



数字图像处理

Digital Image Processing

信息工程学院

School of Information Engineering

7.3 算术编码

郭志强 主讲

7.3 算术编码

算术编码假设，对于一个独立信源来说，任一由信源符号组成的长度为 N 的序列的发生**概率之和等于1**。

根据信源符号序列的概率，把 $[0, 1]$ 区间划分为互不重叠的子区间，**子区间的宽度恰好等于各符号序列的概率**，这样，每个子区间内的任意一个实数都可以用来表示对应的符号。

显然，一串符号序列发生的概率越大，对应的子区间就越宽，表达它所用的比特数就越少，因而相应的码字就越短。

7.3 算术编码

算术编码过程：

1. 建立概率模型，即通过扫描统计，获得各信源符号的概率大小
2. 编码过程，即扫描符号序列，依次分割相应的区间，最终得到符号序列所对应的码字。
3. 编码过程的数学描述

设由 M 个信源符号 $X = x_1 x_2 \dots x_m$ 组成的长度为 N 的输入符号序列中，各信源符号的概率分布为 P_j ($j = 1, 2, \dots, M$; $k = 1, 2, \dots, N$; $M \leq N$) , $[0, 1)$ 为对输入符号序列进行算术编码的初始区间，则对第 k 个输入符号进行算术编码的子分区间 $[low, high)$ 定义为：

7.3 算术编码

$$low = low + range \times rang_low \quad (7.4a)$$

$$high = low + range \times rang_high \quad (7.4b)$$

$$range = high - low \quad (7.4c)$$

其中：等号左边的 low 为当前被编码字符低值

等号左边的 $high$ 为当前被编码字符的高值

等号右边的 low 为上一个被编码字符低值

$rang_low$ 和 $rang_high$ 分别为当前被编码字符已给定出现的概率范围的低值和高值。

7.3 算术编码

举例:设有一个四信源符号的五符号输入序 $a_1a_2a_2a_3a_4$ 。

① 建立信源符号集的概率模型：通过扫描可知信源符号 $a_1a_2a_3a_4$ 的出现概率依次为0.2、0.4、0.2和0.2。

② 编码过程：

符号	概率	范围
a_1	0.2	[0 0.2)
a_2	0.4	[0.2 0.6)
a_3	0.2	[0.6 0.8)
a_4	0.2	[0.8 1)

要求编码的序列为： $a_1 a_2 a_2 a_3 a_4$

$$a_1: \text{low} = \text{low} + \text{range} \times \text{range_low} = 0 + 1 \times 0 = 0$$

$$\text{high} = \text{low} + \text{range} \times \text{range_high} = 0 + 1 \times 0.2 = 0.2$$

$$\text{range} = \text{high} - \text{low} = 0.2 - 0 = 0.2$$

$$a_2: \text{low} = \text{low} + \text{range} \times \text{range_low} = 0 + 0.2 \times 0.2 = 0.04$$

$$\text{high} = \text{low} + \text{range} \times \text{range_high} = 0 + 0.2 \times 0.6 = 0.12$$

$$\text{range} = \text{high} - \text{low} = 0.12 - 0.04 = 0.08$$

$$a_2: \text{low} = \text{low} + \text{range} \times \text{range_low} = 0.04 + 0.08 \times 0.2 = 0.056$$

$$\text{high} = \text{low} + \text{range} \times \text{range_high} = 0.04 + 0.08 \times 0.6 = 0.088$$

$$\text{range} = \text{high} - \text{low} = 0.088 - 0.056 = 0.032$$

$$a_3: \text{low} = \text{low} + \text{range} \times \text{range_low} = 0.056 + 0.032 \times 0.6 = 0.0725$$

$$\text{high} = \text{low} + \text{range} \times \text{range_high} = 0.056 + 0.032 \times 0.8 = 0.0816$$

$$\text{range} = \text{high} - \text{low} = 0.0816 - 0.0725 = 0.0091$$

$$a_4: \text{low} = \text{low} + \text{range} \times \text{range_low} = 0.0725 + 0.0091 \times 0.8 = 0.07978$$

$$\text{high} = \text{low} + \text{range} \times \text{range_high} = 0.0725 + 0.0091 \times 1 = 0.0816$$

$$\text{range} = \text{high} - \text{low} = 0.0816 - 0.07978 = 0.00182$$

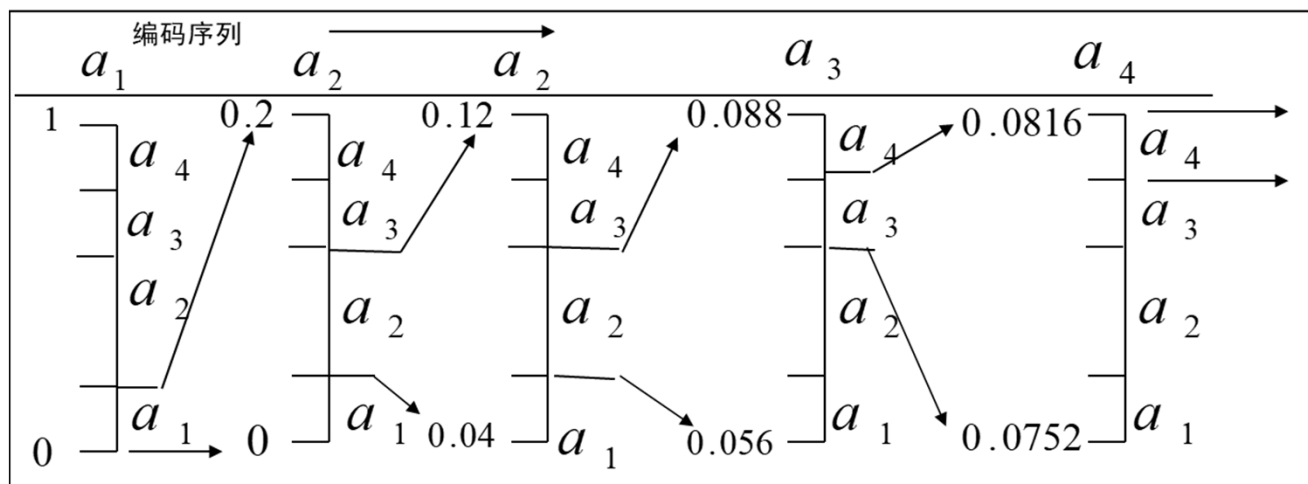


图7.1 算术编码过程图示

7.3 算术编码

```
in='00000011111010101';
pr=[0.5 0.5]; %各字符出现的概率
temp=[0.0 0.5 1.0];
original=temp;
n=length(in);
%编码
for i=1:n
    width=temp(3)-temp(1);
    w=temp(1);
    switch in(i)
        case '0'
            m=1;
        case '1'
            m=2;
    end
    temp(1)=w+original(m)*width;
    temp(3)=w+original(m+1)*width;
    left=temp(1);
    right=temp(3);
    fprintf('left=%.6f',left);
    fprintf(' ');
    fprintf('right=%.6f\n',right);
end
encode=(temp(1)+temp(3))/2
```



谢谢

THANK YOU