

一种基于掩盖效应的感知域图像质量评价方法

魏崇奎 成礼智

(国防科技大学理学院数学与系统科学系, 长沙 410073)

摘要 图像的压缩编码技术及其比特流在信息空间上的传输控制, 为图像质量的评价提出了新课题。通过分析人类视觉系统(HVS)的视觉非线性、多通道及掩盖效应等特性, 建立起相应的图像视觉处理模型, 并将原始图像与降质图像变换到感知域进行误差分析。在处理 HVS 模型的多通道中, 将图像的空间频率按视觉系统的掩盖效应特点, 分成 5 个带分别进行滤波, 以模拟人眼对图像的主观质量评价特性。实验结果表明这种模拟多通道的图像感知域评价模型, 得到的 HVS 值能充分反映人类视觉系统的掩盖效应的特点, 用于对图像质量的评价该模型通用性好, 与主观评价结果高度一致。

关键词 HVS 图像质量 感知域 掩盖效应

中图法分类号: TN911.73 文献标识码: A 文章编号: 1006-8961(2004)02-0195-06

An Image Quality Measure Scheme in the Perceptual Field via Masking

WEI Chong-kui, CHENG Li-zhi

(Mathematics and System Science Faculty, Science Institute, Changsha National University of Defense Technology, Changsha 410073)

Abstract With the technology of image compression and codec booming, and in order to control the transmission of their bit-stream, the term of image quality is given a kind of new degree and challenge. Today the research of Image Quality Assessment has also become one of the basis technologies in the realm of images and information. This paper analyzed the characters of Human Visual System (HVS) at first, that included the visual nonlinear, multi-channel and masking effect as well, then constituted a kind of mathematic model to simulate the course that the human visual system proceeds the image information. In the model, original images and distorted images are transformed respectively to the perceptual field, that is, both of them are processed by the basic visual nonlinear firstly, and with that enter a band pass filter created by the modulate transform function. During the processing of the multi-channel in the HVS model, according to the characteristic of masking, the function of band pass filter is divided into five bands by the varying spatial frequency of initial image. Finally the value comes into being that we gained from the vision pathways, and then the HVS-E (Error) is obtained. The result from experiments indicates the fine agreement between the HVS-E and the MOS (mean opinion score).

Keywords HVS, image quality, perceptual field, masking

1 引言

数字图像系统的发展、性能优化以及相应的视觉信息的获取、显示、存贮与传输都与人类视觉系统对数字化的反应与处理机制有着极为密切的联系。无论是图像压缩、水印信息掩藏, 还是印刷图像处理(如 half-toning patterns、喷绘)等应用都依据了人眼的

视觉特性, 或利用视觉的局限性来进行的, 因此建立 HVS(human visual system)^[1]的计算机模型来进行数字图像(包括印刷图像)的质量评价已经成为该领域中一个重要的研究课题。图像质量的评价方法主要包括均方误差(MSE)、峰值信噪比(PSNR)、HVS、多维模型(multidimensional modeling)、各种范数距离以及模糊距离(fuzzy distance)、主成份与多重回归分析^[2]等。20 世纪 70 年代国外首先对此展开了研

究^[3],近几年来国内也开始进行探讨^[4]。常用的 MSE 或 PSNR,由于其算法的局限性,其客观性往往与主观值有较大的差别,而主观的评价测试,试验条件较困难,且又受到个人情绪等一些无法预测与控制因素的影响,另外,它对图像编码等处理也没有建设性的意义。图像的最终服务对象是人的眼睛,虽然人类视觉系统相当复杂,目前这方面的知识也非常有限,但基于人眼视觉系统的图像质量评价,不仅对评价本身,对图像的各种处理、成像等也都是极有意义的。根据视觉特性的基本原理建立的 HVS 模型有很多,基本上都是先对原图像和编码后的图像作差,然后再对差值建立感知模型,作感知函数处理,得出感知评价价值。本文主要研究静止灰度图像的质量评估,方法是先将两幅图像分别经过一系列基于感知的变换,变换到感知域,将它们分别映射到视觉皮层,然后作差比较,得到 HVS-E(error)误差值。在处理视觉多通道时,将整幅图像按空间频率分成 5 个带,分别用视觉的调制转移函数进行滤波。实验表明,此算法能较好地匹配图像的主观质量,算法简单,易于硬件实现。

2 视觉基本特性

2.1 视觉幅度非线性

视觉系统分辨细节的能力用分辨的相邻两点的视角 θ 的倒数表示,其值与该两点在视网膜上的成像位置有关,以黄斑区为中心向四周作非线性下降,它与亮度的关系取决于相对亮度的变化,而不是取决于整幅图像的基底亮度,那么亮度感觉的增量 ΔS 可用相对亮度的增量来度量,即

$$\Delta S = K \frac{\Delta B}{B} \quad (1)$$

式中, B 为客观亮度, ΔB 为相对亮度增量。

对式(1)积分后得到感觉亮度

$$S = k \ln B + K_0 = K' \lg B + K_0 \quad (2)$$

其中, $K' = K \ln 10$, K_0 为常数, K 为与整个图像的平均亮度相关的常数,图像的平均亮度较大或较小时, K 的值可选得小一些,对于人眼正常的亮度范围, K 可取 1。式(2)说明人眼的感知亮度 S 与实际亮度 B 的对数成线性关系,即韦伯-费赫涅尔(Web-Fechner)定律。如图 1 所示中,图 1(b)是对图 1(a)进行了以自然对数为底的客观亮度变换,再将亮度值放大 35 倍后的效果,它反映的是视觉皮层对图像初级处理的映像。



(a) Lena 原图



(b) Lena 非线性处理图

图 1 视觉非线性对比图

2.2 对比度敏感度带通

2.2.1 对比度基本定义

对比度的基本定义有两种^[1],一种是 Weber-Fechner 定义的,称为 Weber-对比度;另一种是 Michelson 定义的,称为 Michelson-对比度。

人眼视觉系统的响应主要地依赖于相对于背景亮度的局部变化,而不是绝对亮度值,因此对比度就是这种亮度相对变化的测量,Weber-对比度的数学表达式为

$$C_w = \frac{\Delta L}{L} \quad (3)$$

其中, L 为背景亮度值, ΔL 为目标与背景的亮度差值。由于视觉的适应性,在大约 $2 \sim 1096 \text{ cd/m}^2$ 的亮度范围内,人眼刚能分辨的亮度差异 JND(just noticeable difference)所对应的门限对比度值为一常数。也是最低的分辨阈值,当背景亮度较强或较弱时,人眼的对比灵敏度阈值都会增大,分辨亮度差异的能力下降,而 Weber-对比度基本控制在该常数范围之内。

Michelson 认为,实际的图像视觉效果很大程度上依赖于激励的特征,最主要的是颜色(对彩色而言)、空间频率、时间频率(对视频而言),通常用对比度感知函数 CSF(contrast sensitivity function)来量化这种依赖关系,在 CSF 的测量中,认为激励是随着频率而周期变化的,通常是按正弦周期变化,这种意义下的对比度(称为 Michelson-对比度)定义为

$$C_M = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} \quad (4)$$

其中, $L_{\max}$ 、 $L_{\min}$ 分别为目标的最大、最小亮度,式(3)与式(4)给出的两种定义之间的关系表明 CSF 与 JND 互为倒数。用这两种定义直接计算对比度是很困难的,因此实际计算时都用反映其属性的数学模型代替。

2.2.2 带通模型

视觉与对比度的关系用视觉系统的对比度敏感

度函数 CSF 表示,或称为人眼视觉系统空间调制转移函数(modulate transform function,MTF),简称为 MTF 过程,不同实验所得的 CSF 函数形式各异,但基本上都认为 HVS 的对比度敏感性是空间频率的函数,且具有带通滤波器性质,其对高低频端敏感度下降。

基本的 CSF 函数形式有 $H(r)^{[5]}$, $S(\rho, \theta, l, i^2, d, e)^{[6]}$, $|A(r)|H(r)^{[7]}$ 以及 $S(\omega)^{[8,9]}$ 。他们的 CSF 函数响应曲线如图 2 所示,其中曲线 1 为 Ngan 等人提出的 CSF 函数,曲线 2、3 分别为由 Nill 给出的曲线及其改进曲线^[7]。其中横轴为图像的空间频率,单位为 c/d(cycle per degree 周每度),纵轴为 CSF 函数,表明视觉响应的相对幅度。

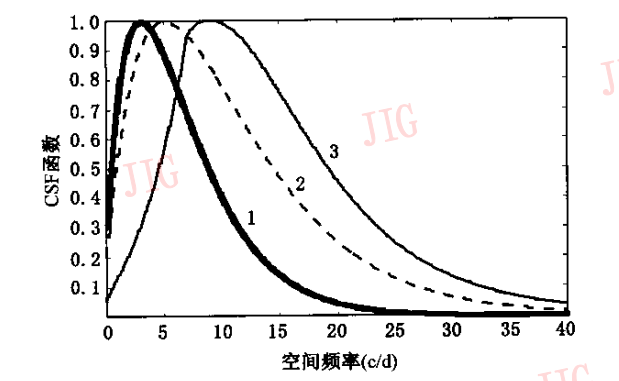


图 2 CSF 响应曲线

从上面的曲线可以看出,曲线(a)的峰值在3c/d处,相对较低,两边的坡度较陡,更适合于本文所用的带通处理模型。其函数表达式为

$$H(r) = (0.31 + 0.69r)e^{-0.29r} \tag{5}$$

其中, $r = \sqrt{(R_f)^2 + (C_f)^2}$, 而 R_f, C_f 分别为水平和竖直方向的空间频率。

2.3 多通道与掩盖(Masking)效应

视觉皮层的细胞对不同的视觉信息或激励,如颜色、频率和方向等有不同的敏感性,而对目标识别、掩盖与自适应的研究认为:所有这些特征激励在人的视觉系统中,是在不同的通道进行处理的,这是早期的多通道理论,而后进一步的理论得出,视觉机制的多通道之间并不是彼此孤立的,而是存在着相互的作用、判决与相互影响,以产生最佳视觉。对于静止灰度图像来说,图像的多通道特性可以由它的空间频率和方向性来表征,只要用足够多的适当的调谐部件,图像在视觉皮层的整个方向带和频率带都可以被完全覆盖,即可以完全模拟视觉系统的多

通道,但多通道之间的相互作用机制尚不明确。

一个激励单独存在时,是很容易辨识的,掩盖效应(masking)是指由于另一个激励的存在导致它完全不能或者不容易被检测到,即被掩盖了。在描述多通道中激励之间的相互作用时,掩蔽效应是必须考虑的一种非常重要的现象。另外,这种掩盖效应导致的视觉系统的探测阈值(JND)的改变,既可以是抑制,也可以是加强,在计算机视觉中,有时为了识别特定的目标,需制作特定的掩盖器(masker),这也是掩盖效应。掩盖效应基本模型的对数曲线如图 3 所示。

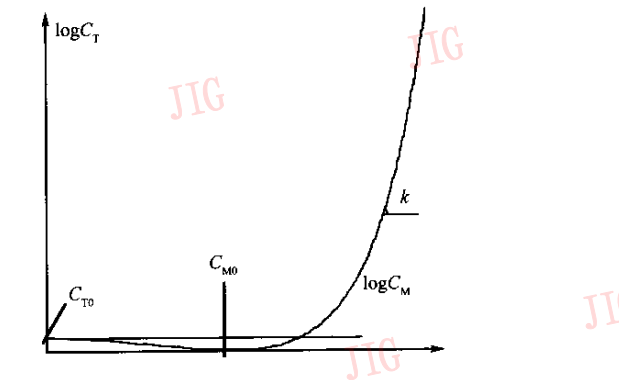


图 3 掩盖效应模型

其中, k 为线性部分的斜率, C_M 是所有产生掩盖效应的背景的对比度, C_T 是存在掩盖效应时目标的对比度检测门限值, C_{M0} 是有利的激励导致视觉门限值下降时的背景对比度值, C_{T0} 是没有任何掩盖时的对比度检测门限值。

基于以上多通道与掩盖效应的这种特殊关系,需要说明以下两点:一是视觉皮层细胞的响应在频域呈带通特性,且有其相应的敏感度的峰值位置和响应带宽;二是人的视觉系统具有将独立的视觉信息聚合起来的能力,且不同的视觉皮层细胞对频率域的不同频段敏感度不一。为此,用一组空间频率分段的滤波器来建立人眼视觉特性的模型,将可见数据按照掩盖效应的特点分别在不同的带通频段内进行滤波。

3 模型与实验结果

3.1 基本框图

由于直接建立多通道模型有一定的难度与复杂性,视觉系统的掩盖效应在多通道之间的相互作用在生物学领域还不完全明朗,因此,这里尝试将图像

按空间频率的不同属性分成 5 个带,然后在每个带中根据掩盖效应的特点,用同一个 CSF 响应函数的不同处理形式,分别对原始图像和降质图像进行频域滤波,最后再作差处理。这样就将 MTF 带通过程、多通道与掩盖效应结合在一起,将图像映射到视觉皮层的感知域再进行处理。

其基本模型如图 4 所示。

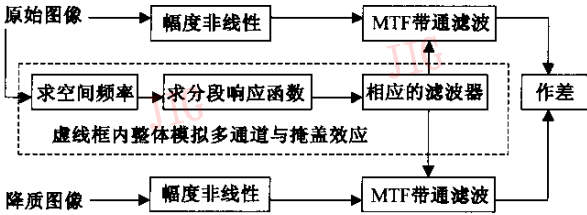


图 4 基本 HVS 模型

3.2 实现方法

3.2.1 空间频率与 HVS-E 值的计算

视频中的空间频率是一个与视角有关的概念^[1],令 H 为屏幕(4:3 的电视)的高度, D 为视距,则视角 α 定义为

$$\alpha = 2\arctan(H/2D)$$

按周每度(cycle per degree)的视角计算,分辨率或者最大空间频率 $f_{\max}$ 为

$$f_{\max} = (L/2\alpha)c/d$$

其中, L 是符合 Nyquist 采样定理的扫描线数。数字图像的空间频率定义方式虽有不同,但都反映其基本属性。由于影响像素的误差可视度的因素是误差周围的局部区域环境,而不是整个图像背景,对给定的 $M \times N$ 的图像块行频、场频分别定义为

$$R_f = \sqrt{\frac{1}{MN} \sum_{j=0}^{M-1} \sum_{k=1}^{N-1} [F(j,k) - F(j,k-1)]^2}$$

$$C_f = \sqrt{\frac{1}{MN} \sum_{k=0}^{N-1} \sum_{j=1}^{M-1} [F(j,k) - F(j-1,k)]^2}$$

则空间频率为 $r = \sqrt{(R_f)^2 + (C_f)^2}$ 。

空间频率函数取式(5)(峰值在 $3c/d$ 处), $M \times N$ 取为 4×4 。按照图 5,将空间频率 r 的 CSF 函数 $h(r)$ 按对比度的探测阈值分为 4 部分: $0 \sim 0.3$, $0.3 \sim 0.5$, $0.5 \sim 0.8$, $0.8 \sim 1$ 。由于 $h(11.396) = 0.3$, $h(8.77) = h(0.3563) = 0.5$, $h(1.186) = h(5.820) = 0.8$,因此空间频率 r 就被相应地分为 6 个频带,并分别根据图像特点定义相应的响应函数

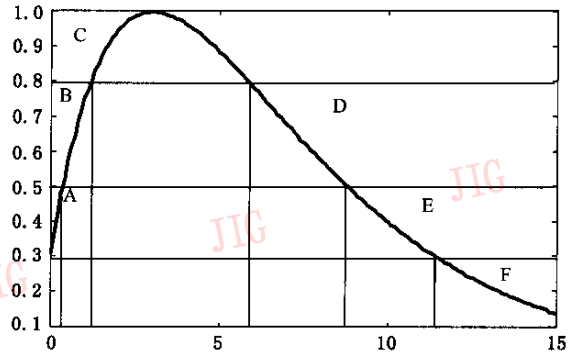


图 5 式(5)曲线分带图

如下:

$0 \sim 0.3563$ A 区,该区图像极平坦,易分辨,影响图像质量, $h_1(r) = 1 - h(r)$;

$0.3563 \sim 1.186$ B 区,为平坦区,不易分辨目标, $h_2(r) = h(r)$;

$1.186 \sim 5.820$ C 区,对比度适中,清晰有目标为重点区, $h_3(r) = 1$;

$5.820 \sim 11.396$ D 区与 E 区,变化较丰富,不易分辨, $h_4(r) = h(r)$;

11.396 以上 F 区,变化最丰富,难分辨,基本不影响图像质量, $h_5(r) = 0$ 。

这样对按照不同的空间频率 r ,对图像分别进行全通 h_3 、全阻 h_5 、带通 h_1, h_2, h_4 等滤波。

图像 HVS-E 误差值的整个计算过程如图 6 所示。

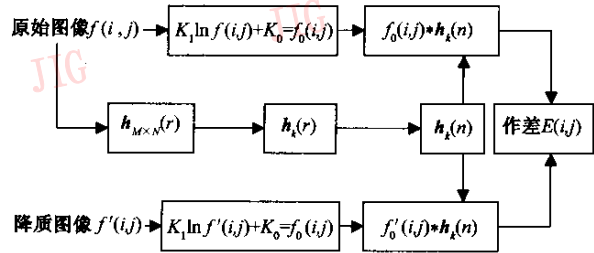


图 6 算法框图

图 6 中, $f(i,j)$ 为原始图像矩阵, $f'(i,j)$ 为降质图像矩阵, K_1 为图像整体亮度的调节系数,整个图像的平均亮度过大或过小时其值可以选得小一些,本文实验用正常图像, K_1 取 1,由于最后是作差处理,故取 $K_0 = 1$,取 $M = N = 4\text{pixels}$,是为了保证得到更有效的空间频率值, K 取 $1 \sim 5$, $h_k(r)$ 与 $H_k(r)$ 分别为各频带的响应函数和相应的滤波器。

最后的图像评测值定义为

$$\text{HVS-E} = \frac{1}{IJ} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} E(i, j) \quad (6)$$

其中, $E(i, j)$ 为绝对差。

因此, HVS-E 值越小, 降质图像越接近原始图像, 处理后的图像质量就越好。反之越差。

3. 2. 2 实验结果比较

本文实验包括 3 个部分: (1) $512 \times 512 \times 8\text{bits}$

的 Lena 图像在不同压缩比下的 HVS-E 值测量; (2) Lena 与 $512 \times 512 \times 8\text{bits}$ 的 Barbara 图像在相同随机噪声下(图 7)的 HVS-E 值比较; (3) $400 \times 543 \times 8\text{bits}$ 的 Missile 图像将相同随机噪声加在不同部位(图 8)的 HVS-E 值比较。实验结果分别列于表 1~表 3。其中 MSE, PSNR 分别为常用的均方误差与峰值信噪比, $Ave-r$ 为整个原图像的平均空间频率。



(a) 原图像

(b) 降质图像

图 7 Lena 与 Barbara 加相同噪声的对比图



(a) 原图像



(b) 噪声加在原图上半部分生成的图像



(c) 相同噪声加在原图下半部分生成的图像

图 8 Missile 图像噪声加在不同部位的对比图

表 1 Lena 图在不同压缩比下的 HVS-E 值与 PSNR 的对比 ($Ave-r=10.9485$)

	用 JPG 压缩时的质量百分比											
	1	5	10	20	30	40	50	60	70	80	90	100
MSE	256.037 2	128.176 1	64.982 6	36.640 5	27.195 6	22.351 2	19.104 5	16.305 8	13.109 8	9.730 4	5.440 4	0.092 3
PSNR	24.047 8	27.052 7	30.002 8	32.491 2	33.785 8	34.637 8	35.319 4	36.007 4	36.954 8	38.249 5	40.774 5	58.476 0
HVS-E	1.809 9	1.198 1	0.814 6	0.597 7	0.511 2	0.461 8	0.431 8	0.405 2	0.369 6	0.321 8	0.252 7	0.030 4

表 1 表明 HVS-E 值与通常意义下的图像客观质量的一致性。

表 2 Lena 与 Barbara 图像加相同噪声下的对比

	MSE	PSNR	Ave-r	HVS-E
Lena+noise	66.803 4	29.882 8	10.948 5	0.512 2
Barb+noise	66.803 4	29.882 8	19.871 1	0.479 4

表 3 Missile 图像将相同噪声加在不同部位的对比

	MSE	PSNR	Ave-r	HVS-E	处理方法
Missile 1	15.520 9	36.221 6	15.690 9	0.209 1	上部加噪声
Missile 2	27.605 1	33.720 9	15.690 9	0.175 2	下部加噪声

图 7 中,(a)为原图像,(b)为加相同随机噪声后的降质图像。不难看出,由于掩蔽效应,相对于原图像,空间频率更大的 Barbara 的视觉效果要比 Lena 的视觉效果好一些,这与表 2 中描述的 HVS-E 值与人眼的主观感知相一致。需要指出的是,作为客观标准的 PSNR 值并不能很好地反映这一现象,并且,在加不同的随机噪声时,PSNR 还会得出与主观视觉相反的结果。

从图 8 对比图看出,Missile 2(图 8(a))的主观质量显然优于 Missile 1(图 8(b)),而从 PSNR 看,却明显与主观质量不符合,由于 HVS-E 是依据空间频率来处理的,它能充分反映视觉感知的掩蔽效应,与主观质量高度一致。

4 结 论

依据人的视觉基本特性,建立了一种基于空间频率分带的 HVS 图像质量测评模型,此算法充分考虑了人眼的视觉非线性与掩蔽效应的特点。实验表明,所得到的 HVS-E 的感知误差能很好地反映图像的主观质量。由于本文方法对空间频率太敏感,故对应用中小范围的感兴趣区域的评价,评价不是太准,进一步的研究是针对不同的应用,采用数值逼近的方法与主观评价相结合,得出与不同应用相应的待定系数,使该方法具有自适应的功能。

参 考 文 献

1 Winkler S. Vision Models and Quality Metrics for Image

Processing Applications [D]. Germany: Diplom-Ingenieur der Elektrotechnik, Technische Universität Wien de nationalité autrichienne, 2000-12-21.

2 Miyahara M, Kotani K, Algazi V R. Objective Picture Quality Scale for Image Coding [J]. IEEE Transactions on Commnication, 1998,46(9):1215~1225.

3 Mannos J L, Sakrison D J. The effect of a visual fidelity criterion on the encoding of images[J]. IEEE Transactions on Inform Theory, 1974,20(7):525~536.

4 汪孔桥. 数字图像的质量评价[J]. 测控技术, 2000,19(5):14~16.

5 Ngan K N, Rao K R, Singh H. Cosine transform coding incorporating human visual system model[A]. In:SPIE fiber'86 [C], Cambridge, MA U S A, 1986:165~171.

6 Daly S. Visible differences predictor: An algorithm for the assessment of image fidelity[J]. SPIE Conference On Human Vision, Visual Processing and Digital Display III, 1992,1666:2~14.

7 Nill K B. A visual model weighted cosine transform for image compression and quality assessment[J]. IEEE Transactions on Commcation, 1985,33(3):551~557.

8 Algazi R. Design of subband coders for high quality based on perceptual criteria[J]. Application of Digital Image Processing XV, 1992,SPIE-1771:396~405.

9 王楠楠,李桂苓. 符合人眼视觉特性的视频质量评价模型[J]. 中国图象图形学报(A 版),2001,6(6):523~527.



魏崇奎 1972 年生,1995 年毕业于华中师范大学理论物理专业,在太原卫星发射中心从事跟踪测量工作,中级职称,现为国防科技大学计算数学硕士研究生。主要研究方向为图像压缩及其质量评价。

成礼智 1962 年生,现为国防科技大学教授,博士生导师,现主持多个科研项目。从事快速算法、图像压缩、遥感图像处理等的研究。发表论文及学术专著共 60 余篇,其中 17 篇为 SCI 检索。