

Compressão de sinais sem perda com o uso de palavras de tamanho variável

Mateus J. Martins¹, Jan Frans W. Slaets¹

¹Depto. de Física e Informática - IFSC/USP
Av. Dr. Carlos Botelho, 1469 - 13560-970 - São Carlos (SP)
E-mail: mateus@uspfc.usp.br

Resumo - Um novo algoritmo de compressão de dados sem perda de informação, baseado na adequação automática do tamanho da “palavra” usada na amplitude do sinal é apresentado e avaliado para armazenamento de sinais de Tomografia por Ressonância Magnética. Uma comparação com outras implementações de compressores de uso geral é apresentada, demonstrando um desempenho superior tanto na taxa de compressão, quanto no tempo de execução.

Abstract - A new data compression algorithm without information losses based on the automatic adaptation of the used word size is presented and evaluated with MRI data. A comparison with others general purpose compression is presented to show the superior performance in the compression rate as well as in execution time.

Introdução

Várias técnicas de compressão, com ou sem perda de informação, são disponíveis atualmente e utilizadas principalmente para diminuir, tanto a ocupação de espaço nos meios de armazenamento, quanto o tempo de transferência dessas informações em redes entre computadores. Normalmente os compressores de dados sem perda, de uso geral (ZIP, ARC, GZIP) utilizam a técnica LZW (Lempel-Ziv & Welch)¹. Esses compressores apresentam uma alta taxa de compressão, porém não são otimizados para sinais que possuam uma grande variação de sua dinâmica durante pequenos intervalos de tempo. Dessa forma, boa parte dos dados armazenados não necessitam de todos os bits da palavra do computador. Assim, desenvolvemos a técnica DAC² (Dados com Compressão) permitindo uma compactação maior para sinais possuindo esta característica.

Metodologia

O uso de “palavras” de comprimento variável é a técnica adotada neste novo método. Como a maior parte dos computadores convencionais não suporta palavras de tamanho variável, o mesmo é implementado de forma artificial. O algoritmo DAC utiliza TAGs de informação para permitir um tamanho de palavra variável. A estrutura TAG (figura 1) contém o tamanho da palavra (b) em número de “bits” e o número de palavras (n) armazenadas em subsequentes registros de b bits. Durante a compressão, a constante avaliação das TAGs adicionadas, permite o armazenamento dos dados em conjuntos diminuindo o espaço utilizado. A última palavra do conjunto é utilizada como um CRC dos dados, para a sua verificação durante o procedimento de descompressão dos mesmos.

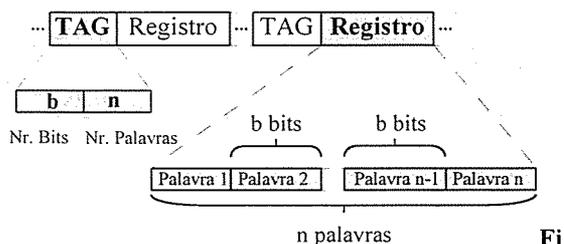


Figura 1 - Formato do TAG e registro utilizado na nova técnica de compressão

Resultados

O método descrito, foi utilizado para a compressão dos dados de Ressonância Magnética do Laboratório de RM do IFSC-USP. Um típico conjunto de dados de 256x256 de 32bits complexo (Real/Imag) apresenta 524292 bytes, sendo seu valor reduzido para 122604 bytes (76,6% de compressão), contra os 167633 bytes (68%) obtidos com o PKZIP, usando o algoritmo Lempel-Zip & Welch do mesmo arquivo. Os outros compressores avaliados (GZIP, ARJ, LHA, ARC e COMPRESS) tiveram uma taxa de compressão ainda inferior ao PKZIP, Tabela I.

Discussão e Conclusões

Embora essa técnica tenha sido implementada para dados de RM, nada impede seu uso com outros tipos de sinais, principalmente biológicos, cuja dinâmica apresente grandes variações. A técnica DAC apresenta uma grande taxa de compressão comparando-se a compressores de uso geral largamente utilizados, como o PKZIP, ARC, GZIP, mesmo utilizando opções de compressão máxima. Além de possuir uma taxa de compressão superior, o tempo de compressão e

descompressão é ligeiramente inferior ao do PKZIP utilizando a opção de velocidade máxima.

Referências

¹ WELCH, TERRY A. A Technique for High performance Data Compression. IEEE Computer., v. 17, n. 6, p. 8-19, 1984.

² MARTINS, M. J. Desenvolvimento de um Tomógrafo de Ressonância Magnética: Integração e Otimização. São Carlos, 1995. Tese (Doutorado) - Instituto de Física de São Carlos: Universidade de São Paulo.

Tabela I - Comparação entre vários programas de compressão em um micro 486DX2 - 66MHz, sobre um arquivo de RM de 256 x 256 aquisições

Arquivos de 524.292 bytes	Tamanho final (bytes)	Tempo de Compressão (seg)	Tempo de Descompressão (seg)	Taxa de Compressão (%)
DAC	122.604	5,0	4,0	76.6
PKZIP (max)	163.928	11,9	0,7	68.7
PKZIP	167.633	7,0	1,0	68.0
GZIP (max)	164.471	109,9	2,3	68.6
GZIP	184.009	11,0	2,4	64.9
ARJ	173.699	16,4	2,2	66.9
LHA	168.711	5,0	0,9	67.8
ARC	193.301	4,3	1,7	63.1
COMPRESS	244.594	16,0	1,1	53.3
TAR	532.480	0,7	0,4	-1.6