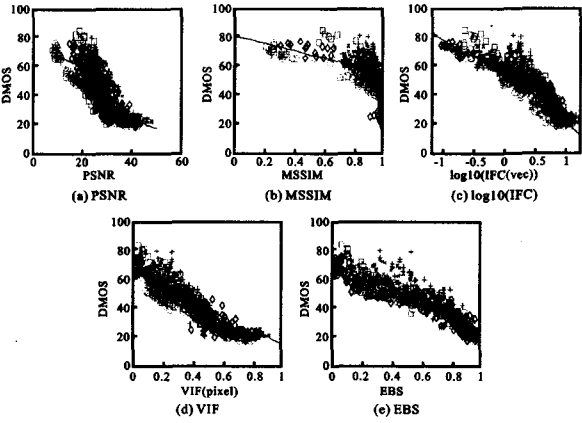


图 5 五种客观图像质量评价标准的散点图
($\times$)JPEG200, ($+$)JPEG, ($\circ$)加入白噪声,
($\square$)高斯模糊, ($\diamond$)传输信道失真(fast fading)

- (1) 非线性映射后的主/客观图像质量之间相关系数(CC)范围:0~1,越接近 1 越好;用来测定客观评价算法结果相对主观评价结果的准确度;
- (2) 秩相关系数(SROCC),值越大越好,用来测定客观评价算法结果相对主观评价结果的单调性;
- (3) 溢出率(OR),值越小越好,用来测定客观评价算法结果与主观评价结果的一致性;
- (4) 平均绝对值预测误差(MAE),值越小越好;
- (5) 均方根预测误差(RMS),值越小越好.

图 6 为经过非线性回归后的各类客观图像质量评价算法的散点图,其中非线性回归函数为 Logistic 函数式(18),而 Logistic 函数中的参数拟合则采用 MATLAB 中的 nlinfit 函数.

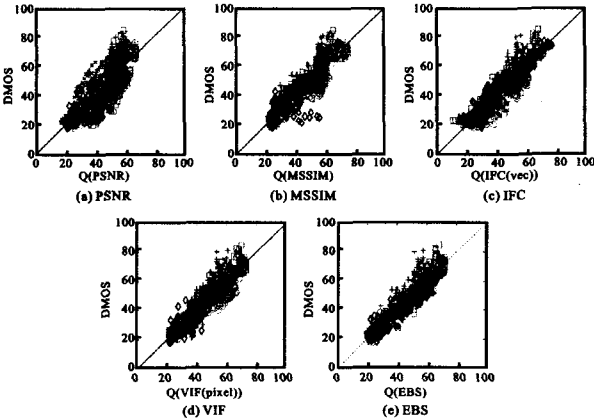


图 6 非线性回归后的五种客观图像质量评价标准的散点图

($\times$)JPEG200, ($+$)JPEG, ($\circ$)加入白噪声,
($\square$)高斯模糊, ($\diamond$)传输信道失真(fast fading)

$$Q(x) = \beta_1 \text{logistic}(\beta_2, (x - \beta_3)) + \beta_4 x + \beta_5 \tag{18}$$

$$\text{logistic}(\tau, x) = \frac{1}{2} - \frac{1}{1 + \exp(\tau x)} \tag{19}$$

表 2 中给出各图像质量评价模型的性能比较,可以看出文中提出的 EBS 算法性能要远远好于完全没有考虑人眼视觉特性的 PSNR 评价方法,相比 MSSIM,IFC 以及基于像素域的 VIF 方法性能也有较大提高.

表 2 各评价方法的性能比较

评价方法	CC	MAE	RMS	OR	SROCC
PSNR	0.826	7.27	9.09	0.114	0.820
MSSIM	0.912	4.97	6.61	0.035	0.911
IFC	0.917	4.92	6.44	0.037	0.915
VIF	0.925	4.64	6.09	0.027	0.926
EBS	0.935	4.33	5.72	0.023	0.934

在 MATLAB 7.0 的运行环境下(AMD Dual Core CPU 2.11GHz, 1.00GB 的内存),我们记录了各种图像质量评价方法运行一幅 512×768 大小的彩色图像的平均执行时间,如表 3 所示,IFC 算法由于涉及到小波分解耗时较长,而文中算法和 PSNR, MSSIM 以及基于像素域的 VIF 方法则比较快速.

表 3 各种评价方法的执行时间(MATLAB)

方法	PSNR	MSSIM	IFC	VIF	EBS
执行时间/s	0.3	1.5	7.3	3.2	4.8

4 结束语

文中所提出的基于人眼视觉特性的彩色图像质量评价方法应用于包含多种类型失真的 779 幅图像质量评价库的实验结果表明,该算法能够比较快速地评价失真图像质量,评价性能较好.另外所提出的方法不仅能够评价彩色图像质量,对于灰度图像质量评价只需改动算法的第一步即直接从灰度图像中用 Sobel 掩模获得图像边缘同样适用.

参考文献:

[1] Eckert M P, Bradley A P. Perceptual quality metrics applied to still image compression[J]. Signal, 2005(70):177-200.

[2] Watson A B, Hu J, McGowan J F. III. DVQ: a digital video quality metric based on human vision[J]. Electron Imaging, 2001, 10(3): 20-29.

[3] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error measurement to structural similarity [J]. IEEE Transactions on image processing, 2004, 13(4): 600-612.

(下转第 67 页)

5 结束语

AEPF 将进化机制引入 PF 算法,利用遗传算法中的选择、交叉和变异操作引入 PF 算法,实现算法性能的改进.并且采用自适应重采样技术,只在需要时才进行重采样,减少重采样的次数,降低好样本被删除的危险,进一步提高算法的性能.将该算法应用于移动机器人定位.

实验结果表明,进化策略和自适应重采样的应用不仅改善了样本的多样性,减轻了样本贫化的影响,而且实现了样本优化,减少了所需样本的数量,使粒子滤波算法具有更高的定位精度.但在采用同样的样本数时,AEPF 算法进行定位的时间长于 PF.

参考文献:

- [1] Rudy N. Robot localization and kalman filters[D]. Utrecht University: Institute of Information and Computing Sciences, 2003.
- [2] Thrun S, Fox D, Burgard W, et al. Robust monte carlo localization for mobile robots[J]. Artificial Intelligence, 2001, 128(1/2):99-141.
- [3] Mo Y W, Xiao D Y. Hybrid system monitoring and diagnosing based on particle filter algorithm[J]. Acta Automation Sinica, 2003, 29(3):641-648.
- [4] Fox D. Adapting the sample size in particle filters through KLD-sampling[R]. Washington: University of Washington, 2003:1-27.
- [5] 殷波. 移动机器人同时定位与地图创建方法研究[D]. 青岛:中国海洋大学, 2006: 46-53.
- [6] 厉茂海,洪炳熔,罗荣华. 用改进的 Rao-Blackwellized 粒子滤波器实现移动机器人同时定位和地图[J]. 吉林大学学报:工学版, 2007, 37(2):401-406.
- [7] 谭宝成,廉春原,徐艾,等. 一种基于改进遗传算法的机器人路径规划方法[J]. 西安工业大学学报, 2008, 28(5):456-459.
- [8] Thrun S, Burgard W, Fox D. Probabilistic robotics[M]. London: MIT Press, 2005:189-279.
- [4] Sheikh H R, Bovik A C, Veciana de G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. IEEE Transactions on image processing, 2005, 14(12): 2117-2128.
- [5] Sheikh H R, Bovik A C. Image information and visual quality[J]. IEEE Transactions on image processing, 2006, 15(2): 430-444.
- [6] Di Zenzo S. A note on the gradient of multi-image[J]. Computer Vision Graphics and Image Processing, 1986, 33: 116-125.
- [7] Sheikh H R, Wang Z, Bovik A C, et al. Image and video

作者简介:

夏益民 女, (1980-), 博士研究生, 讲师. 研究方向为移动机器人定位和地图创建.

杨宜民 男, (1945-), 教授, 博士生导师. 研究方向为智能机器人控制、人工智能.

(上接第 63 页)

- quality assessment research at LIVE. [EB/OL]. [2008-12-10]. <http://live.ece.utexas.edu/research/quality/>.
- [8] VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment. [EB/OL]. [2009-01-10]. <http://www.vqeg.org/>

作者简介:

付伟 男, (1984-), 硕士研究生. 研究方向为图像质量评价、图像压缩编码.

顾晓东 男, (1970-), 博士后, 副教授. 研究方向为人工神经网络、图像处理、模式识别等.