

一种基于视觉熵的图像分割压缩方法

单志广 魏涛 杨扬

北京科技大学信息工程学院,北京 100083

摘要 基于视觉熵概念提出了一种静止图像分割压缩方法.通过对人类视觉系统特性的归纳,总结了基于视觉熵的图像分割原理,提出了用于量化图像特征的数学定义和基于视觉熵的分割算法.实验结果表明,这种基于视觉熵的图像分割压缩算法既提高了压缩比,又能保证压缩后的重建图像整体上具有高的主观视觉感知质量.

关键词 视觉熵; 图像分割; 图像编码

分类号 TP391.41

图像压缩编码的理论基础是信息论.从信息论的角度看,压缩就是去掉信息中的冗余,保留不确定的因素,去除确定的因素,用更接近信息本质的描述来代替原有冗余的描述.这个本质的东西就是信息量(即不确定因素).但信息量不是孤立、绝对的,它与信息的传递密切相关,信息接受者知识世界的改变是信息传递的本质.人眼是一种最优的图像编码系统,图像信息传递的本质特征应该与人眼的视觉特性保持一致.

近年来人们从人类视觉系统 HVS (Human Vision System) 的研究中得到启发,基于 HVS 的压缩方法引起了人们的广泛注意^[1].对于一幅图像,人眼对其中每个区域的敏感度是不同的,从而导致了 HVS 对图像各个部分的注意力不均匀且没有规律,而这决定了人对图像视觉信息的理解.如果不考虑图像表达的意义,不把图像内容与编码方法结合起来,对整个图像一概而论,显然是不合理的.为了衡量图像信息量的大小以及人类视觉系统对图像的敏感程度,人们提出了视觉熵 VE(Visual Entropy)^[2]的概念.

1 视觉熵(VE)

视觉熵是信息熵的概念与人类视觉系统(HVS)特性的结合.

1.1 信息熵概念

在信息论中,信源 X 中目标 x_i ($i=1,2,\dots,n$)

的信息量 $I(x_i)$ 被定义为:

$$I(x_i) = -\log_2 p(x_i),$$

式中, $p(x_i)$ 为 x_i ($i=1,2,\dots,n$) 出现的概率.

信息的熵(Entropy) $H(X)$ 就是信源 X 中所有目标的平均信息量,即:

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i).$$

熵是一种对不确定问题的度量准则.一幅图像的熵值是它的平均信息量的量度,图像中有许多信息是冗余的,去掉冗余的信息,而使压缩后的图像熵值不低于图像信息源的熵值,则不会丢失图像的有用信息.这一结论就是图像压缩编码的依据.

1.2 人类视觉的生理特性和心理特性^[3]

从生理学知,视觉受到大脑的选择性注意机制的控制,人无法同时识别同一复杂场景内部的多个目标,只有进入黄斑区内的景物信息被充分接收,而之外的大量信息最终被中心滤波.另外,视觉选择性与客体的特性也有很大的关系.人眼对空间频率接近于零的平滑区域和空间频率相似的纹理区域有很大的视觉“钝性”,所以“不变”与“规则变化”的场景很容易在人的意识中遗忘,视觉似乎只对“突变”和“极不规则变化”的区域感兴趣.

从心理学和美学角度讲,视觉是具有选择性的.研究表明,人类视觉的选择性使人们的目标集中于周围环境中使生活变得更加美好和那些妨碍生活正常进行的方面.从而进一步说明,视觉是一种积极的感受行为,不仅与物理和生理因素有关,在相当程度上也取决于心理因素.

1.3 VE 的概念

一幅图像的局部视觉熵^[4]定义为:对于一幅图像的特定区域,在保证编码后的图像与原图像无太大视觉感知差异的前提下,可获得的编码后的每个像素的最小比特数。

可见,细节少的区域以较低的比特数就可以编码获得高逼真度,其视觉熵就低;具有重要感知细节的区域视觉熵就高。然而,细节与视觉熵并不等同。因为具有高细节的区域(如随机纹理)也可能压缩到很低的比特率而不会引起感知上的明显扭曲。

视觉熵对于研究具有可变感知级别的编码是一种概念上的指导。视觉熵描述的是人眼对信息的一种主观量度,是对HVS的一种很好的仿真,具有符合人类视觉特性的特点,能够把人眼对图像的敏感度进行量化。

视觉熵理论现在还处在发展阶段。虽然目前还没有精确的方法计算某个图像区域的视觉熵,但是当图像编码的目标是高的感知质量,而不是高的信噪比(SNR)或低的均方误差(MSE)时,视觉熵是一个非常有利的概念。

2 基于 VE 的图像分割

2.1 基于 VE 的图像分割原理

信息源一般是一种不确定的情况,当看到一幅图像时,人类视觉的不确定性导致人眼对整个图像各个部分的注意力存在差别,而视觉熵可以成为量度这些信息的工具。通过对HVS的研究,有以下基本结论:

(1)对于一幅图像,HVS对于每一个区域的敏感度是不同的,这主要是由于区域内部特性与周围区域的特性存在差异。

(2)HVS对图像注意力的分布是参差不齐、不均匀的,这决定了人们对图像视觉信息的理解。

(3)HVS对于亮度的空间频率变化(如纹理结构)其敏感度要大于相同空间频率的其他因素的变化。即人眼对亮度的突变最为敏感,这些区域往往是图像中包含信息量最大,对人们的理解最为重要的部分。

(4)HVS对亮度突然变化的区域的敏感度大于亮度变化慢或变化小的区域。

(5)HVS虽然对图像中亮度变化大的区域比较敏感,但对具有规则变化的区域(如窗帘或

壁纸的花纹),人眼会产生一种适应性,以至于很容易在人的意识中遗忘,使得人眼对这种区域的注意力降低。而且普遍来说这些区域包含的内容意义并不大,对图像理解不起决定性作用。

(6)对于一幅图像,如果某一区域的变化是没有规律的,人眼就会很敏感;如果某一区域的变化规律与周围区域的变化规律差别很大,就会极大地引起人们对这一区域的兴趣,HVS对这种区域的敏感度最高。

从理论和数学的角度描述VE至关重要,但在观察一幅图像时,HVS的结构机理是很复杂的,以至于人们很难掌握其各个区域VE的变化。为了简化研究,我们在对客观进行假设的情况下,对基于视觉熵VE的分割原理作一个原始的数学描述^[5]。

(1)在基于VE的图像分割中,只考虑每个图像区域的内部特性和相邻特性,如下式所示:

$$VE(x_i) = f[\phi(x_i), \varphi(x_i), I] \quad (1)$$

其中, $\phi(x_i), \varphi(x_i)$ 分别表示区域 x_i 的内部特性和周围特性, I 表示包含 x_i 的图像。

(2)我们认为一幅图像中与HVS关系最为紧密的只有2个区域:随机纹理区域和关键点区域。随机纹理区域对于HVS的理解意义较少,而关键点区域却在很大程度上影响了人对图像的理解,后者的视觉熵远大于前者。

2.2 基于 VE 的图像分割算法

(1)基于亮度空间域变化率的第一次分割。

定义 1 对于一幅图像的某个区域,其所有像素亮度分量的平均值称为该区域的平均亮度,用变量 L 表示,可以通过下式求得:

$$L_n = \sum_{k=1}^n l_k / n \quad (n \geq 1) \quad (2)$$

其中, l_k 表示每个像素的亮度分量值, n 表示这个区域中像素的数量。

定义 2 一幅图像中,2个相邻像素亮度分量的差的绝对值称为两个像素间的亮度变化,用变量 D 表示,可以通过下式求得:

$$D_{ij} = |l_i - l_j| \quad (3)$$

其中 l_i, l_j 表示2个相邻像素的亮度分量值。

定义 3 对于一幅图像的某个区域,所有相邻像素亮度变化 D 的总和与该区域中所含像素总数的比,称为该区域的亮度变化率,用变量 C 表示,可以通过下式求得:

$$C_n = \frac{\sum_{i=1}^n \sum_{m=1}^n D_{i,m}}{n} \quad (4)$$

其中， n 表示区域中像素的数量， l, m 表示相邻的2个像素。

亮度变化率 C 反映了图像区域内部亮度变化的大小和快慢，计算出每个区域的 C 后，就可以根据设定的阈值 T 将图像划分成“低信息量区域”($C < T$)和“高信息量区域”($C > T$)，并可通过调节 T 的大小来改变图像的压缩比和压缩质量。

图像分割要将图像划分成若干 $N \times N$ 的子块。如果 N 取得过大，每次运算所需的存储空间和运算时间较大；但当 N 小到一定程度时，子块之间的边界上可能存在被称为“边界效应”的不连续点。通过实践可知，当 $N < 8$ 时，边界效应比较明显，故 $N \geq 8$ 要考虑到后续的处理要采用 DCT 变换，所以在分割中 N 取 8。这样一幅 256×256 的图像被划分为 1024 个子块，分别进行处理。

为了精确的计算每个 8×8 子块的亮度变化率 C ，我们提出了一种“伞式算法”：从 8×8 子块左上角的第 1 个像素开始，计算与它相邻的 3 个像素的亮度变化 D ，然后再以这 3 个像素为起点，依此类推，就可以计算出这个 8×8 子块中所有相邻像素的亮度变化。其运算的轨迹像 1 把打开的伞，所以称其为伞式算法。如图 1 所示。

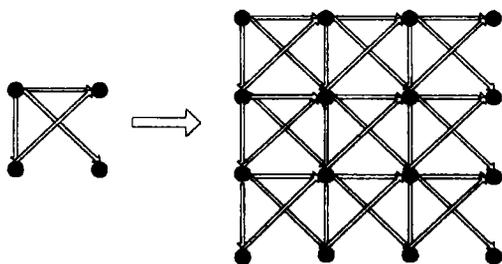


图 1 计算相邻像素亮度变化的伞式算法示意图
Fig.1 Track of umbralla algorithm

(2) 基于 VE 的第二次分割. 根据图像中区域的内部特征对图像进行了初次分割后，我们再根据区域间特性的差异，对图像进一步分割。

为了描述 2.1 节中(5),(6)2条基于视觉熵分块的 HVS 特性，我们给出以下定义。

定义 4 一幅图像中，某一个子块 i 与其所有相邻子块亮度变化率差值的均方根，称为该子块的亮度平均变化率，用变量 R_i 表示。这里相邻的子块有 8 个， R 可以由下式求得：

$$R_i = \sqrt{\sum_{k=1}^8 (C_k - C_i)^2} / 8 \tag{5}$$

其中， C_k, C_i 为相邻 2 个子块的亮度变化率。

定义 5 对于一幅图像的某个特定区域，其中某一个子块 i 的亮度平均变化率 R_i 与该区域的平均亮度 L_n 的比值，称为该子块的亮度相对变化率，用变量 V_i 表示：

$$V_i = \frac{R_i}{L_n}$$

其中， L_n 表示该区域的平均亮度， n 表示该区域的像素数量。

子块的亮度相对变化率 V 反映了该子块与周围子块在亮度变化上的差异。相对变化率较高的区域对人眼理解图像意义非常关键，我们把这种区域称为“关键点区域”；相对变化率较低的区域一般为一些相近的纹理（如墙壁或桌布的花纹），对人类的理解作用不大，我们称为“随机纹理区域”。

(3) 基于 VE 的分割过程. 根据视觉熵分割原理，我们将一幅图像划分为 3 部分：低信息量区域、随机纹理区域和关键点区域，然后再进行压缩。分割过程如图 2 所示。

我们以 256×256 , 24 位真彩色图像 POR-

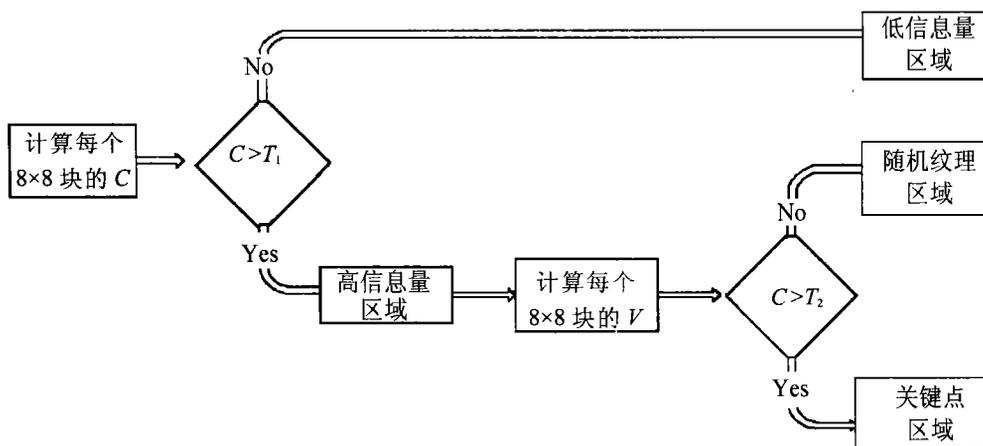


图 2 基于视觉熵分割算法示意图
Fig.2 Procedure of VE-based segmentation

TRAIT为例,将其划分为3个区域,整个分割过程如图3所示.图3显示了对PORTRAIT图像进行2次分割的过程和结果,以及对图像进行的边缘检测.从2种处理结果的结果可以看出,

分割的结果与图像的边缘能够较好的吻合,正好将图像中对人眼最敏感的区域分离出来.这说明我们提出的基于视觉熵的分割方法较好的反映了人类视觉系统在观察一幅图像时的特性.

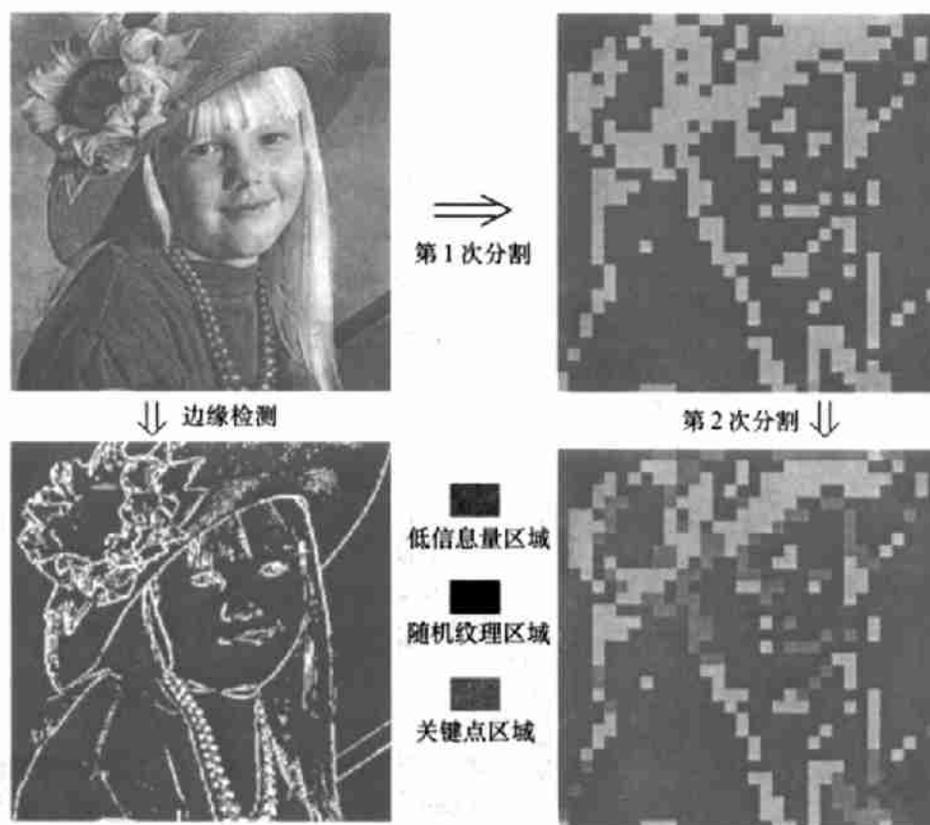


图3 对PORTRAIT图像进行2次分割的示意图

Fig.3 Two steps of segmentation and edge detection for image PORTRAIT

3 实验及结果

由于关键点区域是人类视觉系统最为敏感的区域,对人们理解整个图像的意义和对恢复图像的质量起决定性的作用,因此要求编码后关键点区域有最小的失真.为此,我们采用了DCT变换结合动态 Huffman 编码和游程编码的方法.对于低信息量和随机纹理区域,由于它们携带的信息量少,对人类视觉系统不敏感,对人们理解图像不会起到决定性的作用,所以应提高对这两部分区域的压缩比.对这两部分的编码采用 DCT 编码结合动态 Huffman 编码的方法,加大了调节器步长,并使用了分层编码.这种压缩方法能够取得压缩比与恢复质量二者间的很好折衷,而且压缩比和恢复质量都可以根据要求进行调节,符合我们提出的分割压缩系统对不同区域采用不同压缩比方法的要求.

本文基于 Windows 95/Windows NT 平台,编

程采用 Visual C++ 5.0,系统的编程、调试和运行在 P II 350 计算机上完成.图4给出了用本文所提出的方法对 256×256 , 24 位真彩色图 PORTRAIT 进行压缩的结果, $C_r = 11$, $SNR = 14.8$ db.



图4 基于VE分割算法压缩后重建图像

$C_r = 11, SNR = 14.8$ db

Fig.4 The image PORTRAIT after coding and reconstruction using VE-based segmentation

从压缩结果可以看出,采用基于视觉熵的图像分割压缩方法重建后的图像,能够达到很好的主观视觉效果,尤其是对人类视觉系统最为敏感的区域(如眼部区域).对于人类视觉系统不敏感的部分,即本文所定义的低信息量区域和随机纹理区域,也实现了比较好的主观感知效果.

4 结束语

基于视觉熵的图像分割原理,提出了用于量化图像特征的数学定义和基于视觉熵的分割算法.实验结果表明,这种基于视觉熵的图像分割压缩算法是对人的视觉系统很好的仿真,较好地反映了人类视觉系统的特性,描述了人眼对图像不同区域的不同敏感度,既提高了压缩比,又保证了压缩后的重建图像整体上具有高的主观视觉感知质量.

视觉熵描述了人眼对信息的主观量度,具有符合人类视觉特性的特点,能够把人眼对图像的敏感度进行量化,是一个很有发展前途的理论.但视觉熵理论还处在发展阶段,需要对其进行深入的研究和探讨,主要问题有:

(1) 视觉熵的数学化描述还不完善,需要对视觉熵的量化尺度进行统一的描述,形成一致的评测标准.

(2) 本文在图像分割中使用了固定尺寸的子块,导致对于图像中细节丰富的区域划分的不够精细,而对于变化平淡的部分则又显得效率不高,进一步的完善应考虑使用尺寸能够自适应变化的子块,从而提高分割的效率.

(3) 视觉熵虽然反映了HVS特性,但是它没有包含人类的理解.物体基(Object-Based)的方法和基于内容的压缩编码是当今发展方向,因此应该努力将人类对图像的理解溶入到视觉熵理论中,以更好地仿真人类视觉系统,并使图像压缩提高到一个崭新的水平.

参考文献

- 1 Wu Xiaolin, Fang Yonggang. A Segmentation-Based Predictive Multiresolution Image Coder. IEEE Transaction on Image Processing, 1995,4(1):34
- 2 Fan Guoliang, Zhou Lihua. Visual Entropy-Based Classified Bath Fractal Transform for Image Coding. In: Proceeding of ICSP'96,1996.898
- 3 张正阳,闵范保,吴成柯.基于面部定位的头肩部视觉质量评价方法.西安电子科技大学学报,1998,25(4):403
- 4 Jacques Vaisey, Allen Gersho. Image Compression with Variable Block Size Segmentation. IEEE Transactions on processing,1992,40(8):2040
- 5 魏涛.基于视觉熵理论的图像分块压缩方法研究:[学位论文].北京:北京科技大学,1999

Image Compression with Visual Entropy-Based Segmentation

SHAN Zhiguan, WEI Tao, YANG Yang

Information Engineering School, UST Beijing, Beijing 100083 China

ABSTRACT An image compression method with Visual Entropy-Based segmentation is presented. Firstly the induction of the characteristics of Human Vision System(HVS) and the principles for Visual Entropy-based segmentation is made. Then the mathematic definition for quantification of image character and the algorithm for Visual Entropy-Based segmentation are fully described. The experimental results have shown that image compression with Visual Entropy-Based segmentation can not only gets a rather low bit rate but also gives satisfactory subjective perceptual quality. This method has well emulated the properties of HVS.

KEY WORDS visual entropy (VE); image segmentation; image coding