Tópicos sobre Teoria da informação e codificação de fonte

Sistemas Multimédia

Verão 2008/2009

Tópicos a abordar

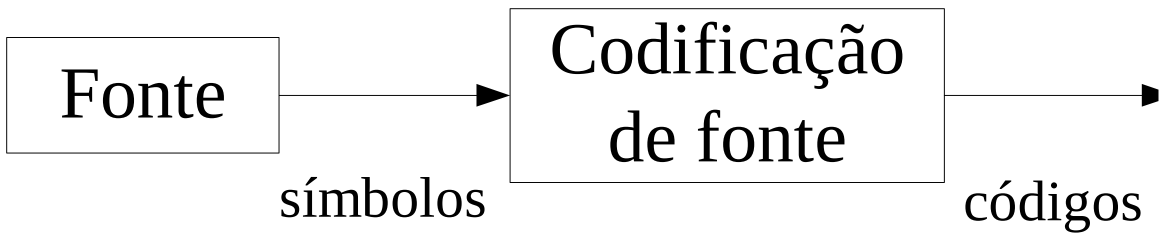
1. Teoria da informação
   - Informação própria
   - Entropia

2. Codificação de fonte
   - Código de Huffman
   - Teorema da codificação de fonte
   - Eficiência do código

3. Análise de resultados

Codificação de fonte

- **Codificação de fonte - *source coding***
  - descrição de acontecimentos produzidos por determinada fonte
  - procura-se minimizar o comprimento médio do código
  - cada símbolo é representado por um código



Medidas da teoria da informação

Seja X uma variável aleatória discreta com concretizações $\{x_1, x_2, \ldots, x_N\}$

A informação própria de cada concretização é dada por:

$$I(x_i) = -\log_2 p(x_i) \ \text{[bit]}$$

A entropia de X é o valor esperado (médio) da informação própria de cada concretização:

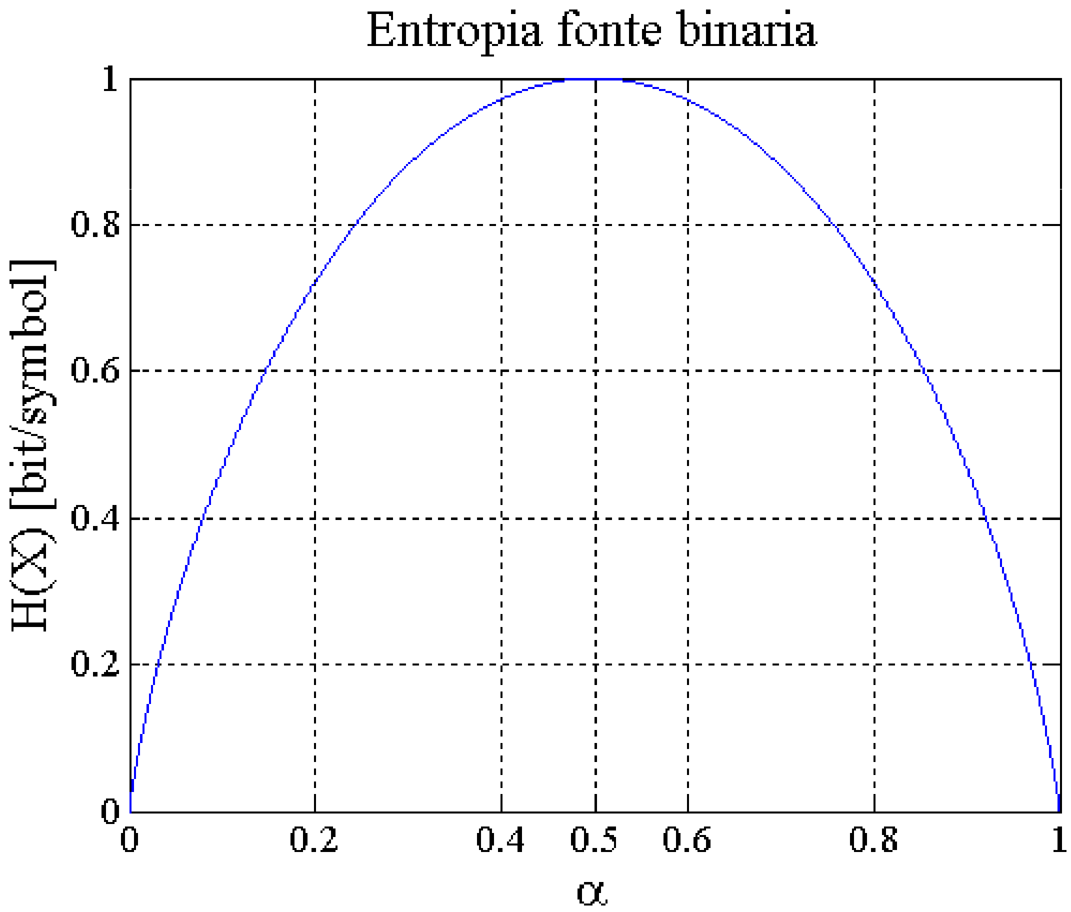
$$\begin{aligned} H(X) &= E[-\log_2 p(x)] \\ &= -\sum_x p(x) \log_2 p(x) \ \text{[bit/symbol]} \end{aligned}$$

Entropia de fonte binária sem memória

Considerando uma fonte binária, em que um dos símbolos tem probabilidade α e o outro tem 1-α tem-se

$$H(X) = \alpha \log_2\left(\frac{1}{\alpha}\right) + (1-\alpha)\log_2\left(\frac{1}{1-\alpha}\right)$$

A situação em que os dois símbolos são equiprováveis, corresponde à máxima incerteza/expectativa



Limites de variação da entropia

Para uma v.a. discreta X com N com concretizações

$\{x_1, x_2, \ldots, x_N\}$ a entropia varia entre os seguintes limites:

$$0 \leq H(X) \leq \log_2(N)$$

• O limite inferior é obtido quando um dos símbolos tem probabilidade 1

• O limite superior acontece quando os símbolos são equiprováveis (distribuição uniforme)

Cálculos de entropia

- Fonte quaternária (quatro símbolos) $\{a, b, c, d\}$
- A probabilidade de cada símbolo é dada por

$$probs = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$$

tem-se $H(X) = \log_2(4) = 2$

- Considerando agora $probs = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{3}, \frac{1}{6}\}$

tem-se $H(X) = 1,96(< 2)$

Teorema da codificação de fonte

Também conhecido como 1º Teorema de Shannon (*noiseless coding theorem*):

"É possível codificar, sem distorção, uma fonte de entropia H bit/símbolo, usando em média N=H+ε bits por símbolo, em que ε é uma quantidade arbitrariamente pequena."

A eficiência da codificação é dada por:

$$\frac{H(X)}{N} = \frac{H(X)}{H(X)+\epsilon}$$

em que N é o comprimento médio das palavras de código e H(X) é a entropia da fonte.

Codificação Entrópica

- O objectivo da codificação entrópica (de fonte) consiste em codificar uma fonte por forma a que o seu *bit rate* médio seja H bits/símbolo.

- Eficiência dum código com N bits por símbolo aplicado a uma fonte de entropia H:

$$\frac{H}{N}$$

- Em 1952, D. Huffman propôs um algoritmo de codificação que atribui códigos de comprimento variável aos símbolos da fonte

- O comprimento da palavra de código que representa um dado símbolo, é função da quantidade da informação associada ao símbolo

Comprimento médio

- O comprimento médio das palavras de um código define-se como:

$$\begin{aligned} N &= E[-l(x_i)] \\ &= \sum_{x_i} p(x_i) l(x_i) \ \text{[digit/symbol]} \end{aligned}$$

em que $l(x_i)$ é o comprimento da palavra de código que representa o símbolo x_i

Representa o número médio de dígitos utilizado para representar cada símbolo da fonte.

Análise do comprimento médio

- Fonte quaternária
  - símbolos $\{a, b, c, d\}$
  - probabilidades $probs = \{\frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{6}\}$
  - entropia $H(X) = 1.96$

Estabelecem-se dois códigos para os símbolos desta fonte:

$$\{0, 10, 110, 111\}$$

com comprimento médio N=2.0833 e eficiência de 94.08 %

$$\{0, 10, 110, 1110\}$$

com comprimento médio N=2.25 e eficiência de 87.11 %

Algoritmo de codificação de Huffman

- Ordenar os símbolos por probabilidade decrescente e considerá-los como nós duma árvore

- Enquanto houver mais do que um nó:
  - Agrupar os dois nós de menor probabilidade e formar um novo nó com probabilidade igual à soma dos dois nós que se agruparam;
  - Arbitrariamente atribuir os códigos 0 e 1 a cada ramo dos nós que se juntaram

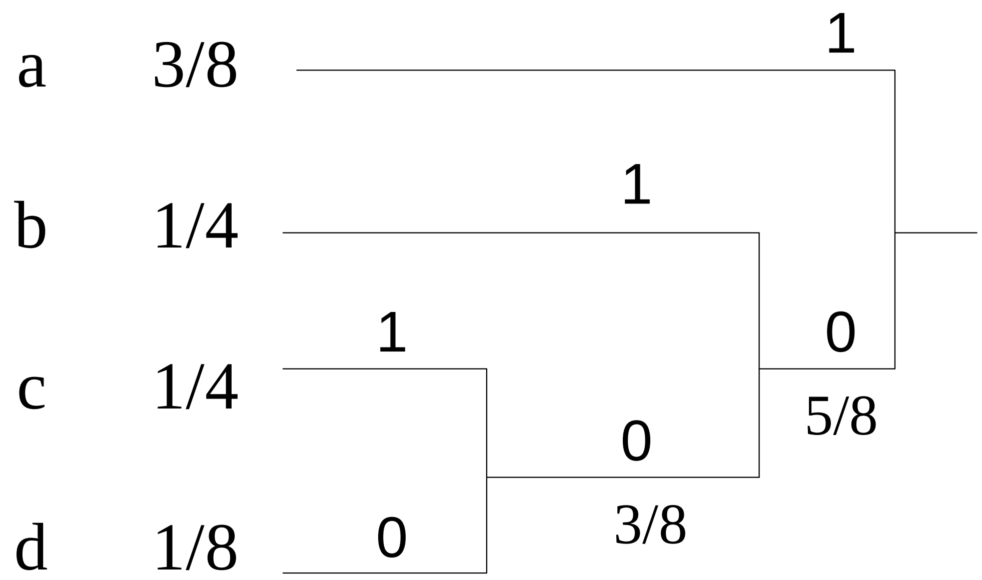
- O código de cada símbolo obtém-se pela leitura sequencial da raiz até ao nó final a que pertence o símbolo em causa

Código óptimo e ideal

- O código de Huffman, é óptimo porque garante que

$$H(X) \leq N < H(X) + 1$$

- No caso em que N=H(X) o código diz-se <u>ideal</u>
- Exemplo de código de Huffman:



a	1
b	01
c	001
d	000

Análise de resultados de codificação

Medidas de avaliação

- Taxa de compressão
  - Razão de compressão

$$\frac{d_c}{d_o}$$

  - Percentagem removida

$$\left(1 - \frac{d_c}{d_o}\right) 100$$

  - Bits por pixel (byte)

$$8\frac{d_c}{d_o} \quad [\mathrm{bpp}]$$

em que d_c é a dimensão em bytes da imagem codificada e d_o é a dimensão em bytes da imagem original.

Medidas de avaliação

- Distorção
  - Relação sinal/ruído (SNR – signal to noise ratio)

$$SNR = 10\log_{10}\left(\frac{P_x}{P_r}\right) \qquad [dB]$$

P_x é a potência da imagem original I , P_r é a potência do ruído (erro) $r = I - \tilde{I}$

$$P_x = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I^2(m,n)$$

$$P_r = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} (I(m,n)-\tilde{I}(m,n))^2 = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} r^2(m,n)$$

Medidas de avaliação

Nalguns casos, na SNR usa-se apenas a potência AC da imagem, como potência do sinal:

$$SNR = 10\log_{10}\left(\frac{P_{x_{AC}}}{P_r}\right) \qquad [dB]$$

$$P_{x_{AC}} = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left(I(m,n) - m_x\right)^2$$

onde m_x é o valor médio da imagem original.

$$m_x = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m,n)$$

Huffman sobre Calgary Corpus

FILENAME	ORIGINAL	PACKED	REM.PERC.(%)	BITS/BYTE	ENTROPY
bib	111261	72933	34.4	5.24	5.20
book1	768771	440041	42.8	4.58	4.52
book2	610856	369216	39.6	4.86	4.79
geo	102400	73394	28.3	5.73	5.64
news	377109	246793	34.6	5.24	5.18
obj1	21504	16415	23.7	6.11	5.94
obj2	246814	195180	20.9	6.33	6.26
paper1	53161	33491	37.0	5.04	4.98
paper2	82199	47833	41.8	4.66	4.60
paper3	46526	27415	41.1	4.71	4.66
paper4	13286	7966	40.0	4.80	4.69
paper5	11954	7549	36.8	5.05	4.93
paper6	38105	24165	36.6	5.07	5.00
pic	513216	122039	76.2	1.90	1.21
progc	39611	26042	34.3	5.26	5.19
progl	71646	43217	39.7	4.83	4.77
progp	49379	30456	38.3	4.93	4.86
trans	93695	65414	30.2	5.59	5.53
AVERAGE	180639	102753	37.6	4.99	4.89

Calgary Corpus: conjunto normalizado de ficheiros de teste
http://corpus.canterbury.ac.nz/descriptions/#calgary

Huffman sobre Canterbury Corpus

FILENAME	ORIGINAL	PACKED	REM.PERC.(%)	BITS/BYTE	ENTROPY
alice29.txt	152089	88116	42.1	4.64	4.56
asyoulik.txt	125179	76010	39.3	4.86	4.80
cp.html	24603	16306	33.7	5.30	5.23
fields.c	11150	7139	36.0	5.12	5.00
grammar.lsp	3721	2271	39.0	4.88	4.63
kennedy.xls	1029744	503121	51.1	3.91	3.57
lcet10.txt	426754	251544	41.1	4.72	4.67
plrabn12.txt	481861	276890	42.5	4.60	4.53
ptt5	513216	122039	76.2	1.90	1.21
sum	38240	26455	30.8	5.53	5.32
xargs.1	4227	2706	36.0	5.12	4.89
AVERAGE	255526	124782	42.5	4.60	4.40

Canterbury Corpus: conjunto normalizado de ficheiros de teste
http://corpus.canterbury.ac.nz/descriptions/#cantrbry

Codificação da imagem lena

- Lena 256x256

H(X) = 7.46 bit / símbolo

H(X,Y)/2 = 6.22 bit / símbolo

Codificador	Taxa de compressão [bpp]	Relação sinal/ruído SNR [dB]
Huffman	7.51	infinita
Huffman Adaptativo	7.49	infinita
Aritmético	7.48	infinita
Aritmético ordem 1	6.26	infinita
JPEG	2.07	26.16
JPEG	1.42	22.77
JPEG	0.96	19.92
JPEG	0.85	19.09
JPEG	0.63	17.32
JPEG	0.35	13.68

Codificação JPEG – imagem lena

Original

2.07 bpp 26.16 dB

0.96 bpp 19.92 dB

0.85 bpp 19.09 dB

0.63 bpp 17.32 dB

0.35 bpp 13.68 dB

Uso do PSPro – Conversão para formato JPEG