

2.0 - Compressão de dados Multimídia

prof. Fábio Engel

fabioe@utfpr.edu.br



1 Compressão de dados Multimídia

2 Algoritmo LZW

- Objetivo principal da compressão.
 - ▶ Redução do espaço de armazenamento necessário.
- Classificação.
 - ▶ Com perdas (*Lossy*).
 - ▶ Sem perdas (*Lossless*).

Terminologia:

- **Compressão** de dados.
 - ▶ Processo de redução da quantidade de bits para representar informação.
 - ▶ Exemplo:
 - Comprimir arquivos com 7-zip, winrar, gzip.
- **Compactação** de dados.
 - ▶ Unir dados.
 - ▶ Em geral não há perdas de dados.
 - ▶ Exemplo:
 - Desfragmentador de disco.
 - Compactar arquivos em um diretório/arquivo único.

Terminologia:

- **Codificação** de dados.

- ▶ Modificação do sinal/dado em um esquema para tornar mais apropriado:
- ▶ Transmissão.
- ▶ Armazenamento.
- ▶ Exemplos:
 - Vídeos codificados no formato YUV.
 - Codificação de caracteres.

Compressão de dados Multimídia

- Os termos são distintos entre si, mas no entanto:
 - ▶ Compactação, Compressão e Codificação normalmente são empregadas em conjunto.
 - ▶ Esquema geral de compressão.



Compressão de dados Multimídia

- Princípios de Compressão.

- 1 Redundância de Dados.

- ▶ **Espacial:** em imagens/*frames*.

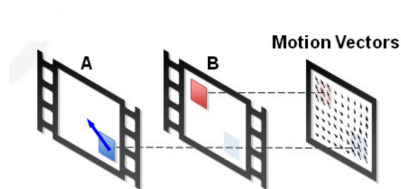
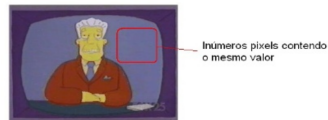
- Semelhanças nos *pixels* adjacentes.
- *Intra-frame prediction*.

- ▶ **Temporal:** vídeos.

- *Motion estimation*.

- 2 Velocidade da Compressão.

- ▶ Complexidade de implementação.
- ▶ Aplicações de tempo-real.



- Taxa de compressão:

- ▶ B_0 = número de bits para representar a informação.
- ▶ B_1 = número de bits após a compressão.
- ▶ Taxa de compressão (T):
 - $T = \frac{B_0}{B_1}$

- ③ Reversível.

- ▶ Descompressão deve recuperar uma exata cópia dos dados.
- ▶ Ou uma aproximação (compressão com perdas).

Técnicas de compressão:

- Codificação de entropia.
 - ▶ Entropia:
 - Medida da desordem em um sistema.
 - ▶ Objetivo é:
 - Identificar símbolos frequentes.
 - Atribuir códigos menores para eles.
 - Em geral possui característica “sem perdas”.

Técnicas de compressão:

- Codificação na origem.
 - ▶ Remove dados irrelevantes.
 - Ênfase na semântica dos dados.
 - Em geral acarreta **perdas**.
- Codificação “Híbrida”.
 - ▶ JPG.
 - ▶ PNG.

- Codificação de entropia.
 - ▶ Codificação *run-length* (de carreira).
 - Forma simples: grupos de símbolos/caracteres que se repetem.
 - ▶ Codificação *variable run-length*.
 - Algoritmo de Shannon-Fano.
 - Algoritmo de Huffman.
 - Codificação aritmética.
 - ▶ Baseada em dicionários.
 - Lempel-Ziv-Welch (LZW).

Algoritmo LZW.

- Algoritmo baseado em “dicionário”.
 - ▶ Tamanho fixo.
 - ▶ Ao invés de codificar caracteres, codifica strings.
 - ▶ Adaptativo.
 - Constrói o dicionário dinamicamente.
- LZW empregado nos formatos:
 - ▶ GIF.
 - ▶ TIFF.

Algoritmo LZW.

- ❶ Inicialmente, o dicionário contém todos os caracteres possíveis.
 - ▶ Uma *string* S inicialmente é vazia.
- ❷ A cada novo caractere C recebido/lido, $S + C$ existe no dicionário?
 - ▶ **Sim.** Então:
 - ❶ $S = S + C$.
 - ▶ **Não.** Então:
 - ❶ Escrever/enviar o código de S .
 - ❷ Adicionar $S + C$ ao dicionário.
 - ❸ $S = C$.

Algoritmo LZW

- Exemplo: w a b b a w a b b a
- Rodada: 0.

Índice	Dicionário
1	a
2	b
3	w

- São adicionados todos os caracteres.

Algoritmo LZW

- Exemplo: **w** a b b a w a b b a
- Rodada: 1.

Índice	Dicionário
1	a
2	b
3	w

- $C = w$
- $S = S + C$, então $S = w$
- $S + C$ já existe.
 - $S = S + C$, então $S = w$

Algoritmo LZW

- Exemplo: w **a** b b a w a b b a
- Rodada: 2.

Índice	Dicionário
1	a
2	b
3	w
4	wa

- $C = a$
- $S = S + C$, então $S = wa$
- $S + C$ Não existe.
 - ▶ Saída (código): 3
 - ▶ Adiciona $S + C$, ou seja, wa ao dicionário.
 - ▶ $S = C$, então $S = a$

- Código: 3

Algoritmo LZW

- Exemplo: w a **b** b a w a b b a
- Rodada: 3.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab

- $C = b$
- $S = S + C$, então $S = ab$
- $S + C$ Não existe.
 - ▶ Saída (código): **1**
 - ▶ Adiciona $S + C$, ou seja, **ab** ao dicionário.
 - ▶ $S = C$, então $S = b$

- Código: **31**

Algoritmo LZW

- Exemplo: w a b **b** a w a b b a
- Rodada: 4.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb

- $C = b$
- $S = S + C$, então $S = bb$
- $S + C$ Não existe.
 - ▶ Saída (código): 2
 - ▶ Adiciona $S + C$, ou seja, bb ao dicionário.
 - ▶ $S = C$, então $S = b$

- Código: 312

Algoritmo LZW

- Exemplo: w a b b **a** w a b b a
- Rodada: 5.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba

- $C = a$
- $S = S + C$, então $S = ba$
- $S + C$ Não existe.
 - ▶ Saída (código): 2
 - ▶ Adiciona $S + C$, ou seja, ba ao dicionário.
 - ▶ $S = C$, então $S = a$

- Código: 3122

Algoritmo LZW

- Exemplo: w a b b a **w** a b b a
- Rodada: 6.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw

- $C = w$
- $S = S + C$, então $S = aw$
- $S + C$ Não existe.
 - ▶ Saída (código): **1**
 - ▶ Adiciona $S + C$, ou seja, **aw** ao dicionário.
 - ▶ $S = C$, então $S = w$

- Código: **31221**

Algoritmo LZW

- Exemplo: w a b b a w **a** b b a
- Rodada: 7.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw

- $C = a$
- $S = S + C$, então $S = wa$
- $S + C$ já existe.
 - $S = S + C$, então $S = wa$

- Código: **31221**

Algoritmo LZW

- Exemplo: w a b b a w a **b** b a
- Rodada: 8.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab

- $C = b$
- $S = S + C$, então $S = wab$
- $S + C$ Não existe.
 - ▶ Saída (código): 4
 - ▶ Adiciona $S + C$, ou seja, wab ao dicionário.
 - ▶ $S = C$, então $S = b$

- Código: 312214

Algoritmo LZW

- Exemplo: w a b b a w a b **b** a
- Rodada: 9.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab

- $C = b$
- $S = S + C$, então $S = bb$
- $S + C$ já existe.
 - ▶ $S = S + C$, então $S = bb$

- Código: **312214**

Algoritmo LZW

- Exemplo: w a b b a w a b b a
- Rodada: 10.

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab
10	bba

- $C = a$
- $S = S + C$, então $S = bba$
- $S + C$ Não existe.
 - ▶ Saída (código): 4
 - ▶ Adiciona $S + C$, ou seja, bba ao dicionário.
 - ▶ $S = C$, então $S = a$

- Código: 3122146

Algoritmo LZW

- Exemplo: w a b b a w a b b a

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab
10	bba

- Stream* finalizou.
- $S = a$
- Como S sobrou, adiciona-se o código de S ao final.
 - ▶ Saída (código): 31221461

- Código: **31221461**

Decodificação LZW.

- ❶ Defina o primeiro código recebido c , e defina $e(c)$ como a entrada (*string*) no dicionário para c .
- ❷ Escreva/exiba $e(c)$.
- ❸ Guarde $p = c$.
- ❹ Em c é recebido o novo código.
- ❺ $e(c)$ existe no dicionário?

► Sim:

- I. Escreva/exiba o caractere decodificado $e(c)$.
 - II. Guarde a entrada anterior $e(p)$ em *ant*.
 - III. Guarde o primeiro caractere de $e(c)$ em *prim*.
 - IV. Adicione $ant + prim$ ao dicionário.
- ❻ Volte ao passo 3 (se houver novo código)

► Não:

- I. Guarde $e(p)$ em *ant*.
- II. Guarde o primeiro caractere de (p) em *prim*.
- III. Escreva/exiba $ant + prim$ e adicione ao dicionário.

- Exemplo: **3** 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w

- Rodada: 1.
 - ▶ $c = 3$
 - ▶ $e(c) = w$
 - ▶ exiba $e(c)$
 - ▶ $p = 3$
- Saída decodificada: w

- Exemplo: 3 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa

- Rodada: 2.
 - ▶ $p = 3$
 - ▶ $c = 1$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva a
 - $ant \leftarrow w$
 - $prin \leftarrow a$
 - Adicione wa ao dicionário.
- Saída decodificada: w a

- Exemplo: 3 1 **2** 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab

- Rodada: 3.
 - ▶ $p = 1$
 - ▶ $c = 2$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva b
 - $ant \leftarrow a$
 - $prin \leftarrow b$
 - Adicione ab ao dicionário.
- Saída decodificada: w a b

- Exemplo: 3 1 2 **2** 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb

- Rodada: 4.
 - ▶ $p = 2$
 - ▶ $c = 2$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva b
 - $ant \leftarrow b$
 - $prin \leftarrow b$
 - Adicione bb ao dicionário.
- Saída decodificada: w a b b

- Exemplo: 3 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba

- Rodada: 5.
 - ▶ $p = 2$
 - ▶ $c = 1$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva a
 - $ant \leftarrow b$
 - $prin \leftarrow a$
 - Adicione ba ao dicionário.
- Saída decodificada: w a b b a

- Exemplo: 3 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw

- Rodada: 6.
 - ▶ $p = 1$
 - ▶ $c = 4$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva wa
 - $ant \leftarrow a$
 - $prin \leftarrow w$
 - Adicione aw ao dicionário.
- Saída decodificada: w a b b a w a

- Exemplo: 3 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab

- Rodada: 7.
 - ▶ $p = 4$
 - ▶ $c = 6$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva bb
 - $ant \leftarrow wa$
 - $prin \leftarrow b$
 - Adicione wab ao dicionário.
- Saída decodificada: w a b b a w a b b

- Exemplo: 3 1 2 2 1 4 6 1

Índice	Dicionário
1	a
2	b
3	w
4	wa
5	ab
6	bb
7	ba
8	aw
9	wab
10	bba

- Rodada: 8.
 - ▶ $p = 6$
 - ▶ $c = 1$
 - ▶ $e(c)$ Existe no dicionário.
 - Escreva a
 - $ant \leftarrow bb$
 - $prin \leftarrow a$
 - Adicione **bba** ao dicionário.
- Saída decodificada: w a b b a w a b b a

- Exercício em sala.

- ① Utilize o algoritmo LZW para codificar A B A B B A B C A B A B B A.
- ② Qual a taxa de compressão, considerando 7 bits ASCII por caractere do exercício anterior?
- ③ Data a tabela de dicionário a seguir, decodifique utilizando o algoritmo LZW o seguinte código: 1 4 1 3 2 7.

Índice	Dicionário
1	a
2	b
3	w

Referências e Contato

- Apresentação baseada em:
 - ▶ Material de aula de prof. Giron.
- Contato:
 - ▶ fabioe@utfpr.edu.br