数据为“ ”或为“ ”的概率均为二分之一“ ”游程与“ ”游程出现的概率相等 在这个前提下 推导随机序列的游程统计规律不再区分“ ”游程与“ ”游程 只用游程的长度来区分游程 如一游程表示游程长度为一的“ ”游程与“ ”游程的总和、二游程表示游程长度为二的“ ”游程与“ ”游程总和 等等 我们利用前面定义 来考察不同游程的出现概率

设一个长度为 序列 首先用穷举法来考察

很明显当 C 时 序列只能为 或 两种可能 一游长所占的概率为 E

当 C 时 序列有四种可能取值 、 、 、 一游长所占的概率为 F 二游长所占的概率为 F

当 C 时 序列有下列八种等概率的样本空间 、 、 、 、 、 、 、 一游长所占的概率为"F 二游长所占的概率为 F 三游长所占的概率为 F

当 C 时 利用穷举法我们可以归纳出一个 长序列包含不同游长概率随 的变化规律 并可以用数学归纳法给出证明 下面我们来进行该公式的推导

对于 长随机序列 > 游长出现的概率记为

> 前面已经推出

当 C 时 C

当 C 时 C F C F

当 C 时 C F C F

C"F

当 C 时 同样可以利用穷举法 写出 长序列个等概率取值

仔细研究这 个 长等概率序列 发现可以将之分解成

个独立的 游长序列

个独立的 游长序列和一个 序列的全排列

个独立的 游长序列和一个 序列的全排列

! 个独立的 游长序列和一个 序列的全排列

…… ……

¦ 个独立的 G ; 游长序列和一个; 序列的全排列

个独立的 游长序列和一个 G 序列的全排列

因此

$$C\frac{\times\ \times\quad H\ \times\ \times\quad H\ \ldots H\quad G\ \times\ \times\quad G\quad H}{\times}$$

$$C\frac{\times\ \times\quad H\ \times\ \times\quad H\ \ldots H\quad G\ \times\ \times\quad G\quad H\ \times}{\times}$$

"

$$C\frac{\times\ \times\quad H\ \times\ \times\quad H\ \ldots H\quad G\ \times\ \times\quad G\quad H\ \times}{\times}$$

$$G\quad C\frac{G\ \times\ \times\quad G\quad G\ H\quad G\ \times}{\times}C\frac{G}{\times}$$

$$C\frac{\times}{\times}C\ \text{—}$$

!

$$;\ C\frac{;\times\ ;\times\quad ;\ ;\ H\ ;H\quad \times\ ;^{H}\times\quad ;H\quad ;\ H\ \ldots H\quad G\quad \times\quad \times\quad G\quad ;\ H\ ;\ \times\ ;}{\times}$$

令 C G 则

$$G\ ;\ C\frac{;\times\ ;\times\quad ;\ ;\ H\ ;H\quad \times\ ;^{H}\times\quad ;H\quad ;\ H\ \ldots H\quad G\quad \times\quad \times\quad G\quad ;\ H;\ \times\ ;}{G\quad \times}$$

所以综合式 到式 ! 可得

;

$$C\frac{\times?\quad G\quad \times\quad \times\quad G\quad ;\ @G\ ;\ \times\ ;\ H\ ;\ \times\ ;}{\times}$$

$$C\frac{G}{\quad}\quad G\ ;\ H\frac{;\times\ ;}{\times}$$

$$;\ C\frac{G}{\quad}\quad G\ ;\ H\frac{;}{\times\ ;^{H}}\quad \geqslant\ \text{且}\ \neq;$$

当 ; C 时 C —显而易见

根据前面的推导可得定理

ç Ø 长随机序列; 游长出现的频度为 ;

$$;\ C\begin{cases}\frac{;}{;^{H}}G\frac{;G\ ;}{;^{H}} & \text{当 I ; I 时}\\ & \text{当 J ; 时}\\ \text{—} & \text{当 C ; 时}\end{cases}$$

£ ü

当 C 时 C F C"F

$$C\frac{}{H}G\frac{G\quad K}{\times\ H}C\ \text{—}$$

$$C\frac{}{H}G\frac{G\quad K}{\times\ H}C\ \text{"—}$$

当 C 时式 成立

设 C 时

$$;\ C\frac{;}{;^{H}}G\frac{;G\quad K;}{\times\ ;^{H}}$$

成立

当 C H 时 根据式

$$H\ ;\ C\frac{}{H}\times\quad ;\ H\frac{;}{H\ \times\ ;^{H}}$$

$$C\frac{}{H}\times\left(\frac{;}{;^{H}}G\frac{;G\quad K;}{\times\ ;^{H}}\right)H\frac{;}{H\ \times\ ;^{H}}$$

$$C\frac{;}{;^{H}}G\frac{;G\quad K;}{H\ \times\ ;^{H}}$$

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

可见当 C H 时 式　成立 证毕

从上式可以看出当 → ∞时

6　; C

对于一个无限长序列 其不同游长出现的概率随着游程长度的增加而呈几何级数下降 如图　所示

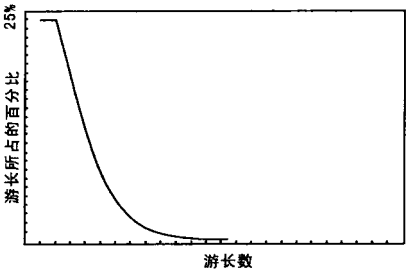

图 3　无限长随机序列的游程统计规律

随机序列游程统计规律 我们用图像数据的游程出现的概率与随机序列游程出现概率的接近程度来描述图像复杂度

ç I　水平复杂度是指在水平方向图像游程出现的频度与随机序列图像出现频度的接近程度 垂直复杂度是指在垂直方向图像游程出现的频度与随机序列图像出现频度的接近程度

对于宽度为) 高度为 % 的图像) ; 表示按行求游程; 出现的频度 % %; 表示按行求的游程; 出现的频度 2); 表示) 长随机序列游程; 出现的频度 2 %; 表示 % 长随机序列游程; 出现的频度 用) % 分别表示图像的水平复杂度与垂直复杂度 则该图像复杂度可用下式表示

) C G " ∑|　　|

C G " ∑|　　(— ×/×)|　"

% C G " ∑|　　|

C G " ∑|　　(— ×/×)|

w, 图像复杂度的上界为　下界为　随机序列组成图像的复杂度为　无限大的全黑或全白图像的复杂度为

这不难理解 对于随机序列组成的图像必有)); C 2); % %; C 2 %; 则

" ∑|　　|C

" ∑|　　|C

所以) % 均为

对于无限大的全黑或全白图像) ; C ;J 9 时)) C % %; C ; J 9 时 % % % C 则

" ∑|　　|C

" ∑|　　|C

所以) % 均为

这正与我们定义图像复杂度的初衷相吻合 同时我们还可以求得 无限大的只有最高单一频率分量的图像 基于频率复杂度定义中最复杂的图像 的复杂度为四分之一 并且随着图像的单一频率下降 复杂度也下降 前面已经计算只有单一直流分量的图像 也就是全黑全白图像 的复杂度为

m^ ¯　¥ ‹ ¨

前面我们研究了图像位平面复杂度 不同的位平面有不同的复杂度 图像每个像素分量一般用八位表示 则可以分解成 ! 个不同的位平面 一般说来位权高的位平面具有较低的复杂度 位权较低的位平面具有较高的复杂度 因为位权高的位平面表示图像的主要成分 位权较低的位平面除了含有图像信息外还包含一定的冗余信息 替换冗余信息不会引起视觉异常 这从另一个角度解释了 4 隐藏算法的合理性

从前面分析可以看出 图像位平面的复杂度越大 图像位平面包含冗余信息的概率越大 容纳隐藏信息的空间就越大 4　隐藏算法[2] [9]就是替代那些复杂度较高的位平面子块 所以合理的复杂度定义在该算法中起着至关重要的作用

文献[2] [9]基于黑白像素变化频率来描述图像位平面复杂度的方法只考虑图像黑白像素变化周期的大小 并没有从根本上反映图像包含冗余信息的情况 而基于游程统计规律的复杂度定义反映的是图像位平面的随机程度 更为准确地反映了图像位平面特性 从不确定性理论而言基于游程统计规律的复杂度定义更为合理 此外 4　算法是用密文直接替代复杂的位平面子块 信息在隐藏前一般都经过压缩等预处理 统计特性更接近随机序列 用接近随机序列的信息去取代接近随机序列的位平面会有较好的不可感知性 而接近随机序列的信息去取代周期性强的位平面容易引起视觉异常

基于黑白像素变化频率来描述图像位平面复杂度的方法只考虑了频率变化的情况 因为实际图像并不是一个规律变化的视觉信号 在实际图像中 周期变化的高频信号出现概率并不是很高 在此定义下适合信息隐藏的子块不是很多 而基于游程统计规律的复杂度定义是从概率角度来考察适合信息隐藏的子块 可挑选更符合条件的子块 在此复杂度定义下实现 4　算法 在同等条件下隐藏

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

容量也更有些优势

在基于游程统计规律的复杂度定义下实现的 4　隐藏算法会比文献? @中实现的算法的性能更优越 图像位平面复杂度定义还可以指导二值图像的信息隐藏 限于篇幅原因 这里不再赘述

前面已经指出图像复杂度与视觉感受有密切关系 把信息隐藏在平坦区和隐藏在纹理区 视觉敏感性和视觉效果会大不相同 这是因为尽管两种隐藏方法的峰值信噪比保持在同一水平 但对图像复杂度的影响却不相同 这说明复杂度的改变也是影响信息隐藏不可感知性的一个重要因素 因此复杂度可以作为图像信息隐藏的客观评价指标

• I Ó D

? @ '5 > (: (* 1- '　'1- (2<　(/　2' 1'6　(<(66 '1(
.'　/　0 4　/. *(　*2(6-3? @　2 1　<' *　0
A　.'7　<(　3/. 7 / (　<　66 '1(.'　/? @4 /. 2
"　!　G
? @ (' -　*　, (　.(　4　/. *(　*2(6-3　/' *　+B
//3 1 7 62 //　< '7 (* / ? @　(.. 2　1 *　'.'　.
. 2/　"　G "!
? @常迥 信息理论基础?A @北京 清华大学出版社
-(*　* L　<(.'　0　0 27 (.'　-　23?A @4 '
5 *　/' *- ($　'8 2'.3　2 //
? @王晓丹 吴崇明 基于 A　4 的系统分析与设计 —— 图像处理?A @西安 西安电子科技大学出版社
B (*) '(M(　B　-　*A ' *　(3/'/ (< < /'*　/3/
. 7 &(/ <　A　4 ?A @) ' (　) ' (　1.2 '1
1-　*3　&'/ -' *,　/

T € e º

R　T　男　年出生于河北赵县 博士 北京电子技术应用研究所副所长 副研究员 中国电子学会高级会员 中国计算机学会高级会员 主要研究方向为多媒体信息安全

7 (' *3&* 9N -　.7 ('　1 7

-• Ð　男　年出生于江苏 中国电子学会通信学分会副主任 多媒体信息安全专委会主任委员 中国计算机学会理事 北京电子技术应用研究所研究员

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net