2ª VERSÃO
Concluída em 28/10/2007
CE 08:001.06/ABNT

Comitê Consultivo para Sistemas Espaciais de Dados

RELATÓRIO AFEITO À PADRÕES DE SISTEMAS ESPACIAIS DE DADOS

COMPRESSÃO DE DADOS LIVRE DE PERDAS

CCSDS 120.0-G-2

LIVRO VERDE

Dezembro 2006

.

N.C.E. – Este documento resulta da tradução de um documento original, escrito em língua Inglesa. O conteúdo e o formato do documento original foram preservados neste documento, sempre que possível. Todo o trabalho de tradução e de edição deste documento foi realizado pela Comissão de Estudos CE 08:001.06, dedicada a Sistemas Espaciais de Transferência de Dados e de Informação, pela ABNT – Associação Brasileira de Normas Técnicas. © CE 08:001.06/ABNT, INPE/MCT, São José dos Campos, SP, Brasil – Outubro 2007.

COLABORADORES DESTA EDIÇÃO

Por ocasião em que foi realizada a editoração deste documento interno participaram deste trabalho:

1) Os seguintes MEMBROS da Comissão de Estudo em Sistemas Espaciais de Transferência de Dados e de Informação – ABNT/CB08/SC01/CE06:

Eduardo W. Bergamini	Coordenador Editorial da Comissão	INPE/MCT
Elvis Paulo da Silva	Membro	EMBRAER
Gerald J. F. Banon	Membro	INPE/MCT
José Bastos Mollica	Membro	ANATEL
Kátia Pádua Thomaz	Membro	ECI/UFMG
Marilia Vidigal da Costa Souza	Membro	EMBRAER
Marco Antonio Grivet M. Maia	Membro	CETUC/PUC-RIO
Reginaldo Palazzo Júnior	Membro	DT/FEEC/UNICAMP
Sérgio Costa	Membro	AEL/IAE/CTA
Valéria Cristina M. N. Leite	Membro	AEL/IAE/CTA

2) As seguintes SECRETÁRIAS da coordenadoria desta Comissão de Estudo, que contribuíram de forma decisiva no extenso trabalho editorial e de correspondência associada a este documento:

Priscila Maria Rodrigues dosa Santos	Secretária	SME/TEC/INPE
Síntique Rodrigues dos Santos	Secretária	SME/TEC/INPE

São José dos Campos, Outubro de 2007

RECOMENDAÇÃO PARA OS PROCEDIMENTOS DE CONTROLE CCSDS PARA A ATRIBUIÇÃO DE CÓDIGOS GSCID

AUTORIDADE

Edição:	Livro Verde, Edição 2
Data:	Dezembro 2006
Local:	Washington, D.C., EUA

Este documento foi aprovado para publicação pelo Conselho de Gerenciamento do Comitê Consultivo para Sistemas Espaciais de Dados (CCSD0S) e representa o consenso dos especialistas técnicos das Agências Membro do CCSDS. O procedimento para revisão e autorização dos Relatórios CCSDS está detalhado no 'Manual de Procedimentos para o Comitê Consultivo para Sistemas Espaciais de Dados'.

Este documento é publicado e mantido por:

Secretariado do CCSDS
Divisão de Integração de Programas (Código M-3)
Administração Nacional de Aeronáutica e Espaço - NASA
Washington, DC 20546, EUA

NOTA DA COORDENAÇÃO EDITORIAL

Este documento representa o resultado da tradução de um documento original, escrito em língua Inglesa. O conteúdo e o formato do documento original foram preservados neste documento, sempre que possível. O conteúdo residual em língua Inglesa que este documento possui é resultante do texto do documento original. Este resíduo em língua Inglesa foi mantido por ter sido este recurso considerado pertinente, para os seus usuários. Este documento é também considerado de utilidade para pessoas não familiarizadas o suficiente com a língua Inglesa e que necessitam: consultar, conhecer e ou utilizar o documento original, escrito em língua Inglesa, a partir do qual foi realizada a tradução que resultou neste documento, presente. Primariamente, de forma equivalente ou correspondente, este documento constitui fonte de informação para o preparo de Normas ou de Documentos Normativos (Anexos, etc.) para o Brasil. A Comissão de Estudos CE 08:001.06, dedicada a Sistemas Espaciais de Dados e Informação, da Associação Brasileira de Normas Técnicas (ABNT), realizou o trabalho de tradução e de edição deste documento. Este trabalho faz parte do esforço de normalização espacial no Brasil, a cargo da Comissão CE 08:001.06/ABNT, dentro do seu escopo de especialização.

O documento original em língua Inglesa que deu origem a esta nova, 2ª edição deste documento, possui o título: "Lossless Data Compression", sob registro 'CCSDS 120.0-G-2', 'Green Book', de Dezembro de 2006, e foi elaborado como 'Recommendation for Space Data System Standards' pelo 'Consultative Committee for Space Data Systems (CCSDS)'. A primeira versão deste documento traduzido para o Português foi editada em Maio de 2002. © CE 08:001.06/ABNT, INPE/MCT, São José dos Campos, SP, Brasil – Outubro 2007

PREFÁCIO

Este documento é um Relatório CCSDS que contêm material básico e explanatório em apoio à Recomendação CCSDS em 'Compressão de Dados Livre de Perdas' (referência [1]).

Através do processo de evolução normal é esperado que a expansão, cancelamento, ou modificação deste Relatório possa vir a ocorrer. Este Relatório está, portanto, sujeito aos procedimentos de gerenciamento e controle de modificação de um documento CCSDS, tal como estão definidos na referência [2]. Versões atuais dos documentos CCSDS encontram-se disponíveis no site Web CCSDS:

http://www.ccsds.org

Questões relacionadas ao conteúdo ou ao estado deste relatório devem ser endereçadas ao Secretariado do CCSDS no endereço dado na Página i.

Quando esta publicação foi realizada, as Agências Membro e Observadoras ativas do CCSDS eram:

Agências Membro

- Agenzia Spaziale Italiana (ASI)/Italy.
- British National Space Centre (BNSC)/United Kingdom.
- Canadian Space Agency (CSA)/Canada.
- Centre National d'Etudes Spatiales (CNES)/France.
- Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)/Germany.
- European Space Agency (ESA)/Europe.
- Federal Space Agency (Roskosmos)/Russian Federation.
- Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
- Japan Aerospace Exploration Agency (JAXA)/Japan.
- National Aeronautics and Space Administration (NASA)/USA.

Agências Observadoras

- Austrian Space Agency (ASA)/Austria.
- Belgian Federal Science Policy Office (BFSPO)/Belgium.
- Central Research Institute of Machine Building (TsNIIMash)/Russian Federation.
- Centro Tecnico Aeroespacial (CTA)/Brazil.
- Chinese Academy of Space Technology (CAST)/China.
- Commonwealth Scientific and Industrial Research Organization (CSIRO)/Australia.
- Danish Space Research Institute (DSRI)/Denmark.
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
- European Telecommunications Satellite Organization (EUTELSAT)/Europe.
- Hellenic National Space Committee (HNSC)/Greece.
- Indian Space Research Organization (ISRO)/India.
- Institute of Space Research (IKI)/Russian Federation.
- KFKI Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
- Korea Aerospace Research Institute (KARI)/Korea.
- MIKOMTEK: CSIR (CSIR)/Republic of South Africa.
- Ministry of Communications (MOC)/Israel.
- National Institute of Information and Communications Technology (NICT)/Japan.
- National Oceanic & Atmospheric Administration (NOAA)/USA.
- National Space Organization (NSPO)/Taipei.
- Space and Upper Atmosphere Research Commission (SUPARCO)/Pakistan.
- Swedish Space Corporation (SSC)/Sweden.
- United States Geological Survey (USGS)/USA.

CONTROLE DO DOCUMENTO

Documento	Título	Data	Estado
CCSDS 120.0-G-1	Compressão de Dados Livre de Perdas, 1ª Edição	Maio 1997	Edição Original, substituída
CCSDS 120.0-G-2	Compressão de Dados Livre de Perdas, Relatório de Informação, 2ª Edição	Dezembro 2006	Edição atual: - Realiza correções menores e atualiza material obsoleto (da 1ª Edição)

CONTEÚDOS

Seção **Página**

1 INTRODUCÃO 1-1
1.1 PROPÓSITO 1-1
1.2 ESCOPO 1-1
1.3 REFERÊNCIAS 1-1

2 VISÃO GERAL 2-1

3 ALGORÍTMO 3-1
3.1 CODIFICADOR POR ENTROPIA 3-3
3.2 O PRÉ-PROCESSADOR 3-6

4 QUESTÕES DE SISTEMAS 4-1

4.1 CARACTERÍSTICAS DE SENSORES 4-1
4.2 PROPAGAÇÃO DE ERRO 4-2
4.3 EMPACOTAMENTO E DESEMPENHO DE COMPRESSÃO DE DADOS 4-2
4.4 REQUISITO DE ARMAZENAMENTO TEMPORÁRIO 4-3

5 IMPLEMENTAÇÃO 5-1
5.1 SEGMENTO ESPACIAL 5-1
5.2 SEGMENTO SOLO 5-2
5.3 DECODIFICAÇÃO 5-3

6 DESEMPENHO 6-1
6.1 MEDIDA 6-1
6.2 EXEMPLOS 6-2

ANEXO A VERIFICAÇÃO DA IMPLEMENTAÇÃO DO CODIFICADOR DE COMPRESSÃO DE DADOS LIVRE DE PERDAS DO CCSDS A-1

ANEXO B ACRÔNIMOS E ABREVIATURAS B-1

Figura

2-1 Sistema de Dados de Telemetria por Pacotes 2-1
3-1 A Arquitetura do Codificador 3-2
3-2 Curva de Desempenho para Vários Valores de *k* 3-4

CONTEÚDOS (CONTINUAÇÃO)

Figura **Página**

3-3 Curva de Desempenho Efetivo....3-4
3-4 Exemplo de Blocos Zero....3-6
3-5 Diagrama de Bloco de Pré-Processador e Pós-Processador Preditivo com Atraso Unitário 3-7
3-6 Preditor Bi-Dimensional de uma Imagem onde x (i,j) é a Amostra a ser Predita 3-8
3-7a Posição Relativa de Amostra de Valor x_i e de Valor Preditor $\hat{x}_i$ 3-10
3-7b Função de Distribuição de Probabilidade (pdf) Típica de Δ_i para Dados de Imagem....3-10
4-1 Esquema de Varredura 'Push-Broom' 4-1
4-2 Probabilidade de Excesso de Capacidade de Memória Temporária ('Buffer Overflow'). 4-4
5-1 O Segmento Espacial 5-1
5-2 O Segmento Solo 5-2
5-3 Diagrama de Blocos do Decodificador 5-3
6-1 Imagem Landsat - 4....6-2
6-2a Um Cubo de Dados do HSI....6-3
6-2b Um Plano Espectral-Espacial VNIR 6-3
6-2c Um Plano Espacial VNIR 6-3
6-3 Imagem de Radiômetro de Mapeamento de Capacidade Calorífica 6-4
6-4 Imagem da Câmera Planetária de Campo Largo....6-5
6-5 Imagem do Telescópio Solar de Raios-X Macios ('Soft') 6-7
6-6 Espectrômetro de Alta Resolução de Goddard 6-8
6-7 Forma-de-Onda de Espectrômetro Acústico-Óptico....6-9
6-8 Forma-de-Onda do espaço GRS 6-10
6-9 Resultados de Compressão com Dados GRS....6-11

Tabela

3-1 Palavras-Código da Seqüência Fundamental de Blocos-Zero em Função do Número de Blocos-Zero Consecutivos 3-5
5-1 Lógica de Decodificação para a Opção de Segunda-Extensão 5-5
6-1 Sumário de Desempenho de Compressão de Dados para Diferentes Conjuntos de Dados Científicos....6-2
6-2 Relação-de-Compressão (CR) para Diferentes Valores de Limiar 6-6
6-3 Relação-de-Compressão (CR) versus Tamanho de Pacote....6-6
6-4 Relação-de-Compressão (CR) para o GHRS 6-8
6-5 Resultado de Compressão do SWAS 6-9

1 INTRODUÇÃO

1.1 PROPÓSITO

Este relatório apresenta um sumário dos conceitos e justificativas operacionais que servem de base para os quesitos da Recomendação CCSDS, *Compressão de Dados Livre de Perdas* (referência [1]). Informações e ilustrações associadas que demonstram o desempenho, são também apresentadas. Este relatório apresenta um extenso tutorial que dá uma visão geral do algoritmo de Compressão de Dados Livre de Perdas do CCSDS, e tem por objetivo ajudar os leitores iniciantes, no entendimento da Recomendação.

1.2 ESCOPO

Este documento provê apenas matéria de apoio e descritiva: **ele não é parte da Recomendação.** No caso de algum conflito entre a Recomendação *Compressão de Dados Livre de Perdas* e a matéria aqui apresentada, a Recomendação deve prevalecer.

1.3 REFERÊNCIAS

[1] *Lossless Data Compression*. Recommendation for Space Data System Standards, CCSDS 121.0-B-1. Blue Book. Issue 1. Washington, D.C.: CCSDS, May 1997.

[2] *Procedures Manual for the Consultative Committee for Space Data Systems*. CCSDS A00.0-Y-9. Yellow Book. Issue 9. Washington, D.C.: CCSDS, November 2003.

[3] Space Packet Protocol. Recommendation for Space Data System Standards, CCSDS 133.0-B-1. Blue Book. Issue 1. Washington, D.C.: CCSDS, September 2003.

[4] AOS Space Data Link Protocol. Recommendation for Space Data System Standards, CCSDS 732.0-B-2. Blue Book. Issue 2. Washington, D.C.: CCSDS, July 2006.

[5] Pen-Shu Yeh and Warner H. Miller. *Application Guide for Universal Source Coding*. NASA Technical Paper 3441. Coding Tutorial. Washington, D.C.: CCSDS, December 1993.

[6] Robert F. Rice, Pen-Shu Yeh, and Warner H. Miller. "Algorithms for High Speed Universal Noiseless Coding." *Proceedings of the AIAA Computing in Aerospace 9 Conference*, San Diego, CA, October 19-21, 1993.

[7] TM Synchronization and Channel Coding. Recommendation for Space Data Systems Standards, CCSDS 101.0-B-3. Blue Book. Issue 1. Washington, D.C.: CCSDS, September 2003.

[8] Pen-Shu Yeh, Robert F. Rice, and Warner H. Miller. *On the Optimality of Code Options for a Universal Noiseless Coder*. NASA/JPL Publication 91-2. February 1991.

[9] Pen-Shu Yeh, Warner H. Miller, and Steve Hou. *Overview of NASA's Lossless Compression Technology Development and Application.* Milcom '95. San Diego. Nov. 1995.

[10] Masud Mansuripur. *Introduction to Information Theory.* Prentice-Hall. Englewood Cliffs, N.J. 1987.

2 VISÃO GERAL

A Compressão de dados livre de perdas tem sido sugerida para muitas aplicações em missões exploratórias em ciência espacial, seja para aumentar o retorno de conteúdo científico ou para reduzir o quesito de memória de bordo, o tempo de contato com a estação de solo e o volume de dados arquivados. Uma técnica de compressão Livre de Perdas garante a reconstrução completa dos dados originais sem incorrer em qualquer distorção no processo.

A técnica recomendada de Compressão de Dados Livre de Perdas preserva a integridade dos dados fonte ao remover a redundância dos dados fonte de aplicação. Nos processos de descompressão o dado fonte original é reconstruído a partir do dado comprimido pela restituição da redundância que foi dele removido, pelo processo de compressão. O dado reconstruído é uma réplica exata do dado fonte original. A quantidade de redundância removida do dado fonte é variável e altamente dependente da estatística dos dados fonte que, frequentemente, não é estacionária.

O algoritmo de Compressão de Dados Livre de Perdas pode ser executado (localmente) sobre o dado fonte de aplicação ou, alternativamente, executado com o uso funcional do sistema de dados de bordo (veja figura 2-1). O desempenho do algoritmo de compressão de dados independe de onde ele é aplicado. No entanto, se o algoritmo de compressão de dados é parte do sistema de dados de bordo, em geral, ele deve obter e armazenar os dados. Em ambos os casos, um re-ordenamento apropriado no sequenciamento dos dados pode se fazer necessário antes da execução do algoritmo de compressão. A re-ordenação dos dados tem o propósito de melhorar a razão de compressão.

Após a realização da compressão, a estrutura de dados de comprimento variável resultante é então transformada em pacotes tal como especificado pela referência [3]. Os pacotes contendo dados comprimidos devem, então, ser transportados através de um enlace de comunicação espaço-solo, a partir da fonte existente a bordo de um veículo espacial, para um destinatário do dados em solo, com o uso de um sistema de pacotes de dados tal como o representado na figura 2-1.

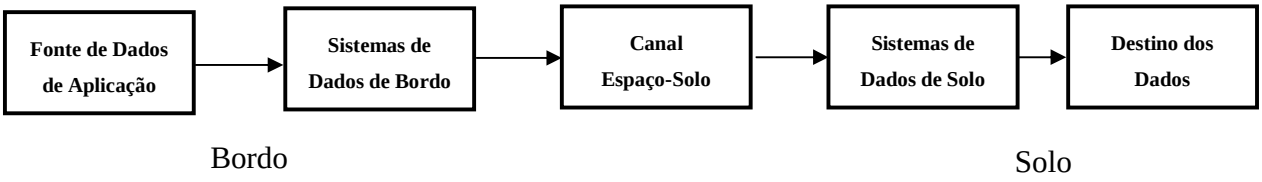


Figure 2-1: Sistema de Dados de Telemetria por Pacotes

O propósito essencial do sistema de telemetria por pacotes é o de permitir que os processos de aplicação de bordo possam criar unidades de dados que são denominadas pacotes. Estes pacotes são então transmitidos através do canal de comunicação espaço-solo de uma forma que o sistema de solo possa recuperar os dados com alta confiabilidade. Os conteúdos dos pacotes são então extraídos e suas unidades de dados são entregues ao destinatário dos dados em solo, em uma seqüência a ser descomprimida.

Este documento apresenta uma descrição de alto nível da arquitetura do algoritmo de Rice, que forma a base para a técnica de compressão (veja referência [6]). Ele, então, aborda questões sistêmicas de importância para implementações de bordo. A parte final do documento fornece exemplos de casos-de-estudo obtidos a partir de um amplo espectro de instrumentos científicos (veja referência [5]). A referência [9] também fornece resultados de Compressão de Dados Livre de Perdas obtidos a partir de três algoritmos diferentes, aplicados a cinco conjuntos de dados.

3 ALGORITMO

A implementação de uma técnica de compressão de dados em uma nave espacial trará benefício, a partir das seguintes considerações:

a) a habilidade do algoritmo em se adaptar às mudanças na estatística dos dados para maximizar o desempenho;

b) os requisitos para minimização do número de passos de processamento, redução de quesito de memória e de uso de potência;

c) a habilidade do algoritmo operar em uma única etapa de execução;

d) a habilidade do algoritmo em poder ter interface com um sistema de dados por pacotes sem necessidade de requerer, a priori, informações para o re-estabelecimento da estatística de dados para cada pacote;

e) a habilidade para limitar os efeitos de um erro de canal.

Existem poucas técnicas de compressão de dados livre de perdas bem conhecidas, o que inclui a codificação de Huffman, a codificação aritmética, a codificação de Ziv-Lempel, e as variantes de cada uma delas. Após estudo extensivo dos critérios adotados por cada uma das técnicas citadas e, comparando o desempenho de cada uma delas em um conjunto de imagens de teste, a Recomendação CCSDS selecionou o algoritmo de Rice (veja referência [9]).

O algoritmo de Rice explora um conjunto de códigos de comprimento variável para obter a compressão. Cada código é próximo do ótimo para uma fonte específica com distribuição geométrica. Códigos de comprimento variável, como os códigos de Huffman e os códigos utilizados pelo algorítmo de Rice, comprimem dados através da designação de palavras-codificadas mais curtas para os símbolos que se espera que possam ocorrer com maior freqüência, tal como ilustrado na seção 3.1.1. Com o uso de diversos códigos diferentes e com a transmissão de identificador de código, o algorítmo de Rice pode se adaptar a muitos tipos de dados fonte, desde aqueles com baixa entropia (mais compressível) até aqueles com alta entropia (menos compressível). Pelo fato de os blocos de amostras de dados fonte serem codificados de forma independente, informações complementares não precisam ser carregadas dentro dos pacotes, de tal forma que o desempenho do algorítmo é independente do tamanho do pacote. Para a discussão do termo entropia, o leitor deverá consultar um livro texto em teoria da informação, tal como o da referência [10].

Um diagrama de bloco do algoritmo de Rice é mostrado na Figura 3-1. Ele consiste de um pré-processador utilizado para decorrelacionar amostras de dados para, subseqüentemente, mapeá-las em símbolos adequados com o estágio da codificação por entropia. Os dados de entrada $\boldsymbol{x}$ para o pré-processador, representam um bloco-J de amostras de n-bits:

$\boldsymbol{x} = x_1, x_2, \ldots x_J$.

O pré-processador transforma os dados de entrada em blocos de amostras pré-processadas, $\boldsymbol{\delta}$, onde:

$\boldsymbol{\delta} = \delta_1, \delta_2, \ldots \delta_J$.

O Codificador Adaptativo converte amostras pré-processadas, $\boldsymbol{\delta}$, em uma seqüência **y** de bits codificados.

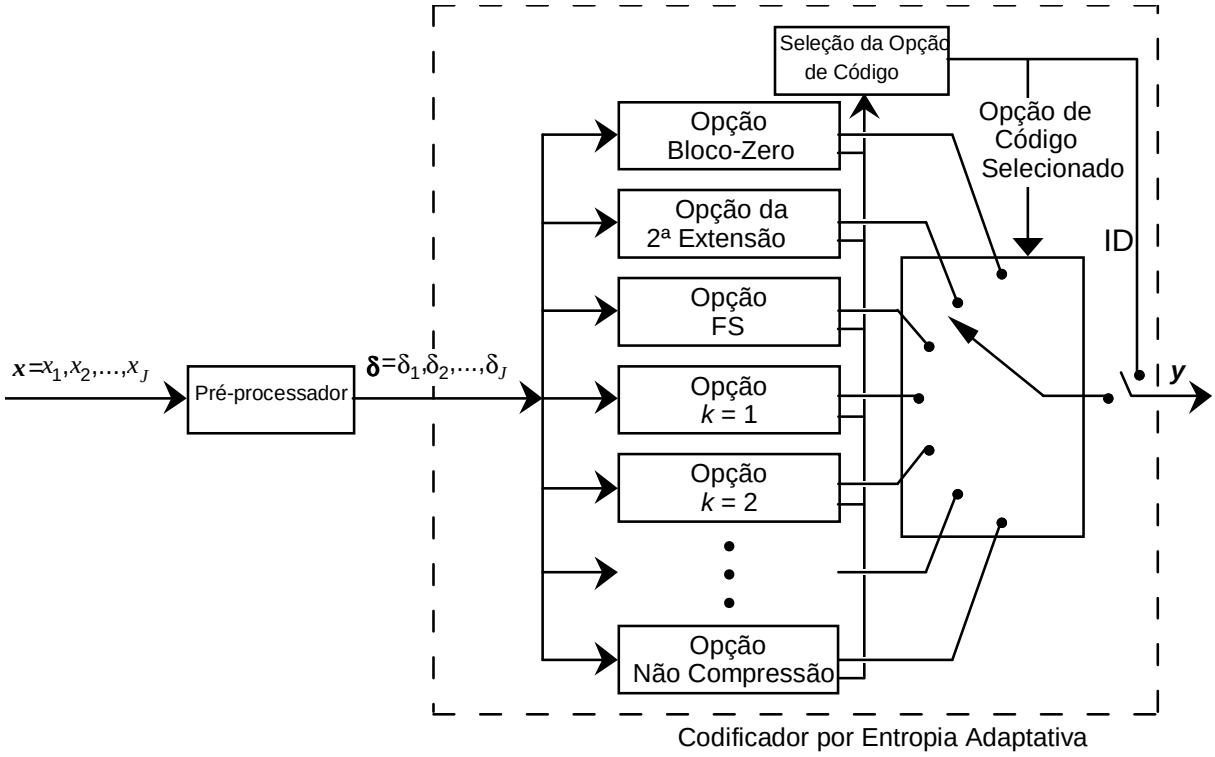


Figura 3-1: A arquitetura do codificador.

O módulo de codificação por entropia é uma coleção de códigos de tamanho variável, operando em paralelo sobre blocos de *J* amostra pré-processadas. Para a transmissão é selecionada a opção de codificação que permite obter a mais alta compressão, juntamente com um padrão de bits ID (de identificação) que é utilizado para identificar a opção adotada no decodificador. Pelo fato de uma nova opção de compressão poder ser selecionada para cada bloco, o algorítmo de Rice pode se adaptar às mudanças estatísticas dos dados fonte. Embora a Recomendação especifique que o parâmetro *J* defina 8 ou 16 amostras por bloco, o valor preferido é 16. O valor preferido de 16 amostras por bloco é o resultado de experimentos executados com várias classes de dados científicos, tanto com dados de imageamento, como com dados diversos. Estes estudos permitiram monitorar a taxa de compressão obtida em função do parâmetro *J*, que foi considerado com os valores 8, 16, 32, e 64 de amostras/bloco para as várias classes de dados científicos. Valores de *J* menores que 16 resultam em uma alta porcentagem de informação adicional, o que resulta em taxas de compressão mais baixas, enquanto que valores de *J* maiores

do que 16 produzem menor quantidade de informação adicional, mas resultam em menor capacidade do codificador em se adaptar às variações estatísticas dos dados fonte.

3.1 CODIFICADOR POR ENTROPIA

3.1.1 CODIFICAÇÃO DA SEQÜÊNCIA FUNDAMENTAL

Para ver como um código de comprimento variável pode realizar a compressão de dados, considere dois métodos de codificação de seqüências a partir de um alfabeto de quatro símbolos:

Símbolo	Código 1	Código 2 (Código SF)
s_1	00	1
s_2	01	01
s_3	10	001
s_4	11	0001

Note que o Código 2 é o código da Seqüência Fundamental (SF).

Sob qualquer um dos dois códigos, a seqüência de bits codificados que corresponde a seqüências de símbolos que são decodificáveis de forma unívoca. Para o Código 1, entende-se que a cada código de dois bits corresponde um símbolo. Para o código SF reconhece-se que cada digito '1' sinaliza o final de uma palavra código, e que o número de zeros que o precede identifica qual é o símbolo que foi transmitido. Este procedimento simples de decodificação permite que as palavras de código SF possam ser decodificadas sem o uso de tabelas de consulta ('lookup tables').

A razão pela qual o código SF permite resultar em compressão é que quando o símbolo s_1 ocorre com grande freqüência e os símbolos s_3 e s_4 ocorrem raramente, na média, iremos transmitir menos de dois bits codificados por símbolo, enquanto que Código 1 sempre requer de dois bits codificados para representar um símbolo. Códigos SF de comprimento de palavra mais longa também permitem a obtenção de compressão de uma maneira similar.

3.1.2 AS OPÇÕES COM AMOSTRA PARTICIONADA

A maioria das opções no codificador por entropia são chamadas de 'opções com amostra particionada'. A *k-ésima* opção de amostra particionada é realizada a partir de um bloco de *J* amostras pré-processadas, que retira os k bits menos significativos de cada amostra e codifica os bits restantes de ordem mais alta com uma simples palavra de código SF, antes de anexar os bits particionados ao conjunto de dados com codificação SF. Cada opção de amostra particionada no algorítmo de Rice representa uma designação para que possam ser obtidos dados comprimidos com intervalo de 1 bit/amostra (aproximadamente k + 1.5 à k + 2.5 bits/amostra); a opção de código que permite produzir a menor quantidade de bits será escolhida para cada bloco pela lógica de seleção-de-opção. Este processo de seleção por opção assegura que o mesmo bloco será codificado com a melhor opção de código disponível para o mesmo bloco de dados em questão, mas isso não implica necessariamente que a entropia dos dados fonte observe este intervalo. O valor real de entropia dos dados fonte poderia ser menor; a estatística dos dados fonte e a

efetividade do estágio de pré-processamento determina quão precisamente a entropia pode ser levada em conta, no processo.

O comprimento teórico esperado para a palavra de código com o uso do esquema de codificação por partição de amostra é apresentado na figura 3-2 para vários valores de k, onde k=0 é a opção de seqüência fundamental (SF). Essas curvas são obtidas sob a hipótese que a fonte de dados obedece, aproximadamente, uma distribuição geométrica discreta. Com este modelo para os dados fonte, traçando apenas porções das curvas próximas da diagonal que representa o desempenho ideal, para qualquer caso de entropia de dados, teoricamente, resulta em uma curva de desempenho com projeção teórica, tal como é apresentada na figura 3-3. Para um sistema prático, tal como o que é apresentado na figura 3-1, os bits de ID com 3 ou 4 bits de comprimento serão necessários para todos os blocos de *J* amostras, para os casos de sistemas com opção-8 ou opção-16. Em um sistema deste tipo, a fonte tem a probabilidade de apresentar um desvio da distribuição geométrica discreta, no entanto, a opção que permite o menor número de bits codificados é aquela que é selecionada. Os resultados práticos da codificação com imagens aéreas, juntamente com a sua aplicabilidade para outros tipos de fonte de dados, com funções de distribuição de Gauss ou de Poisson, são fornecidos na referencia [8].

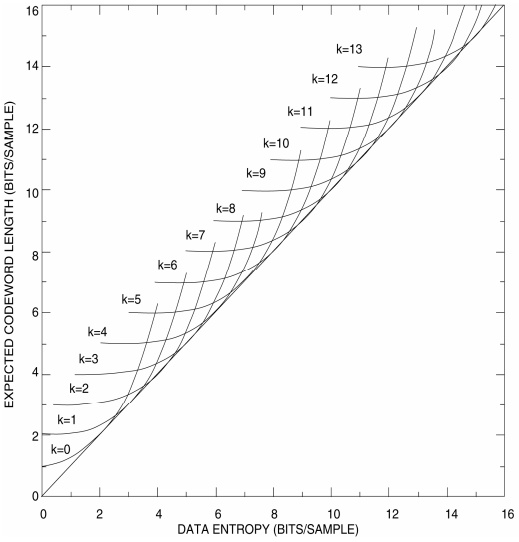


Figura 3-2: Curva de desempenho para vários valores de *k*.

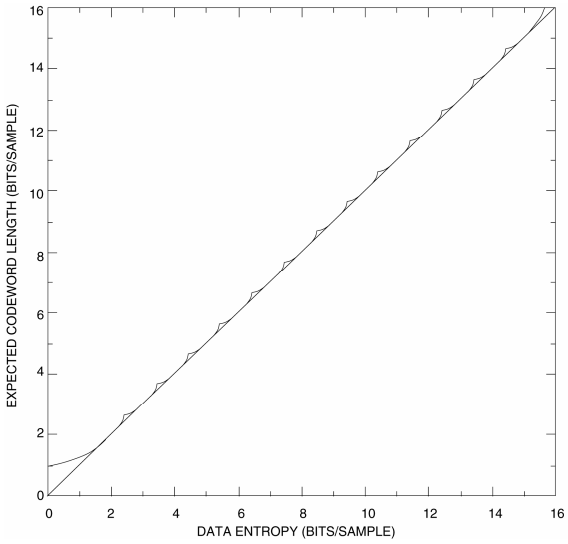


Figura 3-3: Curva de desempenho efetivo.

3.1.3 OPÇÕES DE BAIXA ENTROPIA

Para fornecer uma taxa de compressão menor que 1.5 bits/amostra, o algorítmo inclui duas opções de baixa entropia. As opções de baixa entropia são particularmente eficientes quando as amostras pré-processadas apresentam apenas valores pequenos. Estas opções de baixa entropia são as opções de Segunda Extensão e de Bloco-Zero.

3.1.3.1 Opção de Segunda Extensão

A opção de Segunda Extensão é projetada para produzir taxa de compressão entre 0.5 a 1.5

bits/amostra. Quando esta opção é selecionada, o esquema de codificação cria inicialmente pares de amostras pré-processadas consecutivas do bloco de J amostras e, então, transforma as duplas de amostras em um novo valor que é então codificado com a Seqüência Fundamental FS. As palavras código FS são transmitidas para γ, tal que:

$$\gamma = (\delta_i + \delta_{i+1})\,(\delta_i + \delta_{i+1}+1)/2 + \delta_{i+1}$$

3.1.3.2 Opção de Bloco-Zero

A opção de Bloco-Zero é selecionada quando um ou mais blocos de amostras pré-processadas são todos zeros. Neste caso, o numero de blocos de zeros adjacentes pré-processados são codificados por uma palavra código SF, tal como é apresentado na tabela 3-1.

Tabela 3-1: Palavras-Código da Seqüência Fundamental de Blocos-Zero em Função do Número de Blocos-Zero Consecutivos.

Número de Blocos-Zero	Palavra-Código SF
1	1
2	01
3	001
4	0001
ROS	00001
5	000001
6	0000001
7	00000001
8	000000001
.	.
.	.
.	.
63	0000 . . . 0000000001 (63 0s e um 1)

Os *r* blocos entre as amostras de referencia consecutivas são divididos em segmentos, como descrito na seção 3.4.3.2 da Recomendação "Compressão de Dados Livre de Perdas" (referência [1]). Cada segmento, exceto possivelmente o último, contém *s*=64 blocos. O código Resto-de-Segmento (ROS) é utilizado para indicar que o restante de um segmento consiste de cinco ou mais blocos-zero. Mesmo quando as amostras de referência não são utilizadas, os parâmetros *r* e *s* ainda são utilizados para segmentar os dados para a opção de Bloco-Zero.

Por exemplo, se existem 200 blocos entre as amostras consecutivas de referência, os primeiros três segmentos possuem 64 blocos e o quarto segmento possui 8 blocos. Durante a calibração dos sensores, todos os 200 blocos devem possuir o mesmo valor. Após aplicar o preditor de atraso unitário, todos os 200 blocos correspondentes às amostras pré-processadas terão valores iguais a zero, com exceção da primeira amostra no primeiro bloco, que é a amostra de referência. A relação entre os segmentos não codificados e os segmentos codificados está representada na figura 3-4. A condição ROS está presente no final de cada segmento de 64 blocos, e um novo ID

é atribuído ao inicio de cada segmento. Deve ser notado que a condição ROS também codifica os oito últimos blocos-todo-zero.

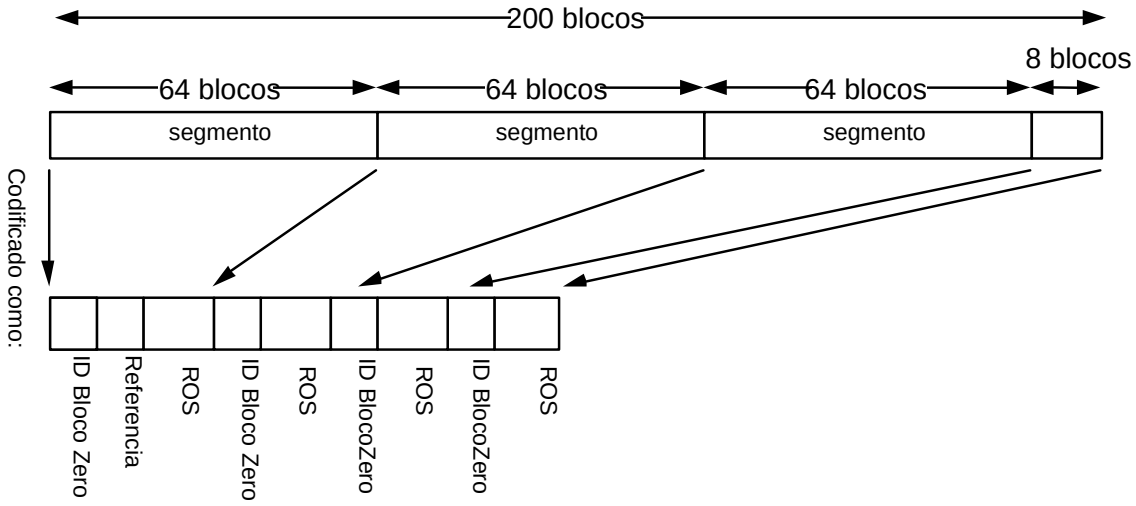


Figura 3-4: Exemplo de Blocos Zero.

3.1.4 OPÇÃO DE NÃO-COMPRESSÃO

Quando nenhuma das opções acima resulta em qualquer compressão de dados em um bloco, a opção de não-compressão é selecionada. Nesta opção, o bloco de dados pré-processados é transmitido sem qualquer alteração, exceto pelo uso de um prefixo de identificação que se faz necessário.

3.1.5 SELEÇÃO DA OPÇÃO

O codificador por entropia seleciona a opção que requer o menor numero de bits necessários para codificar os símbolos de um bloco corrente. Uma seqüência de bits de identificação especifica qual é a opção que foi selecionada.

3.2 O PRÉ-PROCESSADOR

3.2.1 O PREDITOR

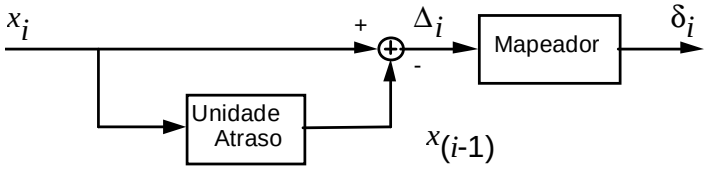
O papel do pré-processador é o de transformar os dados em amostras que possam ser eficientemente comprimidos pelo codificador por entropia. Para garantir que a compressão possa ser Livre de Perdas, o pré-processamento que a executa deve ser reversível. Podemos ver pelo exemplo da seção 3.1.1 que para se obter uma compressão efetiva, é necessário dispor de um estagio de pré-processamento que transforme o dado original de tal forma que palavras código mais curtas ocorram com maior probabilidade do que as palavras mais longas. No caso de

existirem vários estágios de pré-processamento a serem considerados, o melhor seria selecionar aquele que produz, em média, a menor palavra código.

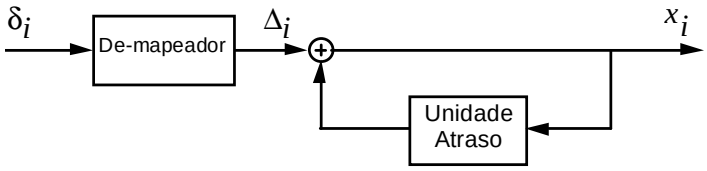
Em geral, um pré-processador que remove a correlação entre amostras no bloco de dados de entrada deverá melhorar o desempenho do codificador por entropia. Na discussão apresentada a seguir, é assumido que o pré-processamento é realizado por um preditor que é sucedido por um mapeador de erros da predição. Para alguns tipos de dados, técnicas mais sofisticadas baseadas em transformadas podem oferecer melhoria de eficiência na compressão, com o custo de se tornarem mais complexos.

A função de descorrelação do pré-processador pode ser implementada por uma escolha criteriosa do preditor para um esperado conjunto de amostras de dados. A seleção de um preditor deveria levar em consideração o tipo de dado que é esperado, assim como também, as possíveis variações do ruído de fundo e o ganho dos sensores que realizam a aquisição de dados. O preditor deveria ser escolhido com o propósito de minimizar a quantidade de ruído resultante da não-uniformidade do sensor.

Um dos tipos de codificador de predição mais simples é o preditor linear de primeira ordem de atraso unitário que é apresentado na figura 3-5 a. A saída, Δ_i, será a diferença entre o símbolo do dado de entrada e o símbolo do dado que o precede.



a) Pré-processador



b) Pós-processador

Figura 3-5: Diagrama de Bloco de Pré-processador e Pós-processador Preditivo com Atraso Unitário.

Existem muitos outros preditores. Em geral, os usuários podem selecionar qualquer esquema de predição que possa ser capaz de melhor descorrelacionar o fluxo de dados de entrada. Por exemplo, pode-se considerar uma imagem com valores $x(i, j)$ onde i representa a linha de varredura e j o numero do 'pixel' na linha de varredura (veja figura 3-6). Os tipos possíveis de preditores incluem:

a) Preditor unidimensional de primeira ordem:

O valor de intensidade preditivo, $\hat{x}(i, j)$ poderia ser igual às amostras anteriores, na mesma linha de varredura, $x(i, j–1)$, ou ao valor da amostra vizinha de uma linha de varredura anterior, $x(i–1, j)$. O preditor de atraso unitário é um exemplo de um preditor unidimensional de primeira ordem.

b) Preditor bidimensional de segunda ordem:

O valor preditivo de intensidade $\hat{x}(i, j)$, poderia resultar da média das amostras adjacentes, $x(i, j–1)$ e $x(i–1, j)$.

c) Preditor bidimensional de terceira ordem:

O valor da intensidade prevista $\hat{x}(i,j)$ poderia resultar da combinação ponderada dos três valores vizinhos de $x(i, j–1)$, $x(i–1, j)$, e $x(i–1, j–1)$.

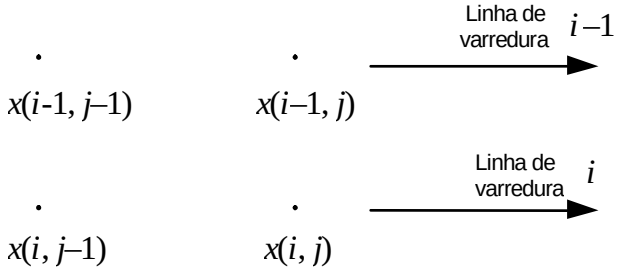


Figura 3-6: Preditor Bidimensional de uma Imagem onde $x(i, j)$ é a Amostra a ser Predita.

Para dados multiespectrais, outra técnica de predição consiste em utilizar amostras de uma banda espectral como entrada para uma função de predição de ordem superior para a banda seguinte.

3.2.2 AMOSTRA DE REFERÊNCIA

Uma amostra de referência é uma amostra de dado inalterado, sobre a qual são realizadas predições sucessivas. Amostras de referência são requeridas pelo decodificador para ele poder recuperar os valores originais a partir de valores diferenciais. Nos casos em que uma amostra de referência não é utilizada no pré-processador ou o pré-processador está ausente, ela não deve ser considerada para inserção. O usuário deve determinar com que freqüência as referências devem ser inseridas. Quando necessária, a referência deve ser a primeira amostra de um bloco de J amostras de entrada. Nos formatos de pacote, a amostra de referência deve estar presente no primeiro Conjunto de Dados Codificados (‘Coded Data Set’ – CDS), no campo de dados do pacote, como é definido no item 5.1.4 da referência [1], e deve ser repetida a cada r de blocos de amostras de dados.

3.2.3 O MAPEADOR

Baseado no valor preditivo $\hat{x}_i$, o mapeador de erro preditivo converte cada valor de predição de erro, Δ_i em um número inteiro não negativo de n-bits, δ_i, apropriado para ser processado pelo codificador por entropia. Para uma compressão mais eficiente pelo estágio do codificador por entropia, os símbolos pré-processados, δ_i, devem satisfazer:

$$p_0 \geq p_1 \geq p_2 \geq \ldots p_j \geq \ldots p_{(2^n-1)}$$

onde p_j é a probabilidade de δ_i ser igual ao inteiro j. O que garante que os símbolos mais prováveis possam ser codificados com palavras de código mais curtas.

O exemplo a seguir ilustra a operação do mapeador de erro de predição após um preditor de atraso unitário ter sido ser aplicado às amostras de dados de 8 bits, com valores entre 0 à 255:

Valor da Amostra x_i	Valor Previsto $\hat{x}_i$	Δ_i	θ_i	δ_i
101	—	—	—	—
101	101	0	101	0
100	101	–1	101	1
101	100	1	100	2
99	101	–2	101	3
101	99	2	99	4
223	101	122	101	223
100	223	–123	32	155

Se x_{min} e x_{max} indicam os valores mínimo e máximo de qualquer amostra de entrada x_i então, claramente, qualquer predição razoável de valor $\hat{x}_i$ deve estar no intervalo [x_{min}, x_{max}]. Conseqüentemente, o erro da predição Δ_i deve ser um dos 2^n valores no intervalo [$x_{min} - \hat{x}_i$, $x_{max} - \hat{x}_i$], como ilustrado na figura 3-7a. É de se esperar que, com uma boa escolha de preditor, possam ocorrer com maior probabilidade valores menores do que maiores que $|\Delta_i|$, tal como é mostrado na figura 3-7b. Conseqüentemente, a função de mapeamento do erro preditivo é dada por:

$$\delta_i = \begin{cases} 2\Delta_i & \text{, com } 0 \leq \Delta_i \leq \theta \\ 2|\Delta_i| - 1 & \text{, com } -\theta \leq \Delta_i \leq 0 \\ \theta + |\Delta_i| & \text{, nos demais casos} \end{cases}$$

onde θ_i = mínimo ($\hat{x}_i - x_{min}$, $x_{max} - \hat{x}_i$),

e possui a propriedade de apresentar a relação de probabilidades $p_i < p_j$, sempre que $|\Delta_i| > |\Delta_j|$.Esta

propriedade aumenta a possibilidade de satisfazer a ordenação de cada probabilidade p_i (descrita acima).

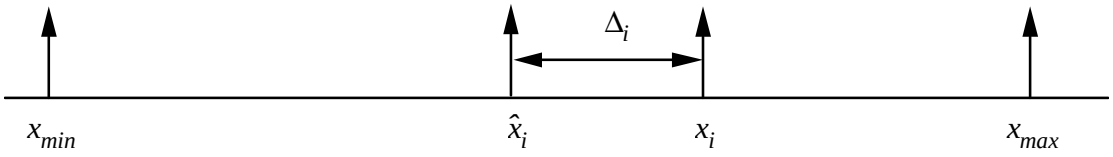


Figura 3-7a: Posição Relativa de Amostra de Valor x_i e de Valor Preditor $\hat{x}_i$

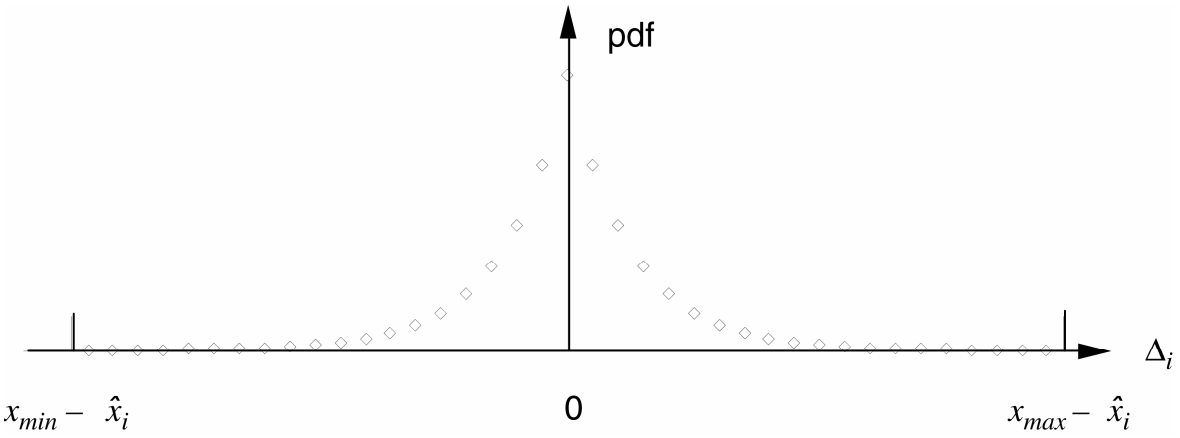


Figura 3-7b: Função de Distribuição de Probabilidade (pdf) Típica de Δ_i para Dados de Imagem.

4 QUESTÕES DE SISTEMAS

Diversas questões sobre sistemas que estão relacionados com a inclusão de esquemas de compressão de dados a bordo de uma nave espacial devem ser consideradas. Tais questões incluem: a relação entre: o arranjo plano focal dos conjuntos de sensores, a abordagem utilizada para a amostragem de dados, a predição da direção, o numero de amostras por pacote, o esquema subseqüente de empacotamento de dados e como ele se relaciona com a propagação de erro quando da ocorrência de um bit de erro no canal de comunicação e, a quantidade de armazenamento temporário de dados que se faz necessária quando os dados comprimidos são repassados para um enlace que opera a uma taxa constante de comunicação (de dados).

4.1 CARACTERÍSTICAS DE SENSORES

Instrumentos avançados de imageamento e espectrômetros freqüentemente utilizam conjuntos de detectores dispostos em uma configuração uni ou bidirecional; um exemplo é o do Dispositivo de Carga Acoplada, (CCD). Este conjunto de detectores individuais tende a ter ligeiras diferenças entre eles quando respondem à mesma intensidade de fótons, na entrada. Por exemplo, os CCDs apresentam geralmente um valor diferente de ganho e de corrente de escuro (‘dark current’), para cada detector. É importante ter os sensores calibrados de tal modo que seus dados refletem os sinais efetivamente recebidos por cada um deles. A investigação da resolução de um sensor CCD com dados de 8-bits, com variação de ganho de 7 por cento e com o nível de corrente de escuro menor do que meio bit, demonstra uma variação de entropia de 2,8 bits, o que torna o esquema de predição ineficaz para a compressão de dados (isto é, para uma taxa de compressão de 2,6 sobre 1,4). Para minimizar o efeito da variação do ganho do detector no ganho da compressão em um esquema de predição unidimensional, o pré-processador (figura 3-5) será muito mais eficaz quando os dados adquiridos por um elemento do detector são utilizados para predição de dados adquiridos pelo mesmo detector, com características estacionárias. Um exemplo é aquele do esquema ‘push-broom’, usado em muitos sistemas de satélite de visada terrestre, e que está representado na figura 4-1. Neste caso, um conjunto de detectores é disposto na direção perpendicular àquela da trilha (‘cross-track direction’) orbital e que também é a direção considerada para a leitura dos dados do detector. A varredura da cena terrestre é obtida pelo movimento do satélite ao longo da direção de sua trilha (‘track’) orbital. Em uma configuração deste tipo, a codificação preditiva será mais eficaz ao longo da direção da trilha orbital da espaçonave, situação em que os dados do mesmo detector são utilizados para predição.

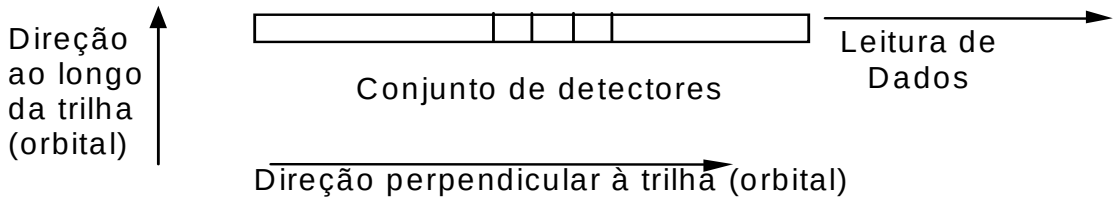


Figura 4-1: Esquema de Varredura ‘Push-Broom’

4.2 PROPAGAÇÃO DE ERRO

Uma preocupação maior na utilização de compressão de dados é aquela devida à possibilidade de propagação de erro quando da ocorrência de erro em um bit, o que pode provocar a propagação de um erro de reconstrução em pontos de dados sucessivos àquele onde se deu o erro original. Este efeito pode ser reduzido com o uso de empacotamento de dados, com o propósito de limitar a extensão da propagação deste tipo de erro e, também, com o uso de codificação de canal, tal como a codificação de Reed-Solomon, que tem o efeito de reduzir e limitar a probabilidade de erro de canal (veja a referência [7]). Um exemplo de aplicação espacial poderia ser o do empacotamento dos dados comprimidos de uma linha de varredura em um formato de pacote recomendado pelo CCSDS (referências [3]) e os dados comprimidos da linha de varredura seguinte em um segundo pacote (e assim por diante). No caso de ocorrência de um de bit de erro em um pacote, o erro de descompressão se propagaria somente da posição onde o evento de erro ocorreu, até o final do pacote.

4.3 EMPACOTAMENTO E DESEMPENHO DE COMPRESSÃO DE DADOS

O número total dos bits codificados que resulta da compressão livre de perdas para um número fixo de amostras de dados, é variável. A arquitetura de dados CCSDS fornece uma estrutura que permite empacotar seqüências de dados de comprimento variável que, por sua vez, podem ser colocadas em quadros de comprimento fixo que para serem transportados do espaço à terra. O processo de empacotamento é essencial para a prevenção da propagação de erro, tal como foi mencionado acima; ele também pode afetar o desempenho da compressão para o conjunto de dados. Para maximizar o desempenho de compressão, o empacotamento deveria ser executado sobre diversos blocos para poder assim minimizar a fração de bits utilizados com a ocupação do cabeçalho. No entanto, para limitar a propagação de erro, um número razoável de blocos de dados deveria ser escolhido para compor cada pacote.

Existem outros algoritmos de compressão disponíveis cujo desempenho depende do estabelecimento de estatísticas de longo prazo, obtidas a partir da coleção de dados do sensor. Tais esquemas, em geral, apresentam um bom desempenho de compressão para uma grande quantidade de dados e, mau desempenho para quantidades de dados relativamente pequenas. E, se o empacotamento é utilizado como meio de prevenção para a propagação de erro em conjunção com estes algoritmos, deve se esperar um desempenho pobre para um pacote mais curto e um desempenho melhor para um pacote maior. O inconveniente é que a perda dos dados devido à propagação de erro pode ser intolerável para pacotes maiores.

A compressão de dados de livre de perdas não é recomendada para os sistemas com pacotes de comprimento fixo porque o desempenho de compressão obtido é menor do que aquele obtido com um sistema de pacotes de comprimento variável. No caso dos projetos para uso em bordo que requerem pacotes de comprimento fixo, os dados comprimidos devem ser seguidos por bits de enchimento (‘fill bits’) que possam garantir que o pacote tenha o seu comprimento fixo. A quantidade de bits de enchimento adicionados devem assegurar que o excesso (´overflow´) de dados de um pacote se torne altamente improvável. Se a contagem de bits comprimidos for maior do que aquela permitida pelo comprimento fixo de um pacote, ocorre o truncamento com perda

dos dados. O decodificador reconhece a condição de truncamento e sinaliza a ocorrência de truncamento com a geração de símbolos falsos. O decodificador, então, gera tantos símbolos falsos quantos forem necessários para preencher um pacote típico, como se ele não tivesse sido truncado. Se um ou mais blocos forem truncados, blocos correspondentes contendo símbolos falsos serão gerados para o preenchimento correto do número de blocos necessários para completar cada pacote. Para reduzir os efeitos do truncamento quando estruturas de pacotes de comprimento fixo são utilizadas, o projeto pode ser realizado de tal forma que a partição ('split') de k-bits menos significativos (LSBs) possam ser os últimos a serem transmitidos. Em tal arranjo, se ocorrer o truncamento, provavelmente somente os bits menos significativos (LSBs) venham a ser perdidos e, desta forma, o impacto no desempenho seria amenizado. Entretanto, (neste caso) a nível de pacote, a compressão não pode mais ser considerada como sendo livre de perdas.

4.4 REQUISITO DE ARMAZENAMENTO TEMPORÁRIO

Dependendo do requisito da missão, os pacotes de comprimento variável resultantes do empacotamento de uma seqüência de bits comprimidos em pacotes CCSDS podem ser armazenados em uma memória de bordo de alta capacidade, ou eles podem ser multiplexados com outros pacotes de dados, antes de serem armazenados no dispositivo de memória. Em qualquer um dos casos, o dispositivo de memória de alta capacidade diminui o efeito da variação no comprimento do pacote. A leitura subseqüente desta memória é realizada a uma taxa fixa. Em outras situações, os pacotes de comprimento variável podem ser temporariamente armazenados antes que seja realizada a transmissão direta para o enlace de comunicação. O armazenamento temporário serve de "buffer" para suavizar a comunicação para o enlace. Ocasionalmente, bits de enchimento são introduzidos na seqüência de dados para proporcionar uma taxa constante de leitura dos dados. A quantidade de armazenamento que é necessária é uma função: da taxa de pacotes de entrada, da estatística do comprimento dos pacotes, e da taxa de leitura.

Uma análise foi realizada, usando (a Recomendação) o *Protocolo CCSDS de Enlace de Dados Espaciais* AOS (referência [4]), para exame de questões, tais como, a do requisito de tamanho máximo de memória temporária de suavização ('smoothing buffer') quando uma Unidade de Dados de Protocolo de Multiplexação (MPDU) de taxa variável é transmitida para o solo com taxa fixa dentro de uma Unidade de Dados de Canal Virtual (VCDU). Nesta análise, a MPDU transportou pacotes de comprimentos variáveis contendo dados comprimidos, com comprimento variado, segundo a distribuição Gaussiana. A Figura 4-2 mostra os resultados desta análise, onde a probabilidade de excesso da capacidade de armazenamento temporário ('buffer overflow') diminui, na medida em que a capacidade de armazenamento temporário aumenta. A duplicação da capacidade de armazenamento temporário normalmente diminui em diversas ordens de grandeza, a probabilidade de ocorrer excesso de capacidade. As MPDUs são deslocadas assincronamente no VCDU sempre que estes são preenchidos e a entrada da MPDU no VCDU ocorre dentro de um período de tempo de geração de pacote igual a *tp;* enquanto que *ts* representa o período de amostragem do VCDU que, por sua vez, provê uma taxa constante de dados para o canal físico.

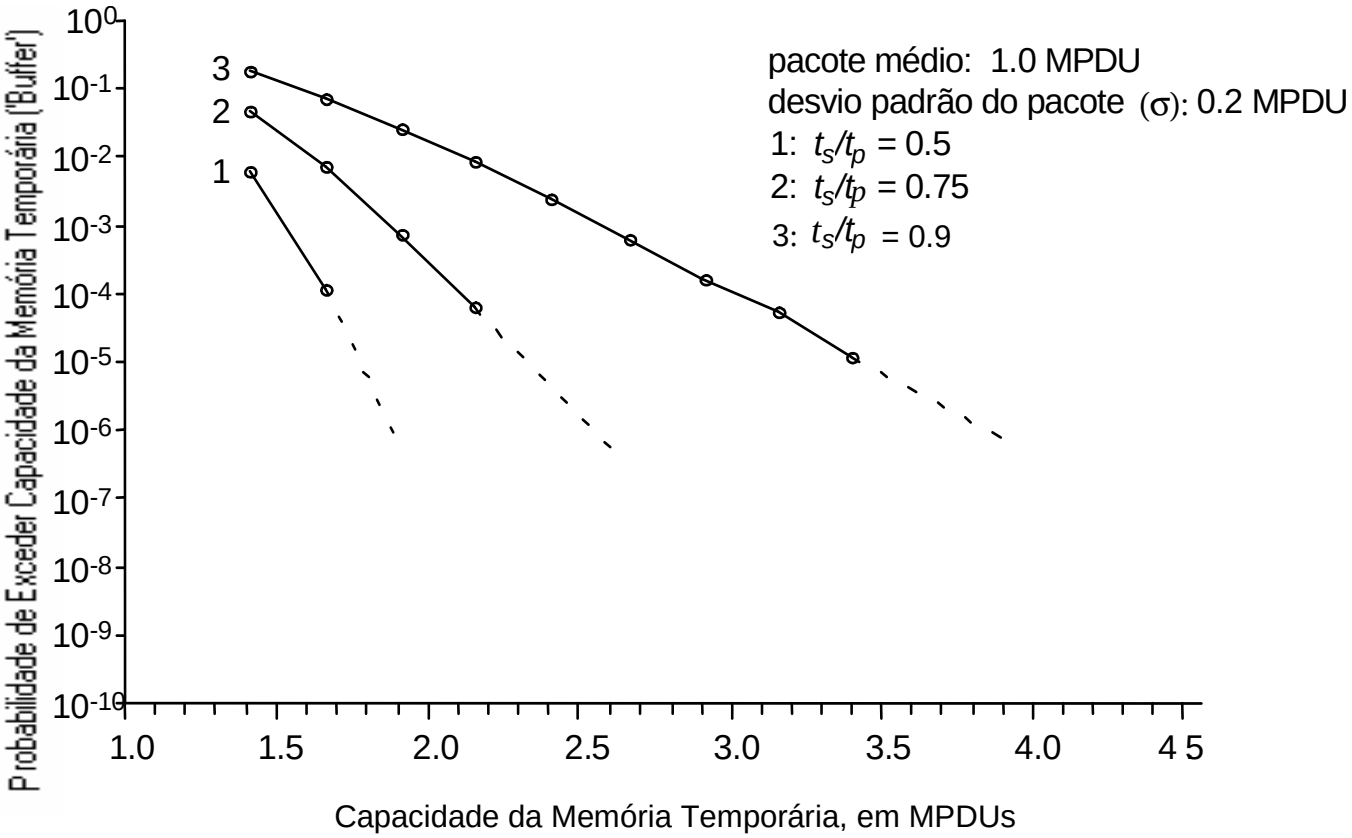


Figura 4-2: Probabilidade de Excesso de Capacidade de Memória Temporária ('Buffer Overflow')

5 IMPLEMENTAÇÃO

Os exemplos apresentados nesta seção estão baseados em um sistema que utiliza o *Protocolo CCSDS de Enlace Espacial de Dados AOS* (referência [4]) em conjunção com *Protocolo CCSDS de Pacotes Espaciais* (referência [3]).

5.1 SEGMENTO ESPACIAL

A Figura 5-1 representa o Segmento Espacial na forma de um diagrama de blocos funcional que identifica cada uma das operações relacionadas com os dados, na medida em que eles são transportados através da arquitetura de camadas CCSDS de bordo. A descrição seguinte assume que os dados de instrumentação de bordo a serem comprimidos resultam de um imageador multiespectral.

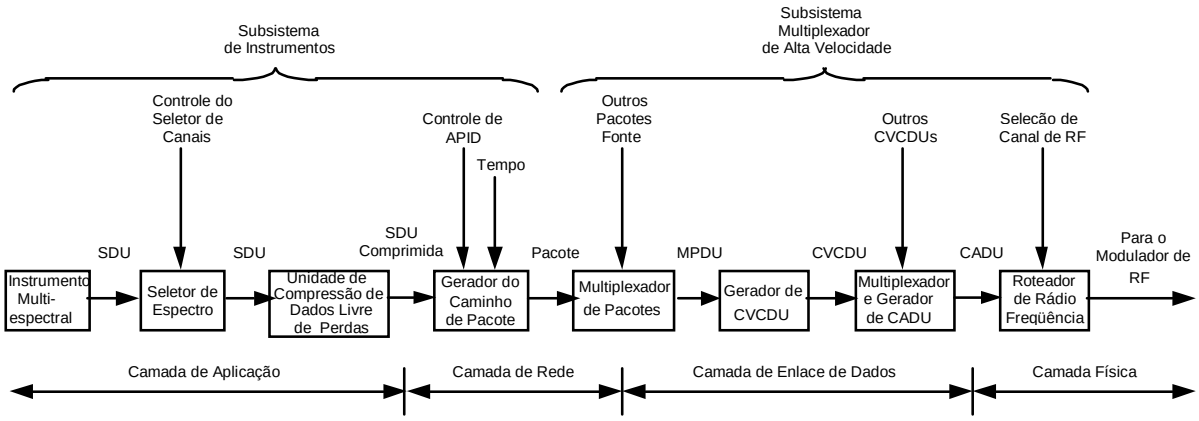


Figura 5-1: O Segmento Espacial

Na camada de aplicação, canais espectrais são selecionados para transporte de imagens destinadas a seus usuários. O seletor fornece Unidades de Dados de Serviço (‘Service Data Units’ - SDUs), que são obtidas a partir de linhas de varredura, na medida em que elas são geradas. Uma SDU é gerada para cada canal espectral individual (i.e., as linhas de varredura, dos canais espectrais não são multiplexados em uma única SDU). Outras funções da camada de aplicação devem realizar a operação de compressão de dados no SDU e identificar o comprimento da SDU, relacionando-a a um número inteiro de octetos (superior) mais próximo, antes de repassá-la para o gerador de pacotes que vem a seguir.

Na camada rede, as SDUs são etiquetadas com Identificadores do Processo de Aplicação (‘Application Process Identifiers’, APIDs) apropriados para a geração de pacotes com o seu comprimento assinalado no cabeçalho do pacote. Se for necessário colocar a etiqueta de tempo, a sinalização de cabeçalho secundário é ativada para que ela possa ser inserida no campo do cabeçalho secundário. Na camada de enlace, os pacotes de dados de imagem são multiplexados com pacotes provenientes de outros instrumentos, para formar um MPDU de tamanho fixo. O cabeçalho da MPDU contém o apontamento para o cabeçalho do primeiro pacote, necessário para de multiplexação. Estas MPDUs são então atribuídas a um canal virtual específico, no qual as

Unidades de Dados de Canal Virtual Codificado (‘Coded Virtual Channel Data Units’, CVCDUs) com comprimento fixo de octetos inteiros são então geradas. As CVCDUs contém o contador de seqüência e os octetos de paridade (do codificador) Reed-Solomon. Estas CVCDUs são então agrupadas aleatoriamente para que possa ser garantida a transição de bits e repassadas ao Multiplexador e ao gerador de Unidades de Dados de Canal de Acesso (‘Channel Access Data Unit’, CADU). Neste ponto, as CADUs de outros instrumentos de dados são conjuntamente multiplexadas e são separadas em canais selecionados para roteamento aos usuários.

5.2 SEGMENTO SOLO

5.2.1 DIAGRAMA DE BLOCOS DO SEGMENTO DE SOLO

A Figura 5-2 apresenta o diagrama de blocos funcional do segmento de solo do usuário.

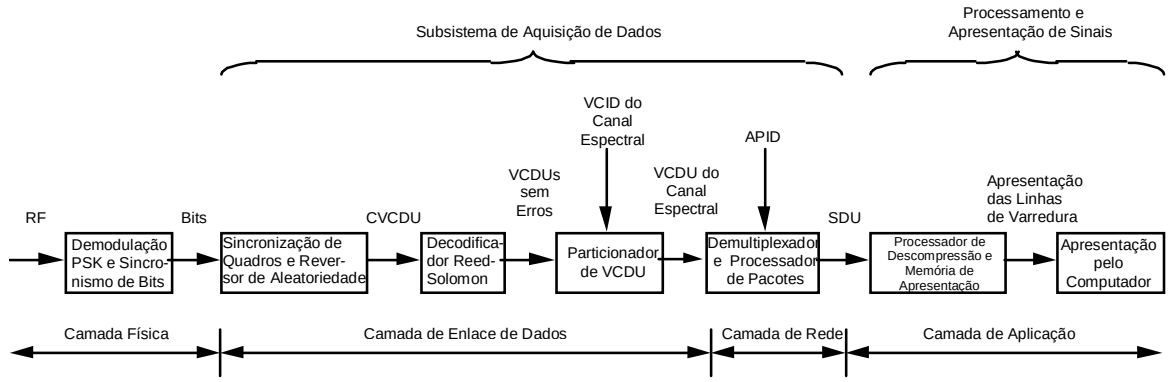


Figure 5-2: Segmento Solo

A camada física detecta o sinal de Rádio Freqüência (RF) do sinal, demodula o sinal de Chaveamento por Deslocamento de Fase (‘Phase Shift Keying’ - PSK) e sincroniza os bits do sinal com ruído. Os bits de Não-Retorno-a-Zero (‘Non-Return to Zero’ - NRZ) detectados e o sinal de sincronismo de relógio são entregues ao Subsistema de Aquisição de Dados (‘Data Acquisition Subsystem’).

A camada de enlace primeiramente recebe a seqüência ou cordão de bits, e então, utilizando o conhecimento do padrão de sincronização de quadro e comprimento de quadro, detecta as CADUs e desfaz a aleatoriedade das CVCDUs. As CVCDUs são posteriormente desembaralhadas e o esquema Reed-Solomon é decodificado para correção de erros de canal. O decodificador Reed-Solomon fornece então as VCDUs virtualmente ‘livres-de-erro’, as quais são então repassadas para o particionador de VCDUs. O particionador de VCDU recebe uma seqüência de VCDUs e delas recupera as VCDUs da imagem espectral. Ele também verifica a contagem de seqüência e passa esta informação para o processador de pacotes.

O processador e demultiplexador de pacotes identifica a localização do cabeçalho do primeiro pacote na MPDU e executa o processamento do pacote. Pacotes administrativos e de telemetria

são os primeiros a serem decodificados, para determinação da configuração e do estado (‘status’) do Subsistema de Instrumentação de bordo. Estas mensagens administrativas permitem que o Subsistema de Aquisição de Solo (‘Ground Acquisition Subsystem’) possa se configurar para poder separar de forma apropriada os dados de entrada e neles poder reconhecer quais são os canais espectrais que estão presentes nos dados. A partir do cabeçalho do pacote são determinados: o tamanho do pacote, o APID, e o estado (‘status’) do cabeçalho secundário. Se o APID identifica o pacote como sendo uma unidade de imagem espectral comprimida, a SDU é então extraída e repassada para o processador de descompressão. Se existir um campo de cabeçalho secundário, ele também é extraído e utilizado para determinar o tempo da geração de imagem. O Processador de Descompressão (‘Decompression Processor’) e a Memória de Apresentação (‘Display Memory’) recebem as imagens espectrais comprimidas ou SDUs, realizam a descompressão das mesmas em tempo real para, em seguida, inserir as linhas de imagem descomprimidas na memória temporária de apresentação.

5.3 DECODIFICAÇÃO

O decodificador Livre de Perdas consiste de duas partes funcionais separadas, o pós-processador e o decodificador de entropia adaptativa, como representado na figura 5-3. O pós-processador realiza ambas as funções de: operação de predição inversa (‘inverse prediction operation’) e o inverso da operação de mapeamento normalizada (veja figura 3-5).

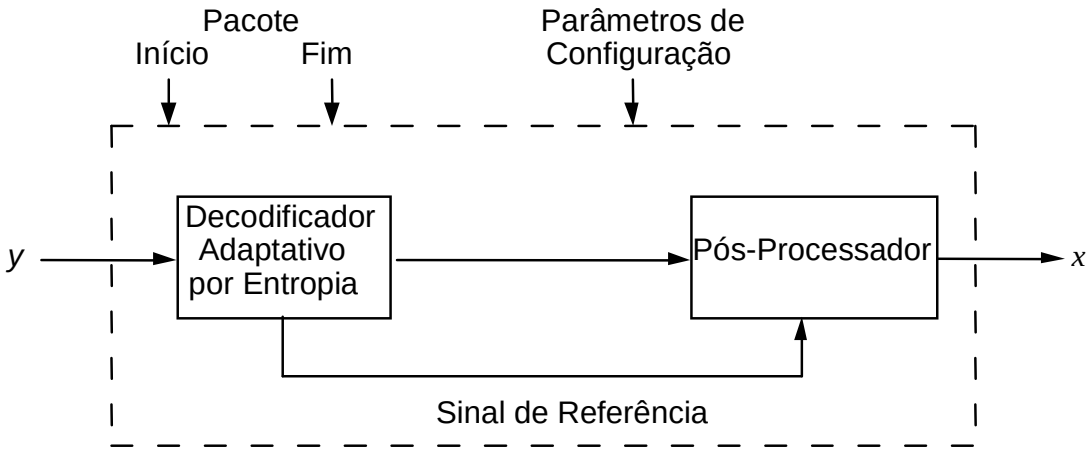


Figura 5-3: Diagrama de Blocos do Decodificador

O primeiro passo para a decodificação é o da definição dos parâmetros de configuração, de tal forma que ambos, o codificador e o decodificador, possam operar no mesmo modo. Estes parâmetros de configuração são específicos para cada missão e fixados para um determinado APID e, ou são conhecidos a priori ou são indicados pelo Pacote de Identificação de Compressão (‘Compression Identification Packet’ - CIP).

Os parâmetros de configuração comuns a ambos, ao codificador e ao decodificador, são os

seguintes:

- resolução em número de bits por amostra (n = 1,..., 32);
- tamanho de bloco, em número de amostras (J = 8 ou 16);
- número de Conjuntos de Dados Codificados (CDSes) contidos em um pacote (l = 1,..., 4096);
- número de CDSes entre referências (r = 1,..., 256);
- tipo de preditor, se utilizado;
- sentido do valor do sinal digital (positivo ou bipolar).

As entradas para o decodificador são: a) o campo de dados do pacote fonte contendo os dados fonte *y*, comprimidos; b) os sinais de início e de fim do campo de dados do pacote fonte, e; c) os parâmetros de configuração do decodificador.

O Decodificador de Entropia Adaptativa utiliza os parâmetros de configuração do decodificador para processar os dados comprimidos de comprimento variável contidos dentro do campo de dados do pacote fonte.

Os bits relacionados para código opcional de ID, que podem ser encontrados no início do CDS, devem ser extraídos em primeiro lugar. Se uma amostra de referência tiver sido inserida no CDS durante a codificação, ela sempre aparecerá no primeiro CDS de um pacote e deverá reaparecer depois de r CDSes, se o comprimento do pacote for maior do que r. Em seguida, a palavra de referência é extraída da seqüência (cordão) de bits do CDS e repassada ao pós-processador. Dependendo dos bits de ID, o decodificador deverá interpretar os bits restantes do CDS como sendo bits codificados por uma das opções especificadas na referência [1].

Para a opção de Não-Compressão, os bits restantes do CDS são repassados diretamente ao pós-processador. Para todas as outras opções, as palavras-de-código FS devem ser decodificadas, de acordo como está descrito na Seção 3.1.1. Para as opções de amostra separadas (‘Split Sample’), cada uma das J palavras-de-código FS devem ser concatenadas com k bits da separação, extraídos da seqüência que utiliza o formato de dados CDS, tal como descrito na Seção 5.1.4, da referência [1].

Para a opção de Segunda-Extensão, somente $J/2$ palavras-de-código FS devem ser extraídas. Em seguida, elas devem ser decodificadas com a execução dos seguintes passos:

Passo 1. obter m contando os 0s da palavra-de-código FS;
Passo 2. obter β e m_s , usando a lógica da tabela 5-1;
Passo 3. calcular $\delta_{i+1} = m - m_s$;
Passo 4. calcular $\delta_i = \beta - \delta_{i+1}$.

Tabela 5-1: Lógica de Decodificação para a Opção de Segunda-Extensão

m	β	m_s
0	0	0
1,2	1	1
3,4,5	2	3
6,...,9	3	6
10,....,14	4	10
15,....,20	5	15
21,....,27	6	21
28,....,35	7	28

Para a opção de Bloco-Zero, a única palavra-de-código FS em seguida ao ID e à referência, indica o número de blocos-de-zero ou a condição ROS, tal como consta na lista da Tabela 3-1. Usando esta tabela, um número apropriado de blocos-de-zero, são gerados como entrada para o pós-processador.

O pós-processador executa a função inversa de mapeamento, dado o valor preditivo de $\hat{x}_i$ Quando o pré-processador não é utilizado durante a codificação, o pós-processador também não é levado em consideração. A função inversa de mapeamento pode ser expressa da seguinte forma:

se $\delta_i \leq 2\theta_i$,

$$\Delta_i = \begin{cases} \delta_i/2 & \text{quando } \delta_i \text{ é par} \\ -(\delta_i + 1)/2 & \text{quando } \delta_i \text{ é ímpar} \end{cases}$$

se $\delta_i > 2\theta_i$,

$$\Delta_i = \begin{cases} \delta_i - \theta & \text{quando } \theta_i = \hat{x}_i - x_{min} \\ \theta_i - \delta_i & \text{quando } \theta_i = x_{max} - \hat{x}_i \end{cases}$$

onde θ_i = minimo $(\hat{x}_i - x_{min}, x_{max} - \hat{x}_i)$

Devem ocorrer J (ou J–1 valores, no caso de um CDS contendo uma amostra de referência)

valores de erro de predição: Δ_1, Δ_2,....Δ_J , gerados a partir de um CDS. Estes valores deverão ser utilizados no pós-processador de atraso unitário (figura 3-5), com o propósito de recuperar os J valores de amostra: x_1, x_2,.....x_J.

Para se executar a decodificação de pacotes que podem incluir bits de enchimento, duas parcelas de informação devem ser entregues ao decodificador. Primeiro, o decodificador deve conhecer o tamanho do bloco, J, assim como o número de CDSes existentes em um pacote, l. Segundo, o decodificador precisa saber quando o Campo de Dados do Pacote Fonte (SPDF) começa e quando ele termina. A partir destas duas parcelas de informação, o número de bits de enchimento ('fill bits'), existentes no fim do campo de dados pode ser determinado.

O número de amostras válidas existentes no último CDS do SPDF, pode ser menor do que J. Se assim for, o usuário obterá as amostras de enchimento para o ultimo CDS para poder compor um bloco completo com J amostras, de tal forma que o usuário precisa extrair as amostras válidas, a partir do último bloco que, então, conterá J amostras.

6 DESEMPENHO

6.1 MEDIDA

Existem duas medidas de desempenho de Compressão de Dados de Livre de Perdas: a taxa de dados codificados, medida em bits por amostra, e o tempo de processamento. A última não será tratada neste documento, uma vez que ela se trata de uma medida altamente dependente da plataforma de processamento.

A taxa mais baixa de dados a partir da fonte pode ser representada como Livre-de-Perdas está relacionada com a entropia da fonte. Isto é, uma fonte com entropia de *H* bits por amostra exige uma média de pelo menos *H* bits para poder representar cada uma das amostras. Por exemplo, se as saídas de uma fonte discreta são independentes e identicamente distribuídas, com probabilidades p_1, p_2,....$p_{(2^n-1)}$, então a fonte possui a entropia:

$$H = -\sum_{0}^{2^n-1} p_i \log_2 p_i$$

Se as saídas das fontes não são independentes, então a entropia depende de distribuições de probabilidade de ordem mais alta. A maioria das fontes no mundo real são suficientemente complexas, de tal forma que é impossível o cômputo exato da entropia da fonte, pois o melhor limite possível de compressão que pode ser obtido de uma fonte qualquer é normalmente desconhecido.

É tentador querer medir a distribuição de probabilidade de primeira ordem de uma fonte e calcular a entropia de 'primeira ordem' com o uso da expressão apresentada acima. Porém, esta abordagem produz uma estimativa precisa somente se os símbolos de saída forem independentes entre si. Ao se aplicar várias técnicas de compressão Livre de Perdas, nós podemos estimar um limite superior para a entropia de uma fonte. Seria também possível desenvolver um modelo probabilístico da fonte, e computar a entropia do modelo.

A Relação de Compressão (CR) é a relação do número de bits por amostra antes da compressão pela taxa de dados codificados, de tal forma que o algoritmo de compressão Livre de Perdas:

$$\mathbf{CR} = \frac{nJ}{\text{comprimento médio dos CDS em}}$$

onde *n* é a resolução de amostra e *J* é o tamanho de bloco.

Normalmente para dados científicos de imageamento, o valor de CR é aproximadamente igual a dois. No pior caso, quando a opção de Não-Compressão é selecionada, a CR resultante será ligeiramente menor do que o valor um, devido aos bits excedentes ('overhead bits') de ID.

6.2 EXEMPLOS

As subseções seguintes contêm resultados de estudos-de-compressão, a maior parte dos quais foram extraídos da referência [5], para oito instrumentos científicos diferentes envolvendo imageamento e não-imageamento. O desempenho de compressão, expressa pela CR, está resumido na Tabela 6-1. Note que a CR é dependente dos tipos de dados de um instrumento. Para um instrumento específico, a CR pode variar de um conjunto de dados, para o (conjunto) seguinte a ele. O leitor deve consultar os detalhes de dados de cada instrumento após consultar cada exemplo individual fornecido pela Tabela 6-1.

Tabela 6-1: Sumário de Desempenho de Compressão de Dados para Diferentes Conjuntos de Dados Científicos

Parágrafo	Conjunto de Dados do Instrumento	Taxa de Compressão
Imageamento	a) Mapeador Temático (TM)	1.83
	b) Imageador Hiper Espectral	2.6
	c) Radiômetro de Imageamento da Capacidade Calorífica	2.19
	d) Câmera Planetária de Campo Largo	2.97
	e) Telescópio Solar de Raios X Macios	4.69
Sem Imageamento	f) Espectrômetro Goddard de Alta resolução	1.72
	g) Espectrômetro Acústico-Óptico	2.3
	h) Espectrômetro de Raios Gama	5 a 26[1]

[1] dependendo do tempo de integração

a) **Mapeador Temático (TM)**

Propósito da Missão: O programa Landsat foi iniciado para o estudo da superfície e recursos da Terra. Os satélites Landsat-1, Landsat-2, e Landsat-3 foram lançados entre 1972 e 1978. Os Landsat-4 e Landsat-5 foram lançados em 1982 e 1984, respectivamente.

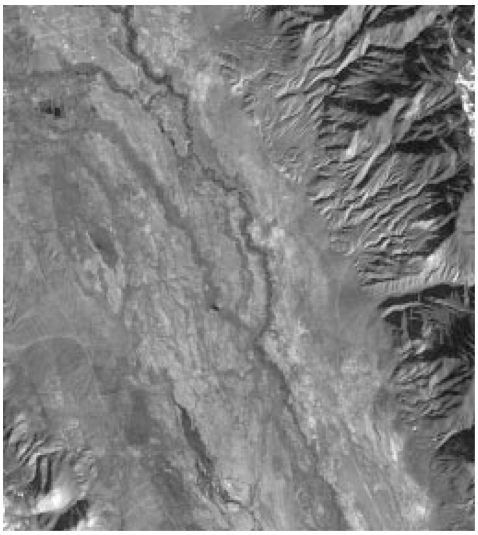


Figura 6-1: Imagem Landsat-4

Mapeador Temático Landsat: Nos satélites Landsat-4 e 5, os dados de TM representam dados típicos de observação terrestre, adquiridos com a projeção de imageamento realizado com varredura de espelho para detectores de estado sólido. O sensor das bandas 1-5 é composto por um conjunto de fotos-diodos com 16 elementos; para a banda 6, foi feito um arranjo com um conjunto de 4 elementos sensores. Uma imagem obtida com o satélite Landsat-4 com resolução de

30m no solo para a banda 1, de comprimento de ondas entre 0.45 - 0.52 μm, mostrado na Figura 6-1. Esta imagem de 8 bits, com 512 x 512 pixels foi obtida sobre a Serra Nevada, na Califórnia e possui uma quantidade de informações relativamente alta sobre a sua área montanhosa.

Estudo de Compressão: Usando um preditor de atraso unitário na direção horizontal, configurado para um bloco de 16 amostras e inserindo uma referência para cada linha de imagem, a compressão de Livre de Perdas obtida corresponde a uma CR de 1.83 para a imagem (com pixels) de 8 bits.

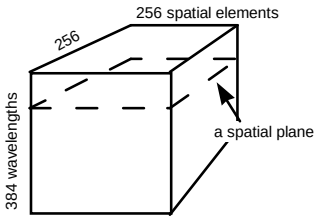


Figura 6-2a:
Um Cubo de Dados do HSI

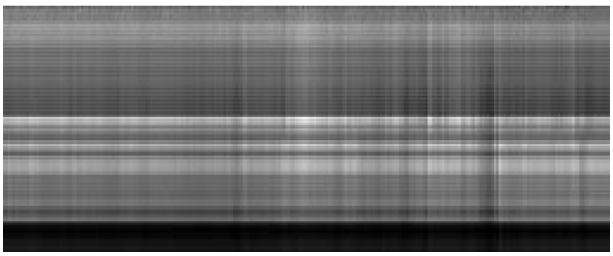


Figura 6-2b: Um Plano Espectral Espacial VNIR

b) **Imageador Hiper Espectral no Pequeno Satélite da Missão de Lewis de Infusão de Tecnologia**

Propósito de Missão: O objetivo da missão de Infusão de Tecnologia com um Pequeno Satélite (SSTI) representa a transição para missões futuras com novas tecnologias que apresentam benefícios significantemente mensuráveis de desempenho. A missão será (foi) lançada em 1997.

O Imageador Hiper Espectral (HSI): O instrumento contém três conjuntos de detectores CCD. Dois deles são utilizados para coletar informações espaciais de espectro na faixa de comprimento de onda de 0.4 até 2.5 μm, com: o imageador de Espectro Visível no Infravermelho-Próximo (VNIR) e o imageador Infra-Vermelho-de-Ondas-Curtas (SWIR). Estes imageadores fornecem uma linha espacial de 256 pixels, com informações em 128 faixas espectrais do detector VNIR e em 256 faixas espectrais do detector SWIR. Ambas as saídas do VNIR e SWIR são quantizadas com dados de 12 bits, com resolução de solo de 30m. O terceiro CCD é um imageador pancromático que fornece resolução de 5m de solo, com saída de dados de 8 bits. Este CCD é composto por um arranjo linear de 2592 elementos.A compressão de dados será aplicada somente sobre os cubos de dados dos detectores VNIR e SWIR, caracterizados por conjuntos de 256 x 256 pixels de informações espaciais abrangendo 384 informações espectrais (veja Figura 6-2a).

Figura 6-2c: Um Plano Espacial VNIR

Estudo de Compressão: Até que os dados de instrumentos possam ser coletados após o lançamento, a simulação é realizada com um cubo de dados simulados a partir dos dados do Espectrômetro Embarcado de Imageamento no Infra-Vermelho Visível (AVIRIS). A compressão é aplicada sobre os dados do VNIR ou do SWIR, separadamente, mas da mesma forma, isto é, ao longo da direção espacial, em um plano espacial-

espectral, tal como é mostrado na Figura 6-2b. Uma relação de compressão média de 2.6 foi obtida com os dados do VNIR e do SWIR. Se a predição multispectral é aplicada entre as linhas espectrais, a CR aumenta para 3.44. A Figura 6.2c mostra um plano espacial que corta o cubo de dados.

c) **Mapeamento Radiométrico de Capacidade Calorífica**

Propósito da Missão: Lançada em 1978, a missão teve o propósito de dar suporte à investigações científicas exploratórias para estabelecer a viabilidade de se utilizar medidas de temperatura da superfície da Terra em intervalos de 12 horas com o uso remoto de sensores de infravermelho termal, com o propósito de se obter a diferença de temperaturas entre a noite e o dia, para determinação da inércia térmica da superfície da Terra.

Radiômetro de Mapeamento de Capacidade Calorífica: O sensor é um foto-detector de estado sólido sensível às regiões do visível ou do infra-vermelho. Uma imagem típica de 8 bits por pixel, tomada na região da luz visível é apresentada na Figura 6-3, com resolução de solo de 500m sobre a área da Baía de Chesapeake, na Costa Leste dos Estados Unidos.

Estudo de Compressão:

Predição uni-dimensional: escolhendo a predição com atraso unitário na direção horizontal, e adotando o tamanho de bloco $J = 16$, a Compressão Livre de Perdas permite obter uma CR de 2.19.

Predição bi-dimensional: usando um predictor bi-dimensional que toma a média da amostra anterior e a amostra da linha anterior de varredura, e mantendo os demais parâmetros os mesmos, é possível obter uma Compressão Livre de Perdas com CR de 2.28, ou seja, cerca de 5% maior do que a que é obtida com um predictor unidimensional.

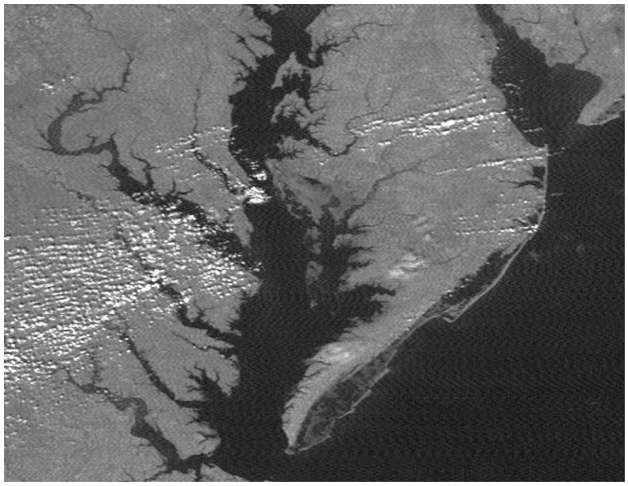


Figura 6-3: Imagem de Radiômetro de Mapeamento de Capacidade Calorífica

d) **Câmera Planetária de Campo Largo (WFPC) do Telescópio Espaçal Hubble**

Propósito da Missão: O **Telescópio Espacial Hubble** (HST), lançado em 1990, tem por objetivo adquirir dados astronômicos com uma resolução que, até então, nunca pôde ser obtida, para se poder observar objetos 100 vezes mais tênues do que aqueles que os demais observatórios (terrestres) de até então têm observado.

Câmera Planetária de Campo Largo (WFPC): A câmera utiliza quatro conjuntos de detectores de área do tipo CCD. Um típico campo estelar observado por esta câmera é mostrado na Figura 6-4. Esta imagem, em particular, tem um valor máximo de resolução de 12 bits (por pixel), incluindo o ruído de fundo de 9 bits.

Figura 6-4: Imagem da Câmera Planetária de Campo Largo

Estudo de Compressão: Dois estudos foram realizados. O primeiro, foi uma investigação do desempenho da Compressão de Dados Livre de Perdas em um campo-de-estrelas com vários níveis de limiar de ruído de fundo. O segundo, foi um estudo do efeito da relação de taxa-de-compressão quando se varia o comprimento dos dados fonte para serem comprimidos em um pacote.

Estudo de Limiar: Foi explorado neste estudo o efeito de situações de limite (limiar) em um campo de estrelas, em relação ao desempenho de compressão. Diferentes valores de limiar foram aplicados a esta imagem e os valores dos dados foram restritos ao valor de limiar ou a um valor maior do que ele. A operação de limiar não alterou a qualidade visual da imagem.Todas as estrelas brilhantes não foram afetadas. A imagem de resultante foi comprimida Livre de Perdas e a compressão fez uso da opção de baixa entropia, o que mostrou ocorrer com freqüência, na cobertura da imagem. A imagem original da Figura 6-4 tem um valor mínimo de quantização igual a 423 para um conjunto de 800 x 800 (pixels). A compressão foi realizada com o emprego de blocos de tamanho igual a 16 e com o uso de um preditor de atraso unitário na direção da linha horizontal. Os resultados estão sumarizados na Tabela 6-2.

Tabela 6-2: Relação-de-Compressão (CR) para Diferentes Valores de Limiar

Valores de Limiar	CCSDS LDC	LZW
Nenhum	2.97	3.30
430	2.99	3.33
440	11.92	17.15
450	41.20	63.53
460	53.52	91.48
470	63.66	110.73

Na Tabela 6-2, as CRs a Compressão de Dados Livre de Perdas CCSDS (LDC) para o estudo da WFPC são comparadas com as CRs obtidas usando o algoritmo de Lempel-Ziv-Welch (LZW). Embora o desempenho do algoritmo de LZW sobre a imagem inteira parece ser melhor do que aquele obtido com o algoritmo de Compressão de Dados Livre de Perdas (LDC) do CCSDS, os resultados são na realidade invertidos após considerar as questões de propagação de erro e de empacotamento, envolvidos na implementação do algoritmo Livre de Perdas.

Estudo do Empacotamento: Quando um formato de empacotamento de dados é utilizado para os dados comprimidos, é esperado que durante a descompressão o erro de reconstrução causado por um único bit de erro no pacote não irá se propagar para o próximo pacote, pois os pacotes de dados são independentes entre si e nenhuma tabela de código ou de estatística de dados é repassada de um pacote para o seguinte. Para explorar o efeito de tamanho do pacote no desempenho de compressão, várias linhas individuais da Figura 6-4 são extraídas e comprimidas, separadamente. Os resultados da Tabela 6-3 são obtidos a partir das linhas de números 101 a 104, que cobrem parcialmente o agrupamento de estrelas mais brilhantes no canto superior esquerdo da Figura 6-4.

Tabela 6-3: Relação-de-Compressão (CR) vs. Tamanho de Pacote

Limiar	Tamanho de Pacote Nº de Linhas	LDC CCSDS	LZW
Nenhum	1 (Linha 101)	2.89	1.94
Nenhum	1 (Linha 102)	2.90	1.92
Nenhum	1 (Linha 103)	2.86	1.90
Nenhum	1 (Linha 104)	2.88	1.90
Nenhum	2 (Linha 101-102)	2.90	2.13
Nenhum	2 (Linha 103-104)	2.87	2.09
Nenhum	4 (Linha 101-104)	2.89	2.29
470	1 (Linha 101)	16.24	6.98
470	1 (Linha 102)	16.92	6.98
470	1 (Linha 103)	15.42	6.94
470	1 (Linha 104)	16.24	6.67
470	2 (Linha 101-102)	16.45	8.70
470	2 (Linha 103-104)	15.80	8.60
470	4 (Linha 101-104)	16.06	10.30

e) **Telescópio de Raios-X Macios (SXT) da Missão Solar-A**

Propósito da Missão: A missão Solar-A, que foi renomeada missão Yohkoh, depois de seu lançamento bem sucedido em Agosto de 1991, foi dedicada ao estudo de "flares" solares, especialmente dos fenômenos de alta energia observados nas faixas de Raios-X e Gama.

Telescópio de Raios-X Macios (SXT): O instrumento de SXT é um telescópio de reflexão com incidência-por-toque para detecção de Raios-X nos comprimentos de onda de 3 a 60 Angstroms. Ele utiliza um conjunto de detectores de 1024 x 1024 CCDs, para cobertura completa do disco solar. Os dados adquiridos utilizados na análise de compressão de dados estão representados na imagem da Figura 6-5.

Estudo de Compressão: O preditor de atraso unitário é aplicado a uma imagem de alto contraste. Uma CR de 4.69 é obtida, o que significa que somente 3.2 bits por pixel são necessários para que se possa garantir a precisão plena de 15 bits, da imagem.

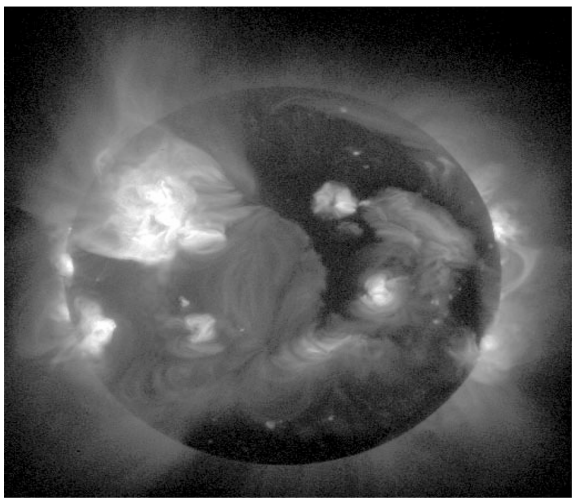


Figura 6-5: Imagem do Telescópio Solar de Raios-X Macios ('Soft')

f) **Espectrômetro de Alta Resolução de Goddard (GHRS) do Telescópio Espacial Hubble**

Propósito da Missão: O Telescópio Espacial Hubble (HST), lançado em 1990, tem por objetivo adquirir dados astronômicos com uma resolução que, até então, nunca pode ser alcançada, para se poder observar objetos 100 vezes mais tênues do que aqueles que os observatórios terrestres, até então, puderam observar.

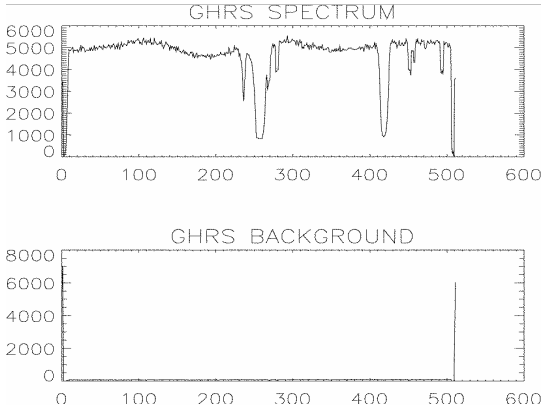


Figura 6-6: Espectrômetro de Alta Resolução de Goddard

Espectrômetro de Alta resolução de Goddard (GHRS): O objetivo científico primário do instrumento GHRS é o de investigar: o meio interestelar, ventos estelares, evolução e fontes extra-galácticas. Seus sensores são compostos por dois conjuntos otimizados de foto-diodos para permitir a cobertura em diferentes regiões de comprimentos de onda, ambos na faixa espectral de UV. A Figura 6-6 ilustra exemplos de um típico espectro de sinal de fundo. Estes sinais de fundo possuem 512 valores de dados; cada um deles é capaz de armazenar contagens digitais numa faixa dinâmica de 10^7. As formas espectrais normalmente não variam muito. Porém, variações sutís do espectro numa faixa tão larga e dinâmica representa um desafio para qualquer algorítmo de Compressão de Dados Livre de Perdas.

Estudo de Compressão: Dois esquemas diferentes de predição foram aplicados: o primeiro utiliza a amostra anterior para o mesmo sinal de fundo e o segundo utiliza o sinal de fundo anterior na mesma categoria (espectro ou fundo) como preditor. Os resultados resumidos estão apresentados na Tabela 6-4, abaixo. O conjunto de dados de teste contém dados com um valor máximo menor do que 16 bits na faixa dinâmica e, no caso, o CR foi calculado em relação a 16 bits.

Tabela 6-4: Relação de Compressão (CR) para o GHRS

Sinal de fundo("Trace")	Predictor	CR
Espectro 1	Amostra Anterior	1.64
Espectro 2	Amostra Anterior	1.63
Espectro 2	Espectro 1	1.72
Fundo 1	Amostra Anterior	2.51
Fundo 2	Amostra Anterior	2.51
Fundo 2	Sinal de Fundo 1	1.53

g) **Espectrômetro Acústico-Óptico no Satélite de Astronômico de Ondas Submilimétricas (SWAS)**

Propósito da Missão: O SWAS é uma missão do Pequeno Explorador (SMEX), programado para lançamento em 1998, a bordo de um lançador Pegasus. O objetivo do SWAS é o de estudar o balanço de energia e as condições físicas das nuvens moleculares em galáxias, através da observação do espectro de ondas de rádio emitidas especificamente por certos tipos de moléculas. Os resultados a serem obtidos serão relacionados com às teorias de formação de estrelas e de sistemas planetários. A plataforma SMEX permite a realização de uma missão científica relativamente barata, com o uso de um curto espaço de tempo de desenvolvimento. A missão SWAS é pioneira que incorpora um observatório completo de rádio-astronomia em uma pequena carga útil.

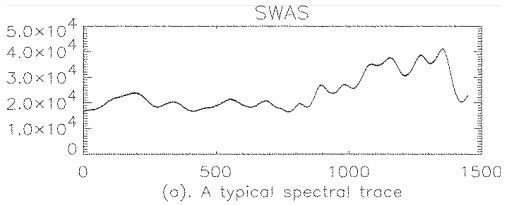


Figura 6-7: Forma de Onda do Espectrômetro Acústico-Óptico

Espectrômetro Acústico-Óptico: O Espectrômetro de Acústico-Óptico utiliza uma célula de Bragg para converter a energia de rádio-frequência de um receptor de ondas submilimetricas SWAS em uma onda acústica que, por sua vez, provoca a difração de um raio laser sobre um conjunto de sensores CCD. O sensor tem 1450 elementos, com leitura de 16 bits. Um espectro típico é apresentado na Figura 6-7a. A visão expandida de uma parte de dois sinais espectrais é apresentada na Figura 6-7b. Devido à não uniformidade do detector, a diferença no ganho de ADC entre um canal par-ímpar e os efeitos causados por variações de temperatura, fazem com que os espectros apresentem valores de "offset" (deslocamento de referência) não uniformes entre os sinais, além de variações em forma dente-de-serra entre amostras, dentro de um mesmo sinal de fundo. Por causa da disponibilidade limitada de memória de bordo, uma CR maior que 2 é exigida para esta missão.

Estudo de Compressão: A larga faixa dinâmica e as variações dentro de cada sinal de fundo representam um desafio para o algoritmo de compressão. A predição de atraso unitário entre amostras é ineficaz quando os canais ímpar e par possuem ganhos diferentes de ADC. Três modos de predição foram estudados; os resultados são apresentados na Tabela 6-5.

Tabela 6-5: Resultado de Compressão do SWAS

Preditor	CR
Amostra Anterior	1.58
Sinal de fundo Anterior	1.61
Sinal de fundo Anterior (Preditor Multi-espectral de Ordem Mais Alta)	2.32

Os resultados mostram que, mesmo que ocorrendo similaridade entre os

sinais espectrais, um preditor que utiliza um sinal de fundo adjacente de uma forma direta não conseguirá melhorar o desempenho de compressão por causa do "offset" desigual entre os sinais de fundo. O modo de operação do preditor multispectral é particularmente efetivo no caso de fontes de dados múltiplos registrados espacialmente, com "offsets" de fundo.

h) **Espectrômetro de Raios Gama do Observador de Marte**

Propósito da Missão: O Observador de Marte foi lançado em Setembro de 1992. O Observador teve o propósito de coletar dados com o uso de vários instrumentos para auxiliar os cientistas na compreensão: da superfície Marciana, das propriedades atmosféricas e das interações entre os vários elementos envolvidos.

Espectrômetro de Raio de Gama (GRS): O espectrômetro usa um detector de germânio de alta pureza para medição de raios Gama. O espectro do vôo é coletado em dezesseis mil canais, cada um deles correspondendo a uma faixa de energia de raios de gama. O faixa total de energia de um espectro se estende de 0.2 MeV até 10 MeV. Espectros típicos para intervalos de tempo de coleta de 5 a 50 segundos são apresentados na Figura 6-8. Estes espectros mostram a natureza aleatória das contagens; algumas das quais têm valor zero sobre vários bits. A faixa dinâmica de contagem espectral é de 8 bits.

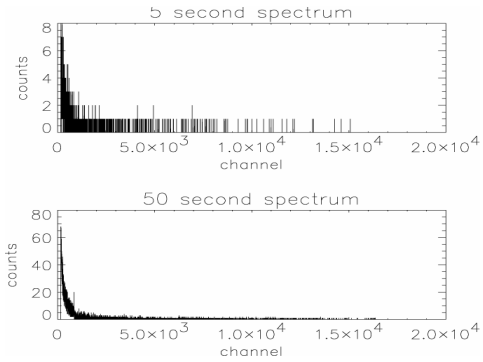


Figura 6-8: Forma-de-Onda do espaço GRS

Estudo de Compressão: Dois esquemas diferentes que utilizam o mesmo algoritmo de compressão foram simulados. Um esquema aplica o algoritmo de Compressão de Dados Livre de Perdas diretamente sobre o espectro e o outro, implementa uma estrutura de codificação em duas camadas.

A aplicação direta do algoritmo (esquema de um - passo): Na implementação em um único-passo, o tamanho de bloco *J* é definido igual a 16 e o modo codificação-por-entropia é selecionado para evitar o uso do preditor e da função de mapeamento; o algoritmo permite obter uma CR maior que 20 com compressão Livre de Perdas para espectros de 5 segundos. Na medida em que o tempo de coleta de raios Gama aumenta, a relação de compressão (CR) que pode ser obtida decresce.

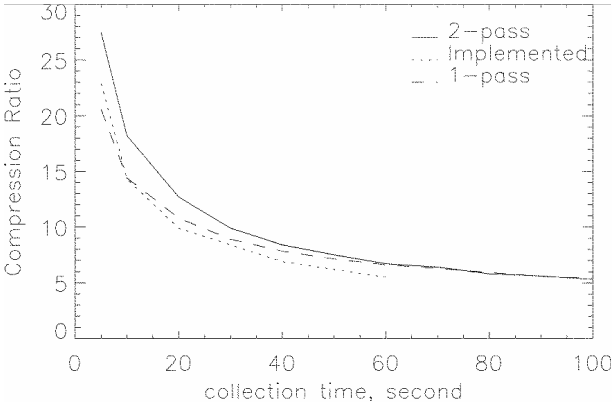


Figura 6-9: Resultados de Compressão com Dados GRS

Esquema de duas camadas: Em um esquema de duas camadas, dois passos se fazem necessários para se realizar a compressão dos dados. O primeiro passo obtém o número de canais que apresentam contagens válidas e fornece qual é o número ("run-lenght") de canais existentes, entre eles. Enquanto isso, um arquivo de contagem é criado para anotar somente as contagens consideradas válidas para dados a serem arquivados. No segundo passo, ambos os arquivos: contendo dados de espaçamento ("run-length") entre canais válidos e entre aqueles contendo as respectivas contagens são comprimidos usando o algoritmo de Compressão Livre de Perdas com um tamanho de bloco de 16 que utiliza o modo de predição com atraso unitário. Os resultados de ambos os esquemas são apresentados na Figura 6-9. O espectro do GRS oferece um exemplo especifico que exige uma técnica eficiente de codificação fonte para dados com baixa-entropia. Para o espectro de 5 segundos, mais de 90% dos dados são codificados com o uso das opções de Bloco-Zero e de Segunda-Extensão.

ANEXO A

VERIFICAÇÃO DA IMPLEMENTAÇÃO DO CODIFICADOR DE COMPRESSÃO DE DADOS LIVRE DE PERDAS DO CCSDS

A1 FINALIDADE

O objetivo deste apêndice é o de apresentar um conjunto limitado de vetores de teste para verificar as implementações de software ou de hardware, no uso do Codificador de Compressão de Dados (Livre de Perdas) do CCSDS. Este conjunto limitado de vetores de teste pode ser obtido a partir do software que pode ser obtido no seguinte endereço (Web):

http://public.ccsds.org/sites/cwe/sls-dc/

NOTA – O programa e os dados são fornecidos aos usuários, como cortesia, tal "como apresentados". Em nenhuma hipótese: o CCSDS, suas Agências (Membro), ou o autor do software são passíveis de responsabilidade por qualquer conseqüência negativa, incidente ou danos indiretos decorrentes do uso ou da inabilidade no uso deste software.

A2 DESCRIÇÃO

O diretório contém arquivos que criam um conjunto limitado de dados de teste para o Codificador de Compressão de Dados Livre de Perdas do CCSDS. Este conjunto de dados permite obter: todas as opções *"split-sample"*, a opção de não compressão e a opção de *"second-extension"*, a ser selecionada uma única vez. A opção de blocos-zero será selecionada no mínimo uma vez para valores com até n=14, que é o caso onde somente 256 pontos de dados se fazem necessários. Se o usuário permitir a escrita de mais de 256 pontos de dados, então, a opção de blocos-zero deverá ser usada para todos os valores de n, até n=16. Os programas atuais de teste fornecem dados com resolução de n entre 4 e 16 bits.

ANEXO B

ACRÔNIMOS E ABREVIATURAS

AOS	Sistema Orbital Avançado
APID	Identificador de Processo de Aplicação
AVIRIS	Espectrômetro de Imagem Infravermelho Visível de Bordo
CADU	Unidade de Dados de Canal de Acesso
CCD	Dispositivo Acoplado por Carga
CCSDS	Comitê Consultivo para Sistemas Espaciais de Dados
CDS	Conjunto de Dados Codificados
CR	Taxa de Compressão
CVCDU	Unidade de Dados Codificados de Canal Virtual
FS	Seqüência Fundamental
GHRS	Espectrômetro Goddard de Alta Resolução
GRS	Espectrômetro de Raios Gama
HSI	Imageador Hiperespectral
HST	Telescópio Espacial Hubble
LDC	Compressão de Dados Livre de Perdas
LSB	Bit Menos Significativo
LZW	Lempel Ziv Welch (algoritmo)
MPDU	Unidade de Dados do Protocolo de Multiplexação
NRZ	Não-Retorno a Zero
pdf	Função de Distribuição de Probabilidade
PSK	Chaveamento por Desvio-de-Fase
RF	Radiofreqüência
ROS	Resto-do-Segmento
SDU	Unidade de Dados de Serviço
SMEX	Pequeno Explorador
SPDF	Campo de Dados do Pacote Fonte
SSTI	Pequeno Satélite de Infusão Tecnológico
SWAS	Satélite Astronômico de Ondas Submilimétricas

SWIR	Ondas Curtas de Infravermelho
SXT	Telescópio de Raios-X Macios
TM	Mapeador Temático
VCDU	Unidade de dados de Canal Virtual
VNIR	Infravermelho-Próximo-Visível
WFPC	Câmera Planetária de Campo Largo