第7期
2001年7月

电　子　学　报
ACTA ELECTRONICA SINICA

Vol. 29　No. 7
July　2001

图像的失真率函数研究

潘旭山, 叶中行
(上海交通大学应用数学系, 上海 200240)

摘　要:　对于无限大图像 I 定义了一个失真率函数 $d(R/I)$. 它是在有限状态编码译码意义下和给定信息率 R 时渐近可达的失真下限. 本文证明了其编码定理和逆定理. 它不必已知 I 的概率特性.

关键词:　图像压缩; 图像复杂度; 有限状态编码器; 有限状态译码器; 失真率函数
中图分类号:　O236　　文献标识码:　A　　文章编号:　0372-2112(2001)07-0990-03

Distortion-Rate Theory for Individual Images

PAN Xu-shan, YE Zhong-xing
(*Department of Applied Mathematics, Shanghai Jiaotong University, Shanghai 200240, China*)

Abstract:　For every individual images I we define a distortion-rate function $d(R/I)$ which is shown to be an asymptotically attainable lower bound on the distortion that can be achieved for I by any finite-state encoder operating at a fixed output rate R. This is done by means of a coding theorem and its converse. No probabilistic characterization of I is assumed.

Key words:　image compression; image complexity; finite-state encoder; finite-state decoder; distortion-rate function

1　引言

信息论中失真率函数反映了原始数据在给定码率(或压缩比)下所能达到的最小失真, 它对数据压缩的指导意义是巨大的. 然而它的研究是建立在原始数据的概率模型基础上的. 一般来说, 这样的概率模型事先是不知道的. 因此, 在不知道概率特性时研究其数据的失真率函数则很有现实意义. ZIV等[1]从有限状态编译码的角度给出了无限序列(一维数据)的失真率函数的表达,并给出了可渐近达到此极限的压缩方法. 本文将此结果推广到图像(二维数据)情形, 得到了图像的失真率函数的表达式,并证明了相应的正反编码定理. 本文下面是这样安排的: 第2节首先给出了有限状态编译码器和失真率函数的定义. 第3节给出了本文的主要结果及证明. 第4节对进一步可能的研究方向作了讨论. 最后, 在第5节作了小结.

2　有限状态编译码器

给定一幅无限大图像 I, 我们考虑其在有限状态编码及有限状态译码意义下的失真—率理论. 如图1所示, 将 I 置于一个第一坐标轴向下, 第二坐标轴向右的坐标系中. (i, j) 代表坐标为(i, j)的像素. u_{ij}表示相应的像素值. 如图1所示, 其中每一个点代表相应的像素点.

设 I 的每个像素取值于字母集 A, 即 $u_{ij} \in A$, $\forall\, i, j \in Z^+$. 编码后的取值字母集为 B. 设$|A| = \alpha < \infty$, $|B| = \beta < \infty$, 其中 $|\cdot|$ 表示集合的势. 则码率 $R = \log_\alpha \beta$. 一个有限状态编码器(二维意义下) E 是一个七元组: (S, A, B, D, g, t, f), 其中, S 是编码器 E 可以取到的所有可能的状态集合, $|S| = s < \infty$; A 是图像所取值的有限字符集; B 是编码后所取值的有限字符集, B 包含一个所谓空字符, 即字符长度为零; D 是扫描移动增量的有限集合; g 是下一

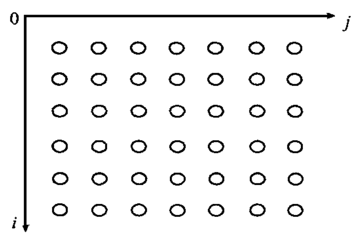


图 1

状态函数, 它将 $S \times A$ 映射到S; t 是扫描增量函数, 它将 $S \times A$ 映射到D; f 是编码函数, 它将 $S \times A$ 映射到B. 当一幅图像被这样一个编码器扫描编码时:

编码器读取一(输入)序列

$$x = x_1 x_2 x_3 \cdots x_i,\ x_k \in A,\ k = 1, 2, \cdots, i;$$

输出一码字序列 $y = y_1 y_2 y_3 \cdots y_i$, $y_k \in B$, $k = 1, 2, \cdots, i$;

扫描增量序列为 $\vec{d} = \vec{d}_1 \vec{d}_2 \vec{d}_3 \cdots \vec{d}_i$, $\vec{d}_k \in D$, $k = 1, 2, \cdots, i$;

状态序列为 $z = z_1 z_2 z_3 \cdots z_i$, $z_k \in S$, $k = 1, 2, \cdots, i$, z_1 为给定的编码器的初始状态.

这些量之间有如下几项关系

$$y_i = f(z_i, x_i),\ \vec{d}_i = t(z_i, x_i),\ z_{i+1} = g(z_i, x_i) \qquad (1)$$

又有, 第 $i+1$ 步扫描所处的位置 $\vec{\Delta}_{i+1} = \vec{\Delta}_0 + \sum_{j=1}^{i} \vec{d}_j$. 其中 $\vec{\Delta}_0$ 是初始扫描位置. 不失一般性, 令 $\vec{\Delta}_0 = (0, 0)$.

译码器 F 也是一个七元组$(S', A, B, D', g', t', f')$, 其中各分量的意义及译码过程与 E 相似. 在译码过程中:

收稿日期: 2000-04-17; 修回日期: 2000-10-12

读取输入序列 $y = y_1 y_2 y_3 \cdots y_i,\ y_k \in B,\ k = 1, 2, \cdots, i$;
译码为输出序列 $x = \hat{x}_1 \hat{x}_2 \hat{x}_3 \cdots \hat{x}_i,\ \hat{x}_k \in A,\ k = 1, 2, \cdots, i$;
扫描增量序列 $\vec{d}' = \vec{d}'_1 \vec{d}'_2 \vec{d}'_3 \cdots \vec{d}'_i,\ \vec{d}'_k \in D',\ k = 1, 2, \cdots, i$;
状态序列 $z' = z'_1 z'_2 z'_3 \cdots z'_i,\ z'_k \in S',\ k = 1, 2, \cdots, i,\ z' = z'_1 z'_2 z'_3 \cdots z'_i \in S'$;
他们之间也有相应的关系:

$$\begin{aligned} \hat{x}_i &= f'(z'_i, y_i) \\ \vec{d}'_i &= t'(z'_i, y_i) \\ z'_{i+1} &= g'(z'_i, y_i) \end{aligned} \tag{2}$$

其第 $i+1$ 步扫描所处的位置 $\vec{\Delta}_{i+1} = \vec{\Delta}_0 + \sum_{j=1}^{i} \vec{d}'_j$. 其中 $\vec{\Delta}_0$ 是译码的初始扫描位置.

记译码后的图像记为 $\hat{I}$. 定义一般的失真测度 $d: A \times A \to R^+$. 设 $\hat{u}_{ij}$ 是译码后图像位置为像素点 (i, j) 的像素值. 且设 $\max_{u_{ij} \in A;\ \hat{u}_{ij} \in A} d(u_{ij}, \hat{u}_{ij}) = d_{\max} < \infty$.

令 I_N 表示 I 左上角大小为 $N \times N$ 的子图, 令 $d(I_N, \hat{I}_N) = \frac{1}{N^2} \sum_{i,j=1}^{N} d(u_{ij}, \hat{u}_{ij})$.
定义

$$d(I, \hat{I}) = \limsup_{N \to \infty} d(I_N, \hat{I}_N) \tag{3}$$

给定一个无限图像 I, 用 $d_s(R, I)$ 表示状态不超过 s 的所有编码译码组合中使 $d(I, \hat{I})$ 所能达到的最小值, 即

$$d_s(R, I) = \inf_{E(s), F(s)} d(I, \hat{I}) \tag{4}$$

其中, $E(s)$, $F(s)$ 分别代表状态数不超过 s 的所有可能的编码器集合及译码器集合. 定义 I 的率失真函数为

$$d(R/I) = \lim_{s \to \infty} d_s(R, I) \tag{5}$$

显然 $d(R/I)$ 是任何有限状态编码译码器在码率为 R 时的最小失真.

3　主要结论及其证明

引理 1[1]　设 u 为任一取值于有限集的无限长序列, 令 $2^{h_l(u) l}$ 表示 u 中不同的 l 长子串的数量, 则 $h_l(u)$ 存在极限.

$$h_l(u) \underset{l \to \infty}{\downarrow} h(u) \triangleq \lim_{l \to \infty} h_l(u) \tag{6}$$

在给出本文主要定理之前先简单介绍一下其中涉及的一种扫描方法, 即所谓标准扫描(standard scans, SS). SS 扫描是由四种基本的扫描方式通过递归的定义方式得到的[3]. 图 2 给出了其直观的表示.

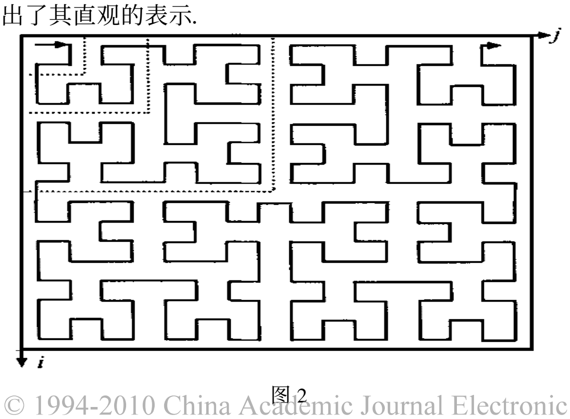


图 2

该扫描的最大特点就是按照边长为 2 的幂次的子块进行扫描, 且只有扫描完某一子块的全部后才进入另一子块进行扫描. 这个特点在证明中得到利用.

定理 1(逆编码定理)　对于如上叙述的无限大图像 I 及相关的有限状态编译码器等定义有

$$d(R/I) \geq \inf_{\tilde{I}:\ h(SS(\tilde{I})) \leq R} d(I, \tilde{I}) \tag{7}$$

其中, $SS(\tilde{I})$ 表示对 $\tilde{I}$ 进行标准扫描后得到的序列[3].

证明　设译码器 $F = (S', A, B, D', g', t', f')$, $|S'| = s$. 考虑 $\hat{I}$ 中大小为 $q \times q$ 的任意子图 $\hat{I}_q$, 令 $l = q \times q$, 设译码器 F 在译出 $\hat{I}_q$ 时共穿越它 m 次. 令相应的输入序列为 $y_{i_1}^{j_1}, y_{i_2}^{j_2}, \cdots, y_{i_m}^{j_m}$, 设 $1 \leq k \leq m$ 则 $\hat{I}_q$ 只与下列三项有关

ⅰ 第 k 次进入 $\hat{I}_q$ 时的状态 $(z'_{i_k}, \Delta'_{i_k})$;

ⅱ 第 k 次离开 $\hat{I}_q$ 时的状态 z_{1+j_k};

ⅲ 字符串 $y_{i_k}^{j_k}$.

ⅰ 中所有可能状态数的上界为 $(s \times 4q)^m$, ⅱ 中所有可能状态数的上界为 s^m, ⅲ 中所有可能字符串数的上界为 $\beta^l \times (l+1)^m$.

因此, 这样的不同 $\hat{I}_q$ 的最大数目(记为 $n(m)$) 不会超过三者的乘积.

$$n(m) \leq \beta^l \times (4s^2 q(l+1))^m \tag{8}$$

下面讨论 m 可能的取值范围: 我们注意到 Δ'_{i_k} 只能在 $\hat{I}_q$ 的边界, 且有 $(z'_{i_k}, \vec{\Delta}_{i_k}) = (z'_{i_r}, \vec{\Delta}_{i_r}) \Rightarrow k = r$, 否则译码器将导致死循环. 因此 m 的取值范围为 $1 \leq m \leq 4qs \triangleq M$. 所以 $\hat{I}$ 中所有可能不同 $\hat{I}_q$ 的数量(记为 $N(l)$) 有:

$$\begin{aligned} N(l) &\leq \sum_{m=1}^{M} n(m) = \sum_{m=1}^{M} \beta^l \times (4s^2 q(l+1))^m \\ &\leq \sum_{m=1}^{M} \beta^l \times (4s^2 q(l+1))^{4qs} = 4qs \times \beta^l \times (4s^2 q(l+1))^{4qs} \end{aligned}$$

因此我们有 $2^{l h_l(SS(\hat{I}))} \leq 16qs \times (4s^2 q(l+1))^{4qs} \times \beta^l$. 从而有

$$h(SS(\hat{I})) = \lim_{l \to \infty} h_l(SS(\hat{I})) \leq \log_2 \beta = R \tag{9}$$

故集合 $\{\tilde{I}: h(SS(\tilde{I})) \leq R\}$ 包含 $\hat{I}$, 从而

$$d(R/I) = \lim_{s \to \infty} \inf_{E(s), F(s)} d(I, \hat{I}) \geq \inf_{\tilde{I}:\ h(SS(\tilde{I})) \leq R} d(I, \tilde{I}) \tag{10}$$

引理 2[2]　设 u 为任一给定的取值于有限字母集 A 的 $(|A| = \alpha < \infty)$ 无限长序列, 则存在块编码器及相应块译码器, 即存在有限状态编译码器 C, D 使得码率为 R 且有

$$d(u, \hat{u}) \leq \inf_{w:\ h(w) \leq R} d(u, w) + \varepsilon \tag{11}$$

其中, $\hat{u}$ 为 u 经编码器 C 和解码器 D 所得序列. 利用上面引理的结果我们得到

定理 2(编码定理)　对于任意 $\varepsilon > 0$ 存在一对有限状态编译码器 E, F, 使得

$$d(I, \hat{I})_{E, F} \leq \inf_{\tilde{I}:\ h(SS(\tilde{I})) \leq R} d(I, \tilde{I}) + \varepsilon \tag{12}$$

证明　我们注意到, 二维情形下的有限状态编译码器与一维情形下的本质差异在于其多了扫描. 现在固定 E, F 的扫描规律为 SS, 则将文[2]中定理 2 的结果就可以应用于 $SS(I)$, 即将扫描后的图像 $SS(I)$ 视为要编译的序列 u, 从而可知结论正确.

4 讨论

(1) 一维时的有限状态编译码器与二维时的差异主要在于: 二维情形下要考虑所谓扫描的方法. 应当指出, 定理中的扫描方式 SS 并不是唯一可行的. 事实上, 任何扫描方法只要其扫描种类随所扫描子图边长增长规律不超过指数形式都可以. 因此, 一般"简单"的扫描方法都是可取的.

(2) 根据定理, 使 $d(R/I)=0$ 时的最小的 R 应该为有限状态编译码下无失真压缩极限. 通过简单的论证可以知道, 这时应有: $R=h(SS(I))$. 此特殊情形结果与文[3]中所得到的结论一致.

(3) 定理中的结论都是建立在极限意义上的(编译码器的状态要趋向无穷), 对于有限状态编译码器的压缩性能还有待研究. 但可以想象, 这与串匹配编码有相当联系[2,4].

(4) 本文的结论可以适当推广到一般的 N 维数据情形.

5 总结

本文讨论了无限大图像的失真率函数的有关理论, 它在不需要给定概率特性的情况下给出了失真率函数简洁的表达, 虽然其中结果是建立在极限的意义上, 但是它对实际图像压缩领域也是有着很强的指导意义.

参考文献:

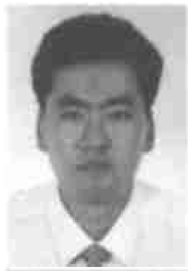
[1] Jacob Ziv. Distortion-rate theory for individual sequences [J]. IEEE TRANSACTIONS ON INFORMATION THEORY, March 1980, IT-26(2).

[2] Jacob Ziv. Coding theorem for individual sequences [J]. IEEE TRANSACTIONS ON INFORMATION THEORY, July 1978, IT-24(4).

[3] Abraham Lempel. Compression of two-dimensional data [J]. IEEE TRANSACTIONS ON INFORMATION THEORY, January 1986, IT-32(1).

[4] Ilan Sadeh. Universal data compression algorithm based on approximate string matching [J]. PROBABILITY IN THE ENGINEERING AND INFORMATIONAL SCIENCES 1996, 10: 465 - 486.

作者简介:



潘旭山 男. 1974 年生于上海, 上海交通大学应用数学系博士生. 主要研究方向: 信息理论与信息处理, 随机场, 图像压缩.

叶中行 男. 1946 年出生于上海, 教授, 美国康奈尔大学博士. 主要研究方向: 信息理论与信息处理, 随机场, 金融数学和智能计算.