

一种基于视觉兴趣性的图象质量评价方法

汪孔桥 沈兰荪 邢昕

(北京工业大学信号与信息处理研究室, 北京 100022)

摘要 图象质量的正确评价是图象信息工程领域内一项很有意义的研究课题,但现有的图象客观评价方法并不完全符合人眼视觉特性.合理地评价图象质量应充分遵循人眼的视觉特性,但由于受到人的心理、文化背景、周围环境、不同的应用场合等多种因素的影响,人眼对同一幅图象中的不同区域往往具有不同的感兴趣程度,即人眼具有视觉兴趣性特性.该文在充分利用视觉兴趣性特性的基础上,针对只存在一个感兴趣区的图象,提出了一种简化的图象质量评价方法.该方法近似地认为人眼对感兴趣区的兴趣程度与其面积成反比,当图象中感兴趣区的面积为零时,该文方法将退化为客观质量评价方法,也就是说客观质量评价方法是该方法的特例.实验结果证明,该方法能够较好地符合人眼的主观视觉特性.

关键词 图象质量评价 感兴趣区 视觉兴趣性

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2000)04-0300-04

A Quality Assessment Method of Image Based on Visual Interests

WANG Kong-qiao, SHEN Lan-sun, XING Xin

(Signal and Information Lab, Beijing Polytechnic University, Beijing 100022)

Abstract Research on right quality assessment of image is very meaningful in the image information project, but the objective quality assessment of image cannot always be in accordance with human vision. Any reasonable quality assessment method of image should abide fully by human subjective vision. Due to many factors such as visual psychology, culture background, surrounding environment, different applications and so on, an image can be divided into different regions of interest (ROI) by vision interests. To best advantage of human visual interests, a simplified method for subjective assessment of image is presented to an image with only one ROI in this paper. The degree of visual interest to the ROI of the image is regarded approximately to be an inverse measure of the acreage of the ROI. As the acreage of the ROI is zero, the method will become the objective quality assessment of image, that is to say, the objective quality assessment of image is a specific case of the method in this paper. The experimental results have shown the effectiveness of the proposal scheme. It is in better accordance with human subjective vision.

Keywords Quality Assessment of Image, Region of interest, Visual interest

0 引言

图象质量的正确评价是一项很有意义的研究课题,目前常用的评价方法主要有两种,即主观质量评价和客观质量评价.主观评价方法就是让观察者对

同一幅图象按视觉效果的好坏进行打分,并对其进行加权平均^[1].该方法由于劳动强度太大,且不能应用于图象实时传输的场合;客观评价方法是用恢复图象偏离原始图象的误差,来衡量图象恢复的质量,最常用的有均方误差(MSE)和峰值信噪比(PSNR).其中MSE的表达式为

$$MSE = \frac{\sum_{0 \leq i < M} \sum_{0 \leq j < N} (f_{ij} - f'_{ij})^2}{M \times N} \quad (1)$$

式中, $\{f_{ij}\}, \{f'_{ij}\}$ 分别表示原始图象和恢复图象, 且 $0 \leq i < M, 0 \leq j < N$, 而 $PSNR$ 本质上与 MSE 相同, 其表达式为

$$PSNR = 10 \lg \frac{256 \times 256}{MSE} \quad (2)$$

式(1)和式(2)看起来直观、严格, 但用它们所求得的结果常与主观视觉不一致. 这是因为它们是从总体上反映原始图象和恢复图象的灰度差别, 并不能反映出少数象点的较大灰度差别和较多象点的较小灰度差别等情况. 显然, 对图象中各个象点同样对待, 不能反映人眼的视觉特性.

图象最终是供人看的, 因此合理地评价图象质量应充分遵循人眼的视觉特性. 由于受到人的心理、文化背景、周围环境、不同的应用场合等多种因素的影响, 人眼对同一幅图象中的不同区域往往具有不同的感兴趣程度, 即人眼具有视觉兴趣性特性^[2~4]. 现有的一些客观质量评价方法^[5]一般都没有考虑到人眼的上述视觉特性, 因此对存在感兴趣区的图象往往得不到与图象主观质量评价非常吻合的结果. 本文在充分考虑到人眼视觉兴趣性的基础上, 提出了一种新的图象质量评价模型, 近似地认为人眼对感兴趣区的兴趣程度与其面积成反比, 当感兴趣区的面积为零时, 本文方法将退化为客观质量评价, 也就是说客观质量评价是本文方法的特例. 实验结果表明, 如果测量图象中存在一个感兴趣区, 则用本文模型测量的结果能较好地吻合图象的主观质量评价结果.

1 图象质量评价模型

1.1 图象质量的测量过程

人眼观看一幅图象时, 有时会对其中的某些区域感兴趣. 但当一幅图象存在多个感兴趣区时, 各感兴趣区间会产生交叉影响, 造成各感兴趣区兴趣程度的差异. 例如, 对于一幅头肩象, 发型设计师除注意人脸区之外, 可能更关注该人的发型, 而对于出现在视频会议等领域内的头肩象, 人们也许更注意人脸区. 有关存在多个感兴趣区的图象质量评价问题, 以后将作进一步的研究. 为简化起见, 本文假设测量图象中只有一个感兴趣区 A_1 , 其面积为 S_1 , 不感兴趣区 A_2 的面积为 S_2 , S 为图象的总面积 ($S = S_1 +$

S_2). 另外还假定人眼分别对感兴趣区和不感兴趣区内各象素点的兴趣程度相同, 且它们的兴趣加权值分别为 λ_1, λ_2 , 加权值越大, 表示人眼对该区感兴趣的程度越高. 将式(1)加以修改可得出本文测量图象的均方误差 ($IMSE$)

$$IMSE = \frac{1}{S} \left| \lambda_1 \sum_{(i,j) \in A_1} (f_{ij}, f'_{ij})^2 + \lambda_2 \sum_{(i,j) \in A_2} (f_{ij}, f'_{ij})^2 \right| \quad (3)$$

$$\lambda_1 S_1 + \lambda_2 S_2 = \lambda_1 S_1 + \lambda_2 (S - S_1) = S \quad (4)$$

由式(3)、(4)可看出, 当 $S_1 = 0, \lambda_2 = 1$ 或 $S_2 = 0, \lambda_1 = 1$ 时, $IMSE$ 将退化为 MSE . 图1给出了用本文方法进行图象质量评价的测试过程, 其中感兴趣区的确定(即 S_1 的确定)可通过检测手段自动完成, 例如对于头肩象, 可通过自动人脸检测^[6,7]来确定感兴趣区, 即人脸区, 检测到的人脸区越大, S_1 就越大.

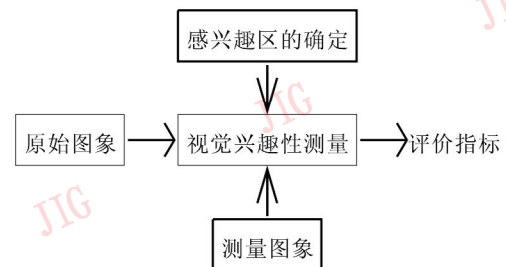


图1 质量评价的测量过程

1.2 视觉的兴趣性测量

视觉经验告诉我们, 对于一幅给定的图象, 人眼对其不同区域感兴趣的程度是不同的, 如头肩象, 人眼对人脸区(感兴趣区)的失真较敏感, 而对其余部位(不感兴趣区), 则能允许较大的失真存在.

对于感兴趣区, 当其面积越来越小时 (S_1 越小), 人眼对其失真就越来越敏感 (λ_1 越大), 也就是说, 人眼对它的兴趣程度大体上与其面积成反比, 当感兴趣区扩张至整幅图象时 ($S_1 = S$), 人眼对其兴趣将降到最低 (λ_1 达到最小), 如不感兴趣区一样.

对于不感兴趣区, 感兴趣区对其兴趣程度有屏蔽作用. 即当感兴趣区的面积较大时, 随着感兴趣区面积的增大 (S_1 越大, S_2 越小), 人眼对不感兴趣区的兴趣将逐渐增强 (λ_2 越大); 而当感兴趣区的面积较小时, 随着感兴趣区面积的减小 (S_1 越小, S_2 越大), 尽管人眼对感兴趣区的兴趣越来越强 (λ_1 越大), 但由于不感兴趣区不断地扩大, 使得观察者越来越不能忽视它的存在, 也就是说, 人眼对不感兴趣区的兴趣程度也越来越强 (λ_2 越大).

在感兴趣区的面积等于零或者感兴趣区扩张至整幅图象的两种情况下, 测量图象的失真对人眼视觉的影响应该是等价的.

将上述视觉经验加以定量化, 便可实现对视觉兴趣性的定量测量.

由式(3)可知, 对视觉兴趣性测量的核心任务就是如何确定加权值 λ_1 、 λ_2 . 根据上述视觉经验, 针对仅存在一个感兴趣区的图象, 我们提出了一种简化的质量评价模型, 其中设定的 λ_1 、 λ_2 需满足如下两个近似条件:

① 当 S_1 由 0 增加到 S 时, λ_1 将由 $+\infty$ 减小到 1;

② 当 S_1 由 0 增加到 $S/2$ 时, λ_2 将由 1 减小到某一不小于 0 的常数, 而当 S_1 由 $S/2$ 增加到 S 时, λ_2 将由该不小于 0 的常数增加到 1.

(1) λ_1 、 λ_2 的构造

根据上述两个条件, 我们将 λ_1 、 λ_2 构造如下:

$$\lambda_2 = 1 - \frac{2k}{S} \sqrt{S_1(S - S_1)} \quad (5)$$

其中 $k \in [0, 1]$ 为调整因子, 它反映了人眼对不感兴趣区或感兴趣区的重视程度. 由式(4)可得

$$\lambda_1 = \frac{S}{S_1}(1 - \lambda_2) + \lambda_2 \quad (6)$$

(2) λ_1 、 λ_2 的验证

由 $\sqrt{S_1(S - S_1)} \leq \frac{S}{2}$ 知: $\lambda_2 \geq 1 - k$, 当 $S_1 = S/2$ 时, $\lambda_2 = 1 - k$, 达到极小; 当 $S_1 = S$ 或 $S_1 = 0$ 时, $\lambda_2 = 1$, 所以式(5)满足 λ_2 的边界条件.

以 S_1 为变量, 对式(5)两边取导数

$$\lambda'_2 = \frac{d\lambda_2}{dS_1} = - \frac{k(S - 2S_1)}{S \sqrt{S_1(S - S_1)}} \quad (7)$$

由式(7)知, 当 $S_1 > S/2$ 时, $\lambda'_2 > 0$, 即当 $S/2 < S_1 < S$ 时, λ_2 是 S_1 的严格单调递增函数. 而当 $S_1 < S/2$ 时, $\lambda'_2 < 0$, 即当 $0 < S_1 < S/2$ 时, λ_2 是 S_1 的严格单调递减函数. 因此 λ_2 的构造函数(式(5))满足假设条件②.

由式(6)可得

$$\lim_{S_1 \rightarrow 0} \lambda_1 = \lim_{S_1 \rightarrow 0} \left[1 + 2k \sqrt{\left| \frac{S}{S_1} - 1 \right|} - \frac{2k}{S} \sqrt{S_1(S - S_1)} \right] = +\infty \quad (8)$$

当 $S_1 = S$ 时, 由 $\lambda_2 = 1$ 可得 $\lambda_1 = 1$, 所以式(6)满足 λ_1 的边界条件.

以 S_1 为变量, 对式(6)两边取导数

$$\lambda'_1 = \frac{d\lambda_1}{dS_1} = - \frac{S}{S_1^2}(1 - \lambda_2) + \left| 1 - \frac{S}{S_1} \right| \lambda'_2 \quad (9)$$

将式(7)代入式(9)得

$$\lambda'_1 = - \frac{k(S + S_1) \sqrt{S_1(S - S_1)}}{SS_1^2} \quad (10)$$

很显然, $\lambda'_1 < 0$, 即 λ'_1 是 S_1 的严格单调递减函数, 因此 λ_1 的构造函数(式(6))满足假设条件①.

2 实验及结论

为验证本文方法的有效性, 我们从电视节目采集到一幅头肩象, 并用文献[5]提出的编码方法、传统小波编码方法以及 JPEG 标准分别对其进行压缩, 适当地调整量化器参数, 重建出 11 幅质量不同的图象(见表 1). 然后用主观评价、 $PSNR$ 、 $ISNR$ 三组参数指标分别对该组恢复图象进行质量评估, 其中 $ISNR$ 的定义为

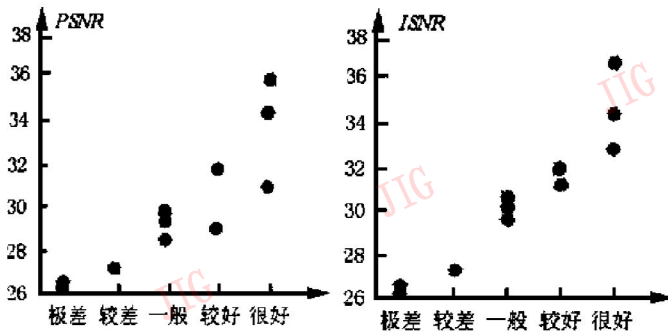
$$ISNR = 10 \lg \frac{256 \times 257}{IMSE} \quad (11)$$

本文中的主观评价方法为: 请 10 位观察者分别对这 11 幅恢复图象进行主观打分, 打分级别分很好、较好、一般、较差和极差 5 个等级. 打分时, 双眼平视, 屏幕且与屏幕保持大约一尺的距离. 最后将这 10 位观察者的打分进行加权平均, 得到表 1 中主观打分一栏的评价级别.

表 1 实验结果

重建图象	主观评价	$PSNR$	$ISNR$	备注
1	很好	31.18	33.14	图象主观质量很好
2	较好	29.08	31.28	人脸区清晰, 背景区稍微有些模糊
3	一般	28.64	30.09	人脸区清晰, 但可感到背景区有些模糊
4	极差	26.32	26.59	图象整体极度模糊
5	一般	29.94	30.39	可感到图象整体有些模糊
6	很好	35.89	37.08	图象主观质量很好
7	极差	26.07	26.10	图象整体方块效应严重
8	较差	27.31	27.38	图象整体方块效应较严重
9	很好	34.36	34.43	图象主观质量很好
10	较好	31.96	32.05	有很轻微的方块效应
11	一般	29.86	29.92	可感觉到方块效应存在

图 2 分别给出了 $PSNR$ 、 $ISNR$ 与主观评价关系示意图, 从中可看出: $PSNR$ 不能很好地与主观视觉相符, 而 $ISNR$ 能较好地与主观视觉相符, 即图象的主观质量越好, $ISNR$ 值越高. 图 3 示出了其中 3 幅恢复图象, 其中(b)的人脸区较清楚, 背景区稍微有些模糊, (c)的人脸区有些模糊, 而(d)的整体出现轻度的方块效应, 它们的主观打分分别为较好、一般、



(a) *PSNR* 与主观评价的关系图 (b) *ISNR* 与主观评价的关系图

图 2 *PSNR*、*ISNR* 与主观分级的关系图

一般,但它们的 *PSNR* 却分别为 29.08dB、29.94dB、29.86dB,显然它们的 *PSNR* 与主观视觉不相符,而用本文的评价方法对这三幅图象的评价结果 (*ISNR*) 分别为: 31.28dB、30.39dB、29.92dB,由此可看出用本文方法的评价结果与主观视觉是相符的。

本文在充分利用人眼视觉兴趣性特性的基础上,针对只存在一个感兴趣区的图象,提出了一种简化的图象质量评价模型,近似地认为人眼对感兴趣区的兴趣程度与其面积成反比,实验结果证明该模型能较好地符合人眼的主观视觉。



(a) 原始图象



(b) 主观: 较好 *PSNR*= 29.08dB *ISNR*= 31.28dB



(c) 主观: 一般 *PSNR*= 29.94dB *ISNR*= 30.39dB



(d) 主观: 一般 *PSNR*= 29.86dB *ISNR*= 29.92dB

图 3 部分测试图象

参 考 文 献

- 1 A M Van Dijk, J-B Martens. Subjective quality assessment of compressed images. *Signal Processing*. 1997, 58: 235~ 252.
- 2 Shen Lansun, Wang Kongqiao, Qing Xia *et al.* A multirate wavelet image coding method based on human face detection. In: *Proc. ICM1'99*, Hong Kong, VI. 9-VI. 13.
- 3 汪孔桥, 沈兰荪. 一种结合人脸检测的静止小波图象编码方法. *电子学报*. 1998, 26(4): 88~ 90.
- 4 汪孔桥, 沈兰荪. 一种结合人脸检测的多码率小波图象编码方法. *高技术通讯*. 1998, 8(9): 6~ 9.
- 5 Westen *et al.* Perceptual image quality based on multiple channel HVS model. In: *ICASSP'95*. Detroit, Michigan, USA, 2351~ 2354.
- 6 Shen Lansun, Wang Kongqiao, Xing Xin. Automatic human face detection and tracing in a complex background. *ISO/IEC JTC1/SC29/WG11 MPEG99/m4422*.
- 7 汪孔桥, 邢昕, 沈兰荪. 复杂背景下的快速人脸检测与跟踪. 见: *第四届中国计算机智能接口与智能应用学术会议论文集*. 山西, 1999, 245~ 249.



汪孔桥 1999 年获中国科学技术大学电子工程与信息科学系信号与信息处理专业博士学位,曾在北京工业大学信号与信息处理研究室从事科学研究. 研究兴趣为模式识别、图象处理、计算机视觉等,发表学术论文十余篇。

沈兰荪 北京工业大学教授,博士生导师. 主要研究兴趣为光谱信号的检测、图象信号处理以及基于 VLSI 的实时信号处理技术,出版“智能仪器与信号处理技术”、“图象编码与异步传输”、“ICP—AES 光谱干扰校正方法的研究”等 10 部专著,国内外发表论文 100 余篇。

邢昕 1974 年生,1997 年毕业于北京工业大学电子工程系电子工程专业及计算机专业. 现为北京工业大学信号与信息处理研究室硕士研究生,研究方向为智能化信息处理。