



ABNT: ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS
CB: 08 COMITÊ BRASILEIRO DE AERONÁUTICA E ESPAÇO
SC: 08.010 SUBCOMITÊ DE ATIVIDADE ESPACIAL
CE: 08:010.70 COMISSÃO DE ESTUDO EM SISTEMAS ESPACIAIS
DE TRANSFERÊNCIA DE DADOS E DE INFORMAÇÃO

Sistemas Espaciais de Transferência de Dados e de Informação

COMPRESSÃO DE DADOS LIVRE DE PERDAS

**Norma Recomendada
CCSDS 121.0-B-1
LIVRO AZUL
Maio 1997**

**Relatório Técnico da Comissão – 02 (RTC-02)
– Versão Editorial - 1: Junho de 2010 –**

**CE 08:010.70 Comissão de Estudos em Sistemas Espaciais de
Transferência de Dados e de Informação**

RECOMENDAÇÃO CCSDS PARA PROTOCOLO DE ENLACE ESPACIAL DE
DADOS DE TELECOMANDO (TC)

COLABORADORES DESTA EDIÇÃO

A editoração deste documento interno, denominado Relatório Técnico da Comissão N° 02, ou RTC – 02, contou com a participação seguintes MEMBROS da Comissão de Estudo de Sistemas Espaciais de Transferência de Dados e de Informação – CE 08:010.70, da ABNT:

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/UFGM
Marília 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

A Secretaria do órgão do INPE (RME/TEC) que coordena esta Comissão de Estudo da ABNT, tem realizado extenso trabalho de apoio de editoração e de correspondência, na elaboração deste documento, neste caso, com a atuação da Secretária:

Síntique Rodrigues dos Santos	Secretária	RME/TEC/INPE
-------------------------------	------------	--------------

As seguintes Secretarias auxiliares, em treinamento, do órgão do INPE (RME/TEC) também prestam assistência na elaboração deste documento:

Helen Joyce Aparecida F. Leite	Secretária Auxiliar	RME/TEC/INPE
Izabela Moraes de Oliveira	Secretária Auxiliar	RME/TEC/INPE

São José dos Campos, Julho de 2010

NOTA DO TRADUTOR

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 Transferência de Dados e de 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 que deu origem a esta Norma Recomendada possui o título: “Lossless Data Compression” – Part 1 sob o código CCSDS 121.0-B-1, Livro Azul, de Maio de 1997, elaborada pelo “Consultative Committee for Space Data Systems” como “Recommendation for Space Data System Standards”.

AUTORIDADE

Edição: Livro Azul, Edição 1
Data: Maio 1997
Local: São José dos Campos
São Paulo, Brasil

Este documento foi aprovado para publicação pelo Conselho de Gerenciamento do Comitê Consultivo para Sistemas Espaciais de Dados (CCSDS) e representa o consenso de entendimento técnico das Agências Membro participantes do CCSDS. O procedimento para revisão e autorização de Recomendações do CCSDS está detalhado no Manual de Procedimentos do Comitê Consultivo para Sistemas Espaciais de Dados (referência [B1]), e o registro de participação da Agência na autorização deste documento pode ser obtido com o Secretariado do CCSDS no endereço dado, abaixo.

Esta Recomendação é publicada e mantida por:

Secretariado do CCSDS
Divisão de Integração de Programa (Código MG)
Administração Nacional de Aeronáutica e Espaço – NASA
Washington, DC 20546, EUA

DECLARAÇÃO DE INTENÇÃO

O Comitê Consultivo para Sistemas Espaciais de Dados (CCSDS) é uma organização oficialmente estabelecida e gerenciada pelas suas Agências espaciais membro. O Comitê se reúne periodicamente para tratar de problemas de sistemas de dados que são comuns a todos os seus participantes, e para formular soluções técnicas seguras para estes problemas. Pelo fato da participação no CCSDS ser plenamente voluntária, os resultados das ações do Comitê são definidos como sendo **Recomendações** que não são consideradas de uso obrigatório para qualquer uma de suas Agências membro.

Esta **Recomendação** é publicada e representada como resultado de consenso do corpo Plenário do CCSDS. O endosso de cada Agência a esta **Recomendação** é inteiramente voluntário. Este endosso, no entanto, implica os seguintes entendimentos:

- Sempre que uma Agência estabelece um **padrão** relacionado com o CCSDS, este **padrão** estará em acordo com a **Recomendação** pertinente ao caso. O estabelecimento de um tal **padrão** não exclui outras disposições que uma agência possa desenvolver para ele.
- Sempre que uma Agência estabelece um **padrão** relacionado com o CCSDS, a Agência deverá prover as outras Agências Membro do CCSDS com as seguintes informações:
 - O **padrão** propriamente dito.
 - Com antecipação, a data da sua capacitação operacional inicial.
 - Com antecipação, a duração operacional do seu respectivo serviço.
- Acordos específicos de serviço devem ser feitos via memorando de entendimento. Nenhuma **Recomendação** ou qualquer **padrão** a ela subsequente substitui um memorando de entendimento.

Em prazo não além de cinco anos após a sua publicação esta Recomendação será revista pelo CCSDS para determinar se ela deverá: (1) Continuar efetiva, sem sofrer modificações; (2) Ser modificada para refletir o impacto de novas tecnologias, novos quesitos, ou novas orientações; ou 3) Ter a sua adoção descontinuada ou cancelada.

Nos casos em que uma nova versão da **Recomendação** for editada, padrões e implementações existentes da Agência relacionadas com o CCSDS não passarão a ser consideradas não compatíveis com o CCSDS, por este motivo. Cabe a cada Agência a responsabilidade de determinar quando tais padrões ou implementações devem ser modificadas ou atualizadas. Cada Agência é fortemente encorajada a orientar o seu planejamento para adoção dos novos padrões e implementações em acordo com uma versão posterior da Recomendação.

PREFÁCIO

Esta Recomendação estabelece uma estrutura e provê uma base comum para um algoritmo de Compressão de Dados Livre de Perdas, aplicável a diversos tipos diferentes de sistemas espaciais de dados.

Através de um processo de evolução natural, é esperado que a expansão, eliminação ou modificação deste documento possa vir a ocorrer. Esta Recomendação é, portanto, sujeita aos procedimentos e controle do CCSDS para o gerenciamento e modificação de documentos e que estão definidos no *Manual de Procedimentos para o Comitê Consultivo para Sistemas Espaciais de Dados* (referência [B1]). Versões atuais dos documentos CCSDS são mantidas no site Web do CCSDS, dado por:

<http://www.ccsds.org>

Questões relacionadas com o conteúdo e o status deste documento devem ser endereçadas ao Secretariado do CCSDS, no endereço indicado na página i.

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

Por ocasião desta publicação, as Agencias Membro e Observadoras ativas do CCSDS, eram:

Agencias 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.
- Deutsche Forschungsanstalt für Luft- und Raumfahrt e.V. (DLR)/Germany.
- European Space Agency (ESA)/Europe.
- Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
- National Aeronautics and Space Administration (NASA)/USA.
- National Space Development Agency of Japan (NASDA)/Japan.
- Russian Space Agency (RSA)/Russian Federation.

Agencias Observadoras

- Austrian Space Agency (ASA)/Austria.
- 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.
- Communications Research Laboratory (CRL)/Japan.
- Danish Space Research Institute (DSRI)/Denmark.
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
- European Telecommunications Satellite Organization (EUTELSAT)/Europe.
- Federal Service of Scientific, Technical & Cultural Affairs (FSST&CA)/Belgium.
- Hellenic National Space Committee (HNSC)/Greece.
- Indian Space Research Organization (ISRO)/India.
- Industry Canada/Communications Research Centre (CRC)/Canada.
- Institute of Space and Astronautical Science (ISAS)/Japan.
- Institute of Space Research (IKI)/Russian Federation.
- KFKI Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
- MIKOMTEK: CSIR (CSIR)/Republic of South Africa.
- Korea Aerospace Research Institute (KARI)/Korea.
- Ministry of Communications (MOC)/Israel.
- National Oceanic & Atmospheric Administration (NOAA)/USA.
- National Space Program Office (NSPO)/Taipei.
- Swedish Space Corporation (SSC)/Sweden.
- United States Geological Survey (USGS)/USA.

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

CONTROLE DO DOCUMENTO

Documento	Título	Data	Status
CCSDS 121.0-B-1	Compressão de Dados Livre de Perdas, ou Edição 1	Maio 1997	Edição Original

SUMÁRIO

<u>Seção</u>	<u>Página</u>
1 INTRODUÇÃO	1
1.1 OBJETIVO	1
1.2 ESCOPO	1
1.3 APLICABILIDADE	1
1.4 CRITÉRIO	1
1.5 NOMENCLATURA E CONVENÇÃO DE NUMERAÇÃO DE BITS	2
1.6 REFERÊNCIAS	2
2 REVISÃO	1
2.1 CONCEITO GERAL	1
2.2 O CODIFICADOR FONTE	1
2.3 EMPACOTAMENTO DOS DADOS CODIFICADOS	3
2.4 CONTROLE DE ERRO	3
3 CODIFICADOR POR ENTROPIA ADAPTATIVA	1
3.1 ESPECIFICAÇÃO DA CODIFICAÇÃO	1
3.2 SEQUÊNCIA FUNDAMENTAL	2
3.3 PARTIÇÃO DE AMOSTRA	3
3.4 OPÇÕES DE BAIXA ENTROPIA	3
3.5 SEM COMPRESSÃO	4
3.6 SELEÇÃO DE CÓDIGO	4
4 PRÉ-PROCESSADOR	1
4.1 FUNÇÃO DO PRÉ-PROCESSADOR	1
4.2 PREDITORES	1
4.3 AMOSTRA DE REFERÊNCIA	2
4.4 MAPEADOR DE PREDIÇÃO DE ERRO	2
5 FORMATO DOS DADOS	1
5.1 ESTRUTURAS DE DADOS LIVRE DE PERDAS	1
5.2 FORMATO DO PACOTE	4
6 PACOTE DE IDENTIFICAÇÃO DE COMPRESSÃO (OPCIONAL)	1
6.1 ESTRUTURA DO PACOTE DE IDENTIFICAÇÃO DE COMPRESSÃO	1
6.2 CABEÇALHO PRIMÁRIO DO CIP	1
6.3 CAMPO DE DADOS DE PACOTE	2
ANEXO A GLOSSÁRIO DE ACRÔNIMOS E TERMOS	A-1
ANEXO B REFERÊNCIAS INFORMATIVAS	B-1

SUMÁRIO (Continuação)

<u>Figura</u>	<u>Página</u>
2-1 Esquema do codificador.....	1
3-1 Codificador por Entropia Adaptativa com um Pré-processador	1
3-2 Formato da Partição da Amostra.....	3
4-1 Um Pré-processador.....	1
4-2 Pré-processador que Utiliza um Preditor com Atraso Unitário	2
5-1 Formato do CDS Quando a Opção de Amostras Particionadas é Seleccionada.....	2
5-2 Formato CDS Quando a Opção de Não-Compressão é Seleccionada.....	3
5-3 Formato CDS Quando a Opção de Bloco-Zero é Seleccionada	3
5-4 O Formato CDS Quando a Opção da Segunda-Extensão é Seleccionada.....	4
5-5 Formato de Pacote para l grupos de dados em formato CDS	4
6-1 Estrutura do Pacote de Identificação de Compressão	1
6-2 Campo da Configuração Fonte	3

<u>Tabela</u>	<u>Página</u>
3-1 Seqüência Fundamental de Palavras-código em Função das Amostras Pré-processadas.....	2
3-2 Palavras-Código da Seqüência Fundamental de Blocos-Zero Como Função do Número de Blocos-Zero Consecutivos	4
5-1 Identificação da Chave de Opção do Código Seleccionado	1

1 INTRODUÇÃO

1.1 OBJETIVO

O objetivo deste documento é o de estabelecer uma Norma para um algoritmo de compressão de dados livre de perdas para utilização na codificação de dados fonte, aplicado a dados digitais de sistemas espaciais de dados, incluindo também a especificação de como os dados comprimidos devem ser inseridos em pacotes de dados fonte, para posterior recuperação e decodificação.

A codificação fonte para a compressão de dados é um método utilizado em sistemas de dados para se obter a redução do volume de dados digitais, com o propósito de se obter os benefícios descritos, mas não limitados à:

- a) redução da largura de banda do canal de transmissão;
- b) redução dos quesitos de armazenamento temporário e de massa, em memórias;
- c) redução do tempo de transmissão de dados para uma certa velocidade.

1.2 ESCOPO

As características dos códigos fonte são especificadas somente na medida necessária para que elas possam assegurar o apoio à missões múltiplas. A especificação não se propõe a quantificar a redução relativa da largura de banda, os méritos de cada abordagem discutida, ou os quesitos de projeto definidos para os codificadores e para os decodificadores associados. Algumas informações de desempenho estão incluídas na referência [5].

Esta Recomendação trata somente de codificação fonte Livre de Perdas, que é aplicável a uma grande variedade de dados digitais, tanto de imageamento como não-imageamento, onde há o quesito de uma redução limitada da taxa de dados que deve permitir que nenhuma distorção venha a ser adicionada ao processo de compressão e descompressão dos dados. O processo de descompressão não é considerado, neste contexto. Veja na referência [5] um resumo de uma aplicação.

1.3 APLICABILIDADE

Esta Recomendação destina-se a aplicações de compressão de dados em missões espaciais que prevêm o apoio cruzado à telemetria por pacotes. Adicionalmente, ela serve como guia para o desenvolvimento, neste campo, de padrões compatíveis por parte das Agências relacionadas com o CCSDS, observada também a boa prática da engenharia.

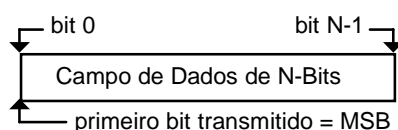
1.4 CRITÉRIO

O conceito e critério utilizados para a codificação fonte Livre de Perdas no algoritmo de compressão de dados descrito a seguir pode ser encontrado na referência [5].

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

1.5 NOMENCLATURA E CONVENÇÃO DE NUMERAÇÃO DE BITS

Neste documento, a seguinte convenção é utilizada para identificar cada unidade de informação digital binária ou ‘bit’, num campo de N-bits. O primeiro bit do campo a ser transmitido (isto é, o que está justificado mais a esquerda quando se desenha uma figura) tem a sua posição definida como sendo a do “Bit 0”; o bit seguinte é definido como sendo o ‘Bit 1’ e, assim por diante, até o ‘Bit N-1’. Quando o campo é utilizado para expressar um valor binário (tal como em um contador), o bit mais significativo (MSB, ‘Most Significant Bit’) será o primeiro bit deste campo a ser transmitido, isto é, ele é o ‘Bit 0’.



De acordo com a prática moderna de comunicação de dados, os campos de dados referentes a uma espaçonave são freqüentemente agrupados em ‘palavras’ de 8 bits que observam a convenção apresentada acima. Ao longo desta Recomendação, a seguinte nomenclatura é utilizada para descrever este agrupamento:

Palavra de 8 Bits = ‘Octeto’

1.6 REFERÊNCIAS

Os documentos seguintes contêm dispositivos que, referenciadas neste texto, são também adotados nesta Recomendação. Na ocasião em que se deu esta publicação, as edições aqui apontadas eram as que estavam em vigor. Todos estes documentos estão sujeitos à revisão e os usuários desta Recomendação são encorajados a pesquisar a possibilidade de utilização de edições mais recentes dos documentos indicados abaixo. O Secretariado do CCSDS mantém o registro das Recomendações CCSDS que se encontram em vigor. A última versão destes documentos encontra-se no site <www.ccsds.org>.

- [1] *TM Synchronization and Channel Coding*. Recomendação para Padronização de Sistemas Espaciais de Dados, CCSDS 131.0-B-1. Livro Azul. Versão 1. Washington, D.C.: CCSDS, Setembro 2003.
- [2] *Packet Telemetry*. Recomendação para Padronização de Sistemas Espaciais de Dados, CCSDS 133.0-B-1. Livro Azul. Versão 1. Washington, D.C.: CCSDS, Setembro 2003.

2 REVISÃO

2.1 CONCEITO GERAL

Esta Recomendação define, para fins de padronização, um algoritmo adaptativo de codificação específico para codificação fonte que possui aplicabilidade muito difundida para vários tipos de dados digitais. Em particular, dados científicos obtidos com diversos tipos de instrumentos, imagens ou não, se adequam bem à aplicação deste algoritmo.

Existem duas classes de métodos de codificação: Livre de Perdas ('Lossless') e Com Perdas ('Lossy').

A técnica de codificação Livre de Perdas preserva a integridade dos dados originais e remove as redundâncias dos dados fonte. No processo de decodificação, os dados originais podem ser recuperados a partir dos dados comprimidos, desta forma, restaurando as redundâncias originais; o processo de descompressão não adiciona distorção. Esta técnica é particularmente útil quando a integridade dos dados não pode ser comprometida. O preço pago é geralmente uma baixa Taxa de Compressão, que é definida como sendo a razão entre o número de bits originais não comprimidos e o número de bits após a compressão, incluindo os bits adicionais que se fazem necessários para o uso de parâmetros de sinalização.

Por outro lado, um método de codificação fonte Com Perdas remove algumas das informações originais, juntamente com a redundância. Os dados originais não podem ser recuperados em sua totalidade, o que provoca a distorção dos mesmos, após a decodificação. Entretanto, se alguma distorção pode ser tolerada, a codificação com perdas geralmente permite a obtenção de uma alta taxa de compressão. Se houver controle da distorção em níveis considerados toleráveis, esta técnica pode permitir aquisição e a disseminação de dados de uma missão, dentro de um prazo considerado crítico.

Esta Recomendação se relaciona apenas à codificação Livre de Perdas e não tem o propósito de explicar a teoria que justifica operação do algoritmo.

2.2 O CODIFICADOR FONTE

O codificador fonte Livre de Perdas consiste de duas partes funcionais separadas: o pré-processador e o codificador com entropia adaptativa, como mostrado na figura 2-1.

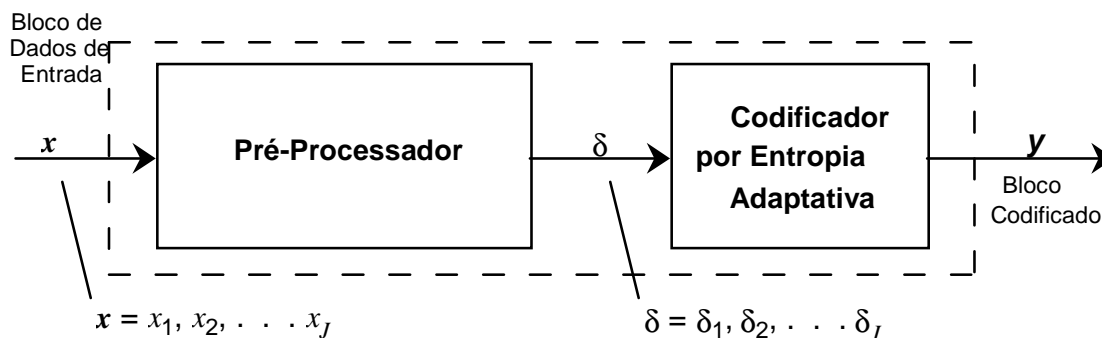


Figura 2-1: Esquema do Codificador Fonte.

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

As entradas para o codificador fonte são

$$\mathbf{x} = x_1, x_2, \dots, x_J$$

que é um bloco de J amostras de n -bits, onde n é um valor constante.

Pré-processador:

O pré-processador aplica uma função reversível às amostras de dados de entrada \mathbf{x} , para produzir uma 'fonte preferencial':

$$\boldsymbol{\delta} = \delta_1, \delta_2, \dots, \delta_i, \dots, \delta_J$$

onde cada dado δ_i é um inteiro de n -bits, $0 \leq \delta_i \leq (2^n - 1)$. Para um pré-processador ideal, $\boldsymbol{\delta}$ terá as seguintes propriedades:

- a) o conjunto de dados $\{\delta_i\}$ é estatisticamente independente e identicamente distribuído;
- b) a probabilidade preferencial, p_m , que indica que qualquer amostra δ_i terá no valor inteiro m é uma função não crescente de valor m , para $m = 0, 1, \dots, (2^n - 1)$.

A função do pré-processador é uma operação reversível e, em geral, o melhor pré-processador Livre de Perdas irá satisfazer as condições indicadas acima e, desta forma, produzir a mais baixa entropia, que é a medida do menor número de bits médio que pode ser utilizado para representar cada amostra.

Esta Recomendação não tentará explicar os métodos que resultam na escolha de um estágio de pré-processamento. Esta Recomendação provê a definição de um estágio básico de pré-processamento que pode ser adequado para muitas aplicações. Entretanto, é importante que o usuário trate cuidadosamente desta questão uma vez que a seleção cuidadosa de um estágio adequado de pré-processamento é essencial na obtenção de uma compressão eficiente e ela depende das características dos dados fonte. Usuários interessados nesta questão específica poderão recorrer à referência [B2].

Codificador por Entropia Adaptativa:

A função do Codificador por Entropia Adaptativa é a de calcular e de decifrar de forma única, as palavras codificadas de tamanho variável correspondentes a cada bloco da amostras de entrada produzida pelo pré-processador. O codificador por entropia incorpora múltiplas opções de codificação, sendo que cada uma delas apresenta um desempenho eficiente sobre diferentes faixas de entropia, ainda que estas possam ser parcialmente sobrepostas. O codificador seleciona a opção de codificação que proporciona a mais alta taxa de compressão entre as várias opções

existentes, para o mesmo bloco de J amostras. Um ‘identificador’ de opção de código, que necessita de apenas alguns bits, é anexado antes do primeiro bit da primeira palavra codificada em um bloco codificado para que possa ser, desta forma, sinalizado ao decodificador de descompressão qual é a opção que foi utilizada na compressão. Uma vez que o tamanho do bloco J é pequeno (16 amostras ou menos) e que uma nova opção de codificação é selecionada para cada bloco, a codificação pode ser ajustada para adaptar à mudanças rápidas na estatística dos dados.

2.3 EMPACOTAMENTO DOS DADOS CODIFICADOS

O fluxo de bits codificados com comprimento variável, representando um bloco de J amostras constitui um Conjunto de Dados Codificados (CDS, ou ‘Coded Data Set’). A estrutura de um pacote de telemetria CCSDS é aquela que é recomendada para transporte dos CDSs que estarão contidos no campo deste tipo de pacote. A informação relacionada com, por exemplo: o sensor, missão, tempo, e outros detalhes específicos da missão serão necessários para o roteamento e a contabilização dos pacotes, e ela estará contida no Cabeçalho Primário do Pacote e (se for o caso) no Cabeçalho Secundário do Pacote (veja referência [2]).

2.4 CONTROLE DE ERRO

Erros em bits de canais individuais têm consequências indesejáveis marcantes quando os dados estão comprimidos. Mesmo assim, as consequências podem ser catastróficas. Por essa razão, para que se possa limitar a propagação de erro quando da utilização do algoritmo de codificação, tal como está descrito neste documento, as seguintes ações são recomendadas:

- a) Usar a codificação do canal de telemetria tal como está descrito na referência [1];
- b) Usar a telemetria por pacotes tal como está descrita na referência [2].

3 CODIFICADOR POR ENTROPIA ADAPTATIVA

3.1 ESPECIFICAÇÃO DA CODIFICAÇÃO

3.1.1 A Figura 3-1 representa o Codificador com Entropia Adaptativa com um pré-processador. Basicamente, tal codificador escolhe um código dentre um conjunto de opções de códigos utilizados para representar um bloco de amostras de dados pré-processados, δ . y representa o vetor de dados de saída, codificados por entropia adaptativa. Uma sequência única de bits de identificação (ID) é anexada ao bloco de códigos para indicar ao decodificador qual é a opção de decodificação que deve ser utilizada.

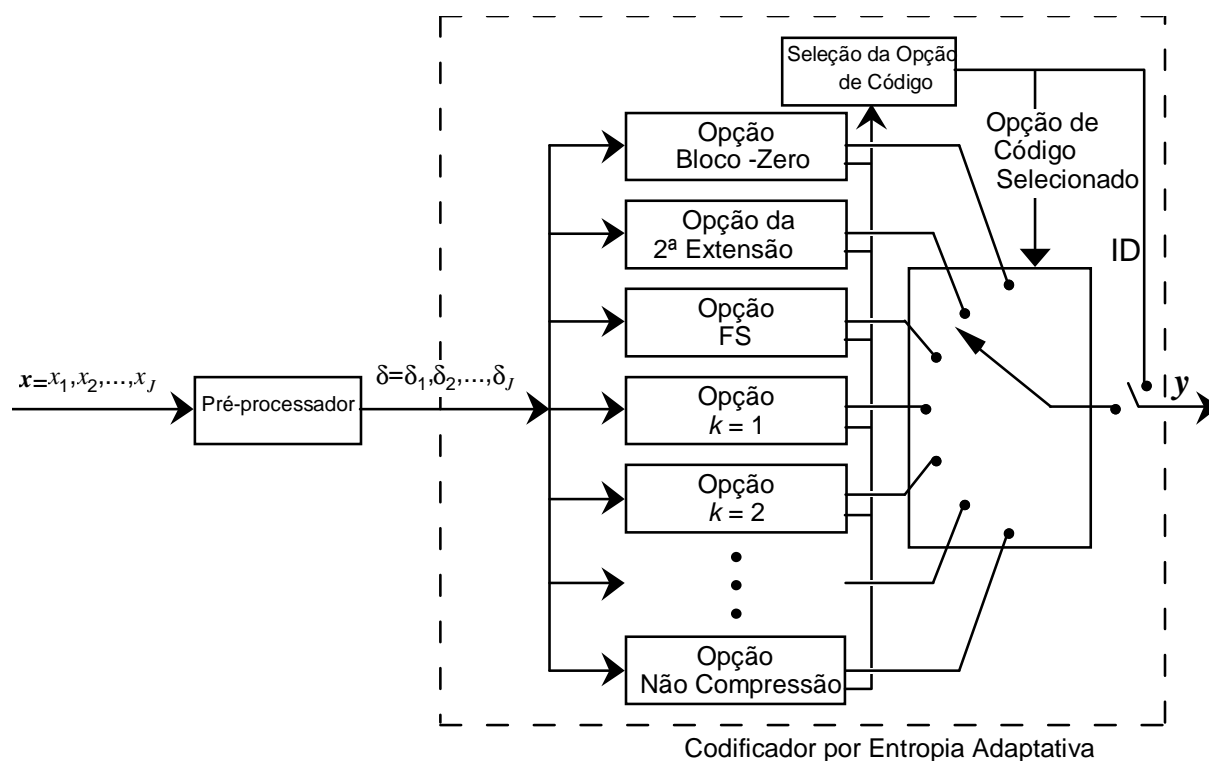


Figura 3-1: O Codificador por Entropia Adaptativa com um Pré-processador

NOTA - A Figura 3-1 ilustra o princípio do Codificador por Entropia Adaptativa com um pré-processador; ela não ilustra uma implementação.

3.1.2 O código básico selecionado é um código de tamanho variável que utiliza a técnica de codificação adaptativa de Rice (consultar referência [B2]). De acordo com a técnica de codificação de Rice, vários algoritmos são aplicados concorrentemente a um bloco de J amostras consecutivas, pré-processadas. A opção de algoritmo que produz a sequência de palavra-codificada mais curta para um dado bloco de dados é aquela que é selecionada para transmissão. A opção de bloco zero é um caso especial em que uma única sequência de palavra codificada representa um ou mais blocos consecutivos de J amostras pré-processadas (veja 3.4.3). Em todas

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

as outras opções, a sequência de uma palavra-codificada é representada por um único bloco de *amostras consecutivas* pré-processadas.

3.1.3 As seguintes variáveis são necessárias para uso da técnica de codificação adaptativa de Rice:

- tamanho do bloco, J ;
- resolução, n (número de bits-de-entrada/amostra);
- a sequência de bits de identificação (ID) da opção de código selecionado.

3.1.4 As seguintes restrições devem ser aplicadas ao esquema de codificação adaptativa de comprimento variável do Codificador com Entropia:

$$\begin{aligned} J &= 8 \text{ ou } 16 \text{ amostras por bloco;} \\ n &= \text{resolução com um máximo de 32 bits por} \\ &\quad \text{amostra com valores digitais com sinal,} \\ &\quad \text{dentro dos limites de } 0 \text{ à } 2^n-1 \text{ (positivos,} \\ &\quad \text{somente), ou de } -2^{n-1} \text{ à } 2^{n-1}-1 \text{ (com sinal,} \\ &\quad \text{positivo ou negativo).} \end{aligned}$$

3.2 SEQUÊNCIA FUNDAMENTAL

A opção mais básica é a de uma palavra codificada do tipo Sequência Fundamental (FS) de comprimento variável, constituída de m zeros seguidos pelo dígito um (1) que resulta de uma amostra pré-processada $\delta_i = m$. A Tabela 3-1 ilustra as possíveis palavras codificadas do tipo FS. Uma Sequência Fundamental resulta da concatenação de J palavras codificadas do tipo FS.

Tabela 3-1: Sequência Fundamental de Palavras Codificadas em Função das Amostras Pré-processadas

Valores Possíveis de Amostras Pré- processadas, δ_i	Palavras do Tipo Codificadas SF
0	1
1	01
2	001
.	.
.	.
.	.
2^n-1	0000 . . . 00001
	$\underbrace{\hspace{1.5cm}}_{(2^n-1 \text{ zeros})}$

3.3 PARTIÇÃO DE AMOSTRA

3.3.1 A k -ésima partição de amostra é obtida com a remoção dos bits menos significativos (LSBs, ou “least-significant bits”) k da representação binária de cada amostra pré-processada, δ_i , e codificando os bits restantes em uma palavra codificada FS (veja figura 3-2). Isto produz uma palavra codificada de tamanho variável. As palavras codificadas FS para o bloco de J amostras pré-processadas em questão são transmitidas sempre com os LSBs removidos, precedidos por um campo de identificação ID que indica qual é o valor de k . Este processo possibilita a adaptação do tamanho da palavra codificada à estatística dos dados fonte.

3.3.2 A opção FS descrita na secção 3.2 pode ser entendida como sendo um caso especial de partição de amostra pré-processada onde $k = 0$.

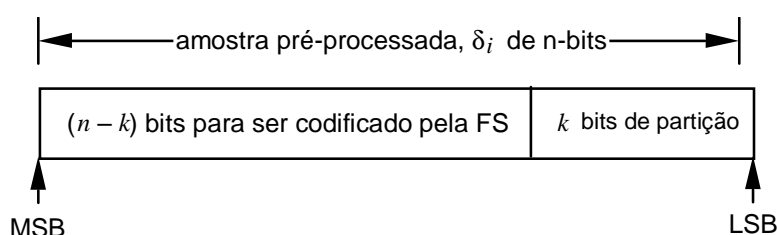


Figura 3-2: Formato da Partição da Amostra

3.4 OPÇÕES DE BAIXA ENTROPIA

3.4.1 CONCEITO GERAL

As duas opções de código, a opção de Segunda Extensão¹ e a opção de Bloco-Zero, resultam em uma codificação mais eficiente do que outras opções quando os dados pré-processados são altamente compressíveis.

3.4.2 A OPÇÃO DE SEGUNDA EXTENSÃO

Quando a opção de Segunda Extensão é selecionada, cada par de amostras pré-processadas a partir de um bloco de amostras J é transformado e codificado com o uso de palavras-código FS. Entende-se que δ_i e δ_{i+1} são os pares de amostras adjacentes de um bloco de amostras J , formado por dados pré-processados. Cada par é transformado em um único novo símbolo γ_j pela seguinte equação:

$$\gamma_j = (\delta_i + \delta_{i+1}) (\delta_i + \delta_{i+1} + 1) / 2 + \delta_{i+1}, \text{ aonde } j = 1, 2, \dots, J/2$$

Os $J/2$ símbolos transformados de um bloco são codificados usando palavras codificadas FS segundo a tabela 3-1. O processo acima requer que J seja um número par, inteiro que observe os valores recomendados em 3.1.4 ($J = 8$ ou 16).

¹A Primeira Extensão de uma amostra pré-processada se identifica com a própria amostra, pré-processada.

3.4.3 OPÇÃO DE BLOCO-ZERO

3.4.3.1 A opção Blocos-Zero é selecionada quando um ou mais blocos de amostras pré-processadas são compostas por zeros. Neste caso, uma única palavra codificada pode representar vários blocos de amostras pré-processadas, diferentemente das outras opções nas quais uma palavra codificada FS representa somente uma ou duas amostras pré-processadas.

3.4.3.2 O conjunto de blocos r existente entre amostras de referências consecutivas, tal como está descrito em 4.3, é particionado em um ou mais segmentos. Cada segmento, exceto possivelmente o último, contém s blocos. O valor recomendado para s é 64.

3.4.3.3 Dentro de cada segmento, cada grupo de blocos-zero adjacentes entre si é codificado com o uso das palavras codificadas FS, tal como especificadas na tabela 3-2, o que fornece a identidade do tamanho de cada grupo. A palavra codificada do Resto-do-Segmento (ROS, ou “Remainder-Of-Segment”) na tabela 3-2 é utilizada para indicar que o restante de um segmento consiste de cinco ou mais blocos-zero.

Tabela 3-2: Palavras-Código da Sequência Fundamental de Blocos-Zero Como Função do Número de Blocos-Zero Consecutivos

<u>Número de Blocos-de-Zeros</u>	<u>Palavra-Codificada FS</u>
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)

3.5 SEM COMPRESSÃO

A última opção é aquela em que a compressão de dados não é aplicada. Se esta for a opção selecionada, o bloco de amostras pré-processadas não sofre alteração e recebe um campo de identificação que é a ele anexado.

3.6 SELEÇÃO DE CÓDIGO

O Codificador por Entropia Adaptativa inclui uma função de seleção de código que é aquela que seleciona a opção de código que, por sua vez, resulta no melhor desempenho quando aplicada sobre o bloco de amostras, sendo considerado. A seleção é feita com base no número de bits que

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

a opção selecionada irá utilizar para codificar o bloco de amostras, em questão. Uma seqüência de bits especifica qual é a opção que foi utilizada para codificar o conjunto de palavras codificadas as quais ela se refere. As seqüências de bits de identificação, IDs são apresentadas na tabela 5-1.

4 PRÉ-PROCESSADOR

4.1 FUNÇÃO DO PRÉ-PROCESSADOR

4.1.1 Dois fatores que contribuem para a medida de desempenho da taxa de bits de codificação (bits/amostra) na técnica de compressão Livre de Perdas são: a quantidade de correlação removida dentre as amostras de dados no estágio de pré-processamento, e a eficiência na codificação exercida pelo codificador por entropia. A função do pré-processador é a de decorrelacionar os dados de entrada e a de reformatá-los em inteiros não negativos de forma compatível com a distribuição de probabilidade preferencial. Existem situações em que um pré-processador não é necessário (veja referência [B2]), diante das quais ele pode ser evitado para prover melhor desempenho de compressão.

4.1.2 Um pré-processador possui duas funções: de predição e a de mapeamento, como mostrado na Figura 4. O pré-processador subtrai o valor de predição, \hat{x}_i , do valor do dado corrente, x_i . O erro resultante de predição com $(n+1)$ -bits, Δ_i , é então mapeado para um valor inteiro, δ_i , de n -bits, baseado no valor de predição, \hat{x}_i . Quando o preditor é escolhido adequadamente, o erro de predição tende a ser pequeno e, para alguns tipos de dados fonte, apresenta uma probabilidade de distribuição próxima da Laplaciana, para o qual o Codificador com Entropia Adaptativa é ótimo. Existem várias técnicas de pré-processamento, das quais apenas uma, o Preditor com Atraso Unitário, tal como descrito em 4.2, é apresentado por esta Recomendação (veja referência [B2] para exemplos de preditores).

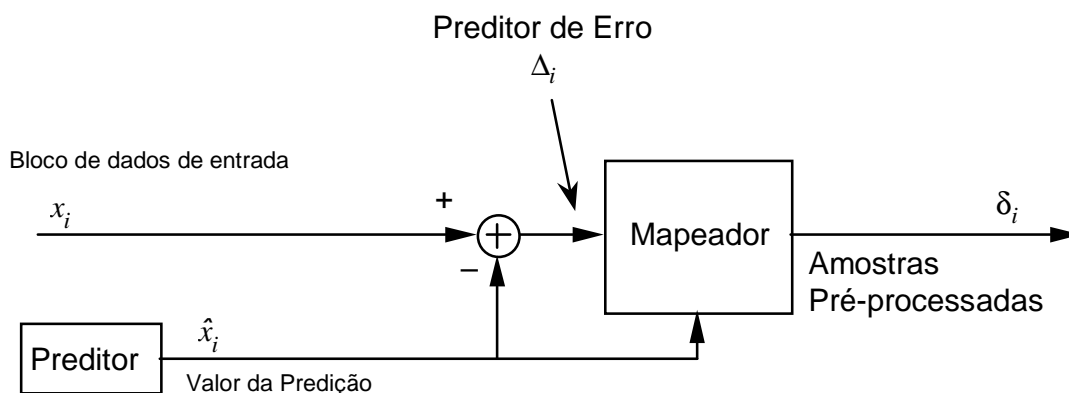


Figura 4-1: Um Pré-processador

4.2 PREDITORES

4.2.1 TÉCNICAS DE PREDIÇÃO

Várias técnicas de pré-processamento podem ser utilizadas com o Codificador com Entropia Adaptativa. Uma técnica, aquela que utiliza o Preditor com Atraso Unitário, está especificada no item 4.2.2, abaixo. Um preditor definido para uma aplicação específica pode ser utilizado em lugar de um Preditor com Atraso Unitário, mas tal tipo de preditor seria único para uma aplicação em particular e, por este motivo, ele não é especificado por esta Recomendação.

4.2.2 PREDITOR COM ATRASO UNITÁRIO

A técnica do preditor com atraso unitário ilustrado na figura 4.2 utiliza o sinal de dado de amostra de entrada com atraso unitário como sendo o valor de predição para o sinal de dados corrente, de tal forma que a predição de erro é um valor passado para o mapeador, proveniente do preditor, para mapeamento em um número inteiro.

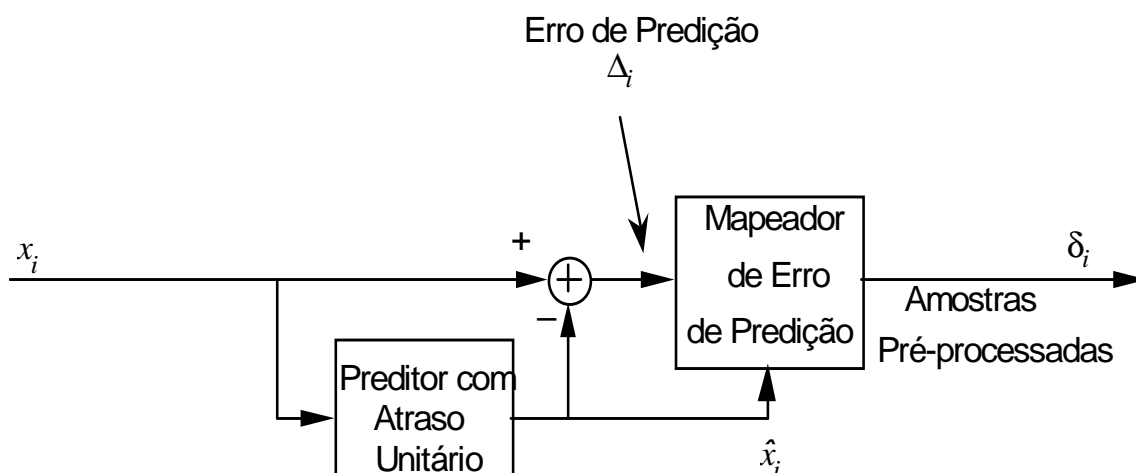


Figura 4-2: Pré-processador que Utiliza um Preditor com Atraso Unitário

4.3 AMOSTRA DE REFERÊNCIA

Uma amostra de referência é uma amostra da entrada de dados inalterada sobre a qual se baseia a amostra de predição que a ela sucede. Quando um preditor de atraso unitário ou outros preditores de ordens mais altas que utilizam dados atrasados para predição são empregados, se faz necessário o uso de referências pelo decodificador para que ele possa ter condições de recuperar os valores das amostras, a partir dos erros decodificados. Quando a amostra de referência é inserida, existem $J - 1$ amostras pré-processadas no CDS. O usuário precisa determinar com que frequência as referências devem ser inseridas. O intervalo da amostra de referência, r , é limitado a um valor máximo de 256 CDSes (isto é, 4096 amostras, quando $J = 16$). Quando a referência não é utilizada no pré-processador, o parâmetro r serve para definir um intervalo dentre os blocos de entrada da amostra dos dados que serão segmentados na opção de bloco zero, tal como descrita em 3.4.3.2.

4.4 MAPEADOR DE PREDIÇÃO DE ERRO

O mapeador de Previsão de Erro toma os valores do erro previstos e os mapeia em número inteiros não negativos adequados ao Codificador com Entropia Adaptativa. A previsão de erro Δ_i que resulta da diferença entre um valor, x_i e um valor de predição \hat{x}_i sendo, ambos, números inteiros de n -bits, terão uma excursão de $(n+1)$ -bits dentro do intervalo dinâmico existente entre os extremos $[-2^{n+1}+1, 2^n-1]$. No entanto, para todo valor de predição existem apenas 2^n valores de erros de predição possíveis. Com a escolha apropriada do preditor, o valor mais provável do

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

erro de predição é zero, seguido por +1 e -1, +2 e -2, ..., etc. O menor valor de erro de predição é a diferença entre o valor mínimo de sinal, x_{min} , e o valor de predição, \hat{x}_i , igual a $x_{min} - \hat{x}_i$. O maior valor do erro de predição é aquele dado pela diferença entre o valor máximo, x_{max} , e o valor de predição, \hat{x}_i , igual a $x_{max} - \hat{x}_i$. Para mapear os 2^n possíveis valores de erro previstos em inteiros não negativos, a seguinte equação é utilizada:

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

onde

$$\theta i = \text{minimum} (\hat{x}_i - x_{min}, x_{max} - \hat{x}_i);$$

e para o valor de sinal com n -bits e com sinal aritmético (+ ou -),

$$x_{min} = -2^{n-1}, \quad x_{max} = 2^{n-1} - 1;$$

para valores de sinal com n -bits, não negativos,

$$x_{min} = 0, \quad x_{max} = 2^n - 1.$$

5 FORMATO DOS DADOS

5.1 ESTRUTURAS DE DADOS LIVRE DE PERDAS

5.1.1 GERAL

É necessária a especificação de vários parâmetros para a transferência sem perda de dados, codificados em comprimentos variáveis, de forma adaptativa, entre o codificador e o formatador de pacotes do canal de telemetria.

5.1.2 IDENTIFICAÇÃO DA OPÇÃO

5.1.2.1 O Campo ID especifica qual das opções foi utilizada para o conjunto de amostras a ele pertinente. As diversas chaves de código-ID para cada uma das opções estão apresentadas na tabela 5.1.

Tabela 5-1: Identificação da Chave de Opção do Código Selecionado

Opção	$n \leq 8$	$8 < n \leq 16$	$16 < n$
Bloco-Zero	0000	00000	000000
Segunda Extensão	0001	00001	000001
FS	001	0001	00001
$k = 1$	010	0010	00010
$k = 2$	011	0011	00011
$k = 3$	100	0100	00100
$k = 4$	101	0101	00101
$k = 5$	110	0110	00110
$k = 6$	—	0111	00111
$k = 7$	—	1000	01000
$k = 8$	—	1001	01001
$k = 9$	—	1010	01010
$k = 10$	—	1011	01011
$k = 11$	—	1100	01100
$k = 12$	—	1101	01101
$k = 13$	—	1110	01110
$k = 14$	—	—	01111
$k = 15$	—	—	10000
.	.	.	.
.	.	.	.
.	.	.	.
k=29	—	—	11110
sem compressão	111	1111	11111
NOTA – ‘—’ indica que o valor não se aplica, no caso.			

5.1.2.2 Para aplicações que não exigem desempenho pleno no limite de entropia oferecido pelas opções específicas, um subconjunto de opções pode ser implementado, na fonte. A identificação - ID, da opção, é sempre exigida, mesmo que apenas um subconjunto de opções venha a ser utilizado.

5.1.3 AMOSTRA DE REFERÊNCIA

Quando o preprocessador está presente e a amostra de referência exigida, o primeiro Conjunto de Dados Codificados (CDS, ou 'Coded Data Set') do Campo de Dados do Pacote Fonte ('Source Packet Data Field') deve conter a amostra de referência. Referências devem, portanto, ser inseridas no Campo de Dados do Pacote Fonte ('Source Packet Data Field'), pelo menos, a cada 256 CDSes, tal como está especificado no próprio CDS.

5.1.4 FORMATO DO CONJUNTO DE DADOS CODIFICADOS

5.1.4.1 O formato do CDS está representado na figura 5.1, no caso de seleção da opção com amostras particionadas. A Figura 5.1a representa o caso em que ocorre uma amostra de referência; A figura 5.1b representa o formato quando não ocorre uma amostra de referência. O formato CDS possui a seguinte estrutura quando a opção de amostras divididas é selecionada: 1) A seqüência opcional de bits de ID (Identificação) seguida por uma amostra de referência de n -bits, 2) dados comprimidos, e 3) os k bits menos significativos de cada amostra, concatenados.

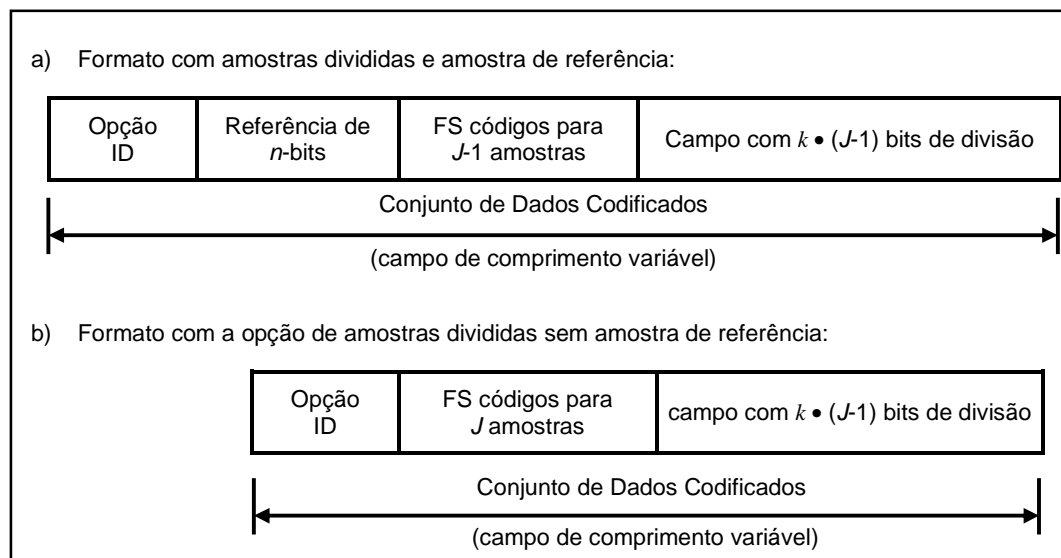


Figura 5-1: Formato do CDS Quando a Opção de Amostras Particionadas é Seleccionada

5.1.4.2 Quando a opção de não compressão é selecionada, o formato do CDS tem comprimento fixo e contém o campo opcional de ID, opcionalmente seguido de uma amostra de referência de n -bits, além de um número J de amostras pré-processadas. O caso em que uma referência está presente está ilustrado na figura 5-2a; o caso de não-referência está ilustrado na figura 5-2b.

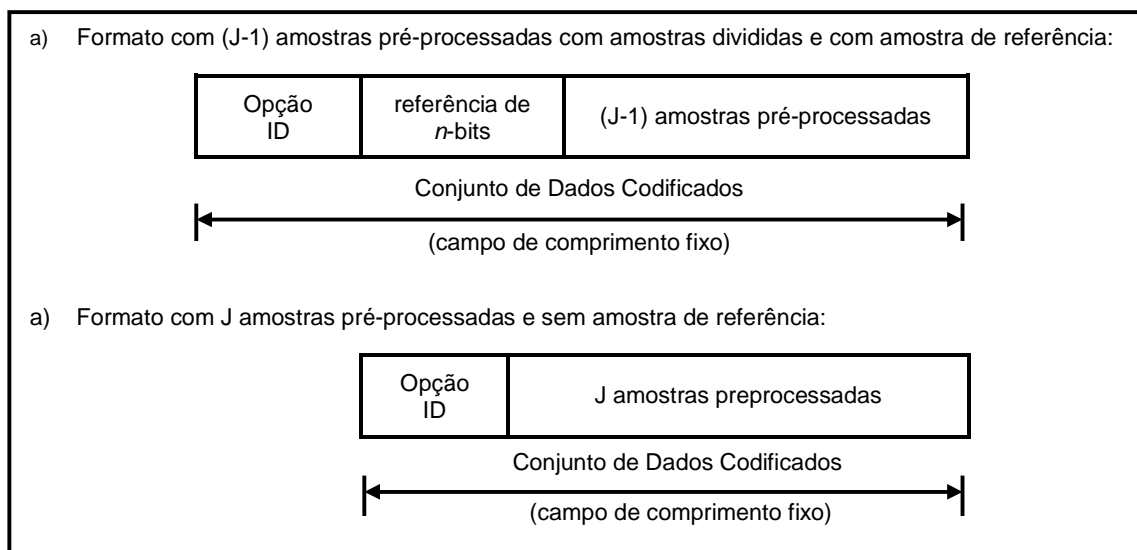


Figura 5-2: Formato CDS Quando a Opção de Não-Compressão é Seleccionada

5.1.4.3 Quando a opção de Bloco-Zero ('Zero-Block') é seleccionada, o formato CDS deve conter o campo opcional ID, opcionalmente seguido por uma amostra de referência de n -bits, e por uma palavra-código FS que é necessária para especificar o número de blocos concatenados com valor zero ou concatenados na forma de um Resto-de-Segmento (ROS, ou 'Remainder-of-Segment'), tal como descrito na seção 3.4.3, deste documento. O caso aonde uma (amostra de) referência é apresentada está ilustrado na figura 5-3a; o caso de não-referência está ilustrado na figura 5-3b.

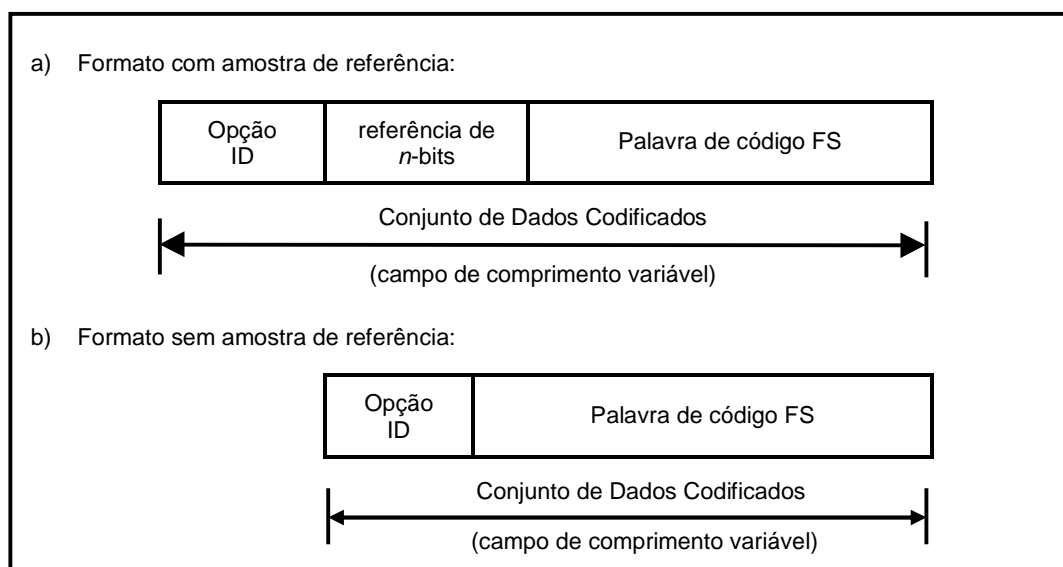


Figura 5-3: Formato CDS Quando a Opção Bloco-Zero é Seleccionada

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

5.1.4.4 Quando a opção de Segunda Extensão ('Second Extension') é selecionada, o formato CDS do campo opcional ID que é, opcionalmente, seguido por uma amostra de referência de n -bits e pelas palavras-código FS, exigidas para as $J/2$ amostras transformadas. O caso em que uma referência ocorre está ilustrado na figura 5-4a; o caso sem-referência está ilustrado na figura 5-4b. No caso em que a inserção de referência ocorre, uma amostra '0' é adicionada em frente das $J-1$ amostras pré-processadas, de tal forma que $J/2$ amostras são produzidas após a transformação.

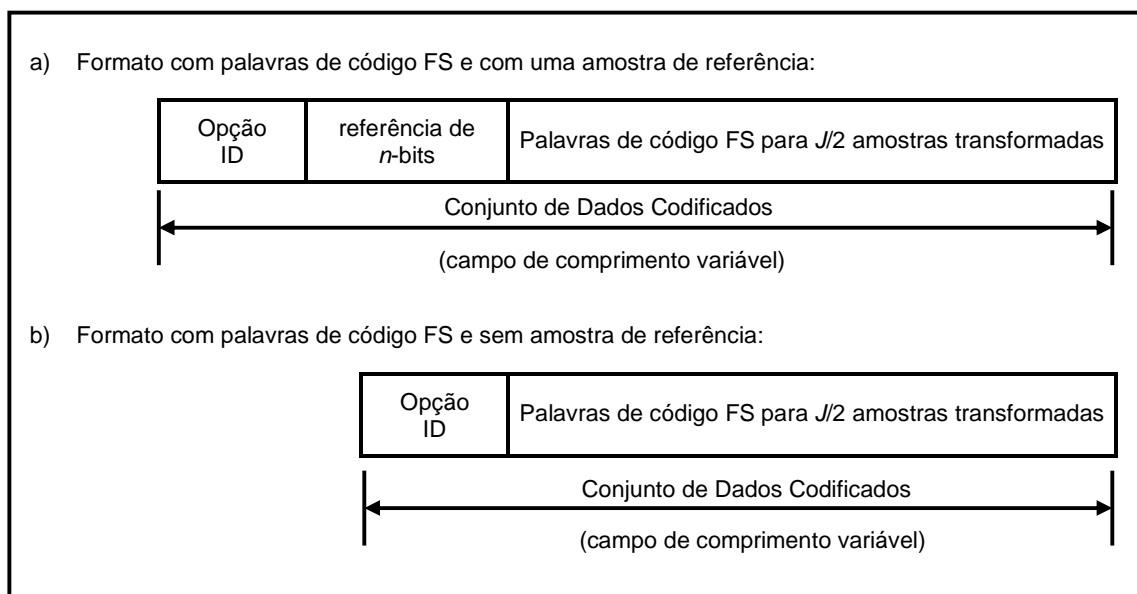


Figura 5-4: O Formato CDS Quando a Opção da Segunda-Extensão é Selecionada

5.2 FORMATO DO PACOTE

5.2.1 FORMATO DO PACOTE SEM PERDA

Pacotes contendo dados comprimidos sem perda estão ilustrados na figura 5-5 (vide referência [2]). O formatador de pacotes utiliza o parâmetro fornecido pelo codificador de dados fonte para formar um ou mais grupos de dados com formato CDS, o que determina o tamanho do pacote em octetos. Bits de enchimento com o valor zero podem ser necessários para forçar a terminação do pacote dentro do limite de um octeto.

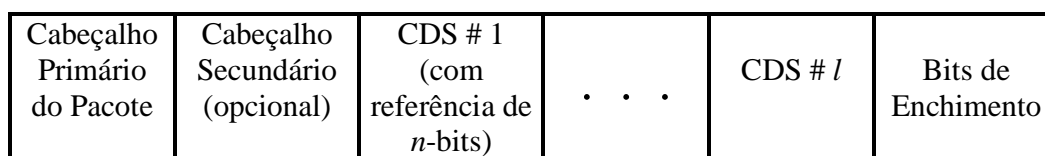


Figura 5-5: Formato de Pacote para l grupos de dados em formato CDS

5.2.2 REQUISITOS DOS PACOTES

5.2.2.1 O Campo de Dados de Pacote Fonte deve observar os seguintes requisitos:

- O formato CDS de um pacote precisa observar os quesitos definidos na Subseção 5.1;
- Quando uma amostra de referência é utilizada, o Campo de Dados de Pacote Fonte deve ser iniciado pelo formato CDS que contém esta referência, seguido por um ou mais formatos do tipo CDSes; Quando a amostra de referência não é exigida pelo preprocessador, ou porque o preprocessador não é utilizado, a amostra de referência não deve ser inserida no primeiro formato CDS, no Campo de Dados de Pacote Fonte;
- Vários formatos CDS podem ser inseridos na seqüência, dentro de um pacote fonte;
- Bits de preenchimento são permitidos somente no final do Campo de Dados de Pacote Fonte e, não, dentro do corpo dos dados comprimidos;
- Cada pacote precisa terminar dentro do limite de um octeto.

NOTA - Algumas implementações podem exigir a adição de um número adequado de bits de preenchimento com o propósito de terminar o pacote delimitado por um número par de octetos.

5.2.2.2 A menos que a opção de uso do Pacote de Identificação de Compressão - CIP ('Compression Identification Packet') seja escolhida (vide seção 6), com o propósito de decodificar os pacotes que podem incluir os bits de preenchimento, diversas partes de informação precisam ser dadas ao decodificador, 'a priori'. Esta informação deverá ser específica para a missão e deve ser pré-fixada para um dado Identificador de Processo de Aplicação (APID, ou 'Application Process Identifier'), para cada missão, nos seguintes termos:

- l , o número de formatos CDS existentes em um pacote;
- r , o intervalo da amostra de referência, igual ao número de formatos CDS, contados a partir de um CDS que contém a amostra de referência até, mas sem incluir, os CDSs consecutivos que contêm a amostra de referência;
- n , a resolução;
- J , o número de amostras, por bloco.

5.2.2.3 O Cabeçalho de Pacote Secundário é opcional e pode ser utilizado, por exemplo, para relacionar o momento de observação e a posição da informação para o usuário (veja a referência [2]).

5.2.2.4 O uso de Sinalizações de Seqüência ('Sequence Flags') no Campo de Controle de Sequenciamento de Pacote é opcional e pode ser utilizado, por exemplo, para sinalizar a existência de um grupo de pacotes de dados comprimidos. O uso do Agrupamento de Sinalizações está descrito na referência [2].

6 PACOTE DE IDENTIFICAÇÃO DE COMPRESSÃO (OPCIONAL)

6.1 ESTRUTURA DO PACOTE DE IDENTIFICAÇÃO DE COMPRESSÃO

6.1.1 Quando os dados comprimidos são transmitidos na forma de pacotes fonte agrupados, o Pacote de Identificação de Compressão (CIP, ou ‘Compression Identification Packet’) é um pacote opcional que, se utilizado, deve preceder e prover a informação de configuração para um grupo de pacotes de dados comprimidos de aplicação. O CIP será transmitido a partir de um processo de aplicação que se encontra no espaço para um ou vários processos destinatários, em solo.

6.1.2 O CIP deve ser o primeiro pacote do grupo.

6.1.3 O CIP deve consistir dos dois campos maiores, contiguamente posicionados, na seguinte sequência: Cabeçalho Primário de Pacote (‘Packet Primary Header’) e o Campo de Dados de Pacote (‘Packet Data Field’). Vide Figura 6.1.

6.1.4 O CIP deve conter informação que possa permitir ao descompressor de dados se configurar automaticamente para que ele possa processar um grupo de pacotes de dados de aplicação comprimidos sem que haja a necessidade dele gerenciar este tipo de informação, ‘a-priori’. O CIP deve ser utilizado para que o descompressor possa se configurar automaticamente somente se, associado a ele, estiver também disponível um sistema confiável para a execução de transferência de arquivos.

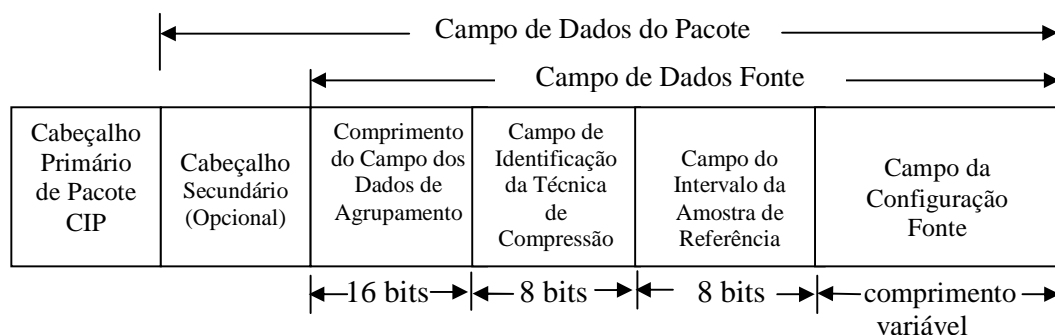


Figura 6-1: Estrutura do Pacote de Identificação de Compressão

6.2 CABEÇALHO PRIMÁRIO DO CIP

6.2.1 CONCEITO GERAL

O Cabeçalho Primário de Pacote é mandatório para o CIP e sua estrutura deve estar em conformidade com o Livro Azul do CCSDS em Protocolo de Pacote Espacial (‘CCSDS Space Packet Protocol’), de acordo com a referência [2]. O campo de Cabeçalho Primário de Pacote (‘Packet Primary Header’) do CIP deve conter os dados fonte APID (Identificador de Processo de Aplicação ou ‘Application Process Identifier’). O uso do CIP é específico para uma determinada missão e é prefixado para um dado APID.

6.2.2 SINALIZAÇÕES DE AGRUPAMENTO

6.2.2.1 As Sinalizações de Sequenciamento ('Grouping Flags') estão presentes no campo Seqüência de Controle ('Sequence Control') do pacote, tal como especificado na referência [2]. Este campo se situa no Cabeçalho Primário de Pacotes que encapsulam os dados comprimidos do usuário. Como indicado abaixo, este campo é sempre definido pelo valor '01' para o Cabeçalho Primário do CIP.

6.2.2.2 As Sinalizações de Seqüência devem obedecer às seguintes configurações:

- '01' para o primeiro pacote do grupo, que é o CIP;
- '00' para os pacotes fonte seguintes, que contém os dados comprimidos do grupo;
- '10' para o último pacote fonte que contém os dados comprimidos do grupo.

6.2.2.3 Para um pacote fonte não pertencente a um grupo de pacotes fonte com dados comprimidos, as Sinalizações de Seqüência devem observar o valor em '11'.

6.3 CAMPO DE DADOS DE PACOTE

6.3.1 CONCEITO GERAL

O Campo de Dados de Pacote de um CIP deve consistir de dois campos posicionados de forma justaposta, na seguinte seqüência: Cabeçalho Secundário de Pacote ('Packet Secondary Header'), que é opcional, e o Campo de Dados Fonte ('Source Data Field').

6.3.2 CABEÇALHO SECUNDÁRIO (OPCIONAL)

O Cabeçalho Secundário é um recurso para se poder incluir dados ancilares no CIP, tais como informação de tempo e de posição/atitude da espaçonave.

6.3.3 CAMPO DE DADOS FONTE

6.3.3.1 CONCEITO GERAL

O Campo de Dados Fonte ('Source Data Field') do CIP deve ser constituído por de quatro campos posicionados contiguamente, obedecendo à seguinte seqüência:

	<u>Comprimento (bits)</u>
- O Comprimento do Agrupamento de Dados	16
- Identificação da Técnica de Compressão	8
- Intervalo da Amostra de Referência	8
- Configuração Fonte	(Variável)

6.3.3.2 COMPRIMENTO DOS DADOS DE AGRUPAMENTO

O Comprimento dos Dados de Agrupamento é um campo de 16 bits dos quais os primeiros quatro bits são reservados. Os 12 bits restantes do campo devem conter um número binário igual ao número de pacotes que contem os dados comprimidos dentro do grupo, subtraído de um, sendo que o número de pacotes que contém os dados comprimidos deve estar em uma faixa de

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

valor entre 1 e 4096. O número de pacotes no grupo com o CIP incluído deve variar entre 2 e 4097.

6.3.3.3 CAMPO DE IDENTIFICAÇÃO DA TÉCNICA DE COMPRESSÃO

6.3.3.3.1 O campo de Identificação da Técnica de Compressão (CTI, ou ‘Compression Technique Identification’) deve indicar a técnica de compressão em uso para o grupo de pacotes fonte identificados pelo CIP.

6.3.3.3.2 Quando nenhuma técnica de compressão para o grupo de dados sendo processado está sendo utilizada, todas as posições do campo do CTI devem possuir o valor zero.

6.3.3.3.3 Somente a técnica de compressão de dados Livre de Perdas é definida atualmente, e é sinalizada com o valor 1 no campo do CTI. Os demais valores são reservados para uso futuro pelo CCSDS, e não podem ser utilizados, por este motivo.

6.3.3.4 CAMPO DE INTERVALO DA AMOSTRA DE REFERÊNCIA

6.3.3.4.1 O intervalo da amostra de referência, r , corresponde ao número de formatos CDSes, contados a partir de um CDS que contém uma amostra de referência até, mas não incluído, o CDS seguinte que contém uma amostra de referência. Quando o preprocessador não estiver sendo utilizado, ou quando uma amostra de referência não for necessária, a amostra de referência não deve ser inserida no CDS; apesar disso, o parâmetro r serve para definir o intervalo dos blocos de amostra dos dados de entrada para a opção de bloco-zero, tal como está descrito em 4.3.

6.3.3.4.2 O campo r de 8-bits deve conter um número binário igual a $r-1$, com o valor de r variando de 1 a 256.

6.3.3.5 CONFIGURAÇÃO FONTE

6.3.3.5.1 PARTIÇÕES DO SUBCAMPO

6.3.3.5.1.1 O campo de Configuração Fonte deve ser particionado em três subcampos: Pre-processador, Codificador de Entropia (‘Entropy Coder’) e Configuração de Instrumento (vide Figura 6.2). O uso dos subcampos do Pre-processador e do Codificador de Entropia são obrigatórios, considerando que o subcampo de Configuração do Instrumento (‘Instrument Configuration’) é opcional. Cada um desses subcampos de Configuração de Compressão (‘Compressor Configuration’) deve ter como cabeçalho os dois primeiros bits, para identificação do tipo do subcampo.

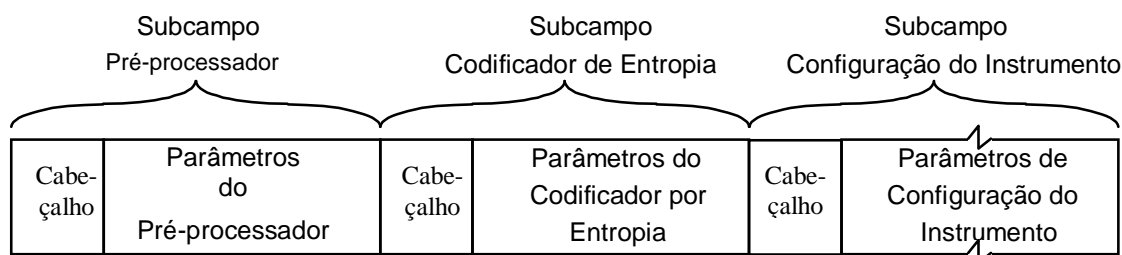


Figura 6-2: Campo da Configuração Fonte

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

6.3.3.5.1.2 Os sinalizadores do tipo de subcampo devem ser configurados como segue:

- 00 – Pré-processador
- 01 – Codificador de Entropia
- 10 – Configuração de Instrumento
- 11 – Reservados

6.3.3.5.2 PRÉ-PROCESSADOR

6.3.3.5.2.1 O comprimento do subcampo Pré-processador deve possuir dois octetos, dos quais, os dois primeiros bits devem servir como cabeçalho, tal como descrito na Subseção 6.3.3.5.1.2.

6.3.3.5.2.2 Os parâmetros do Pré-processador para o compressor de dados de Livre de Perdas devem ser particionados em seis áreas e devem se posicionar de forma justaposta, contígua, em seguida aos 2 bits do cabeçalho do Pré-processador. Ver 3.1 e 4 para a definição de parâmetros do Pré-processador. As seis áreas são as seguintes:

a) Estado do Pré-processador (1 bit):

- 0 – ausente
- 1 – presente

b) Tipo de Preditor (3 bits); ignora se o estado do Pré-processador está em ‘0’:

- 000 – evita o uso do preditor
- 001 – preditor com atraso unitário
- 111 – preditor para aplicação específica

Todos outros códigos são reservados pelo CCSDS para futuras opções de pré-processamento.

c) Mapeador (2 bits); ignora se o estado do Pré-processador é ‘0’:

- 00 – Mapeador de Predição de Erro, descrito em 4.4
- 01 – reservado
- 10 – reservado
- 11 – mapeador aplicação específica

d) Tamanho de bloco (J) (2 bits):

	<u>Número de</u> <u>Amostras/Bloco</u>
00	8
01	16
10	reservado
11	aplicação específica

e) Sentido dos dados (1 bit):

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

0 – complemento de dois

1 – positivo; obrigatório se o Pré-processador não está sendo utilizado

f) Resolução (n) da amostra de dados de entrada (5 bits):

Os 5 bits do campo da Amostra de Dados de Entrada devem conter um número binário igual à resolução da amostra de dados de entrada subtraída de um, supondo que a resolução da amostra de dados de entrada pode variar de 1 a 32.

6.3.3.5.3 CODIFICADOR DE ENTROPIA

6.3.3.5.3.1 O comprimento do subcampo do Codificador de Entropia deve ser de dois octetos, sendo que os seus dois primeiros bits devem servir de cabeçalho, tal como descrito na Subseção 6.3.3.5.1.2.

6.3.3.5.3.2 Os parâmetros do subcampo do Codificador de Entropia devem ser particionados em duas áreas e devem ser posicionados de forma contígua, seguindo os dois bits do cabeçalho do Código de Entropia. As duas áreas são as seguintes:

a) Opção do Codificador de Entropia (2 bits):

Um máximo de 32 opções são permitidas (ver Tabela 5.1 para a identificação da opção de chaves):

00 — Reservado

01 — para $n \leq 8$

10 — para $8 < n \leq 16$

11 — para $16 < n \leq 32$

b) Número de CDSes por pacote, l (12 bits):

O campo de 12 bits que indica qual é o número de CDSes por pacote (l) que devem conter um número binário igual a $l - 1$.

6.3.3.5.4 CONFIGURAÇÃO DO INSTRUMENTO

O Subcampo de Configuração de Instrumento (CIS) é um campo unívoco ao instrumento, que é utilizado para endereçamento dos parâmetros de configuração unívocos ao instrumento. Os conteúdos deste campo são específicos a uma missão e estão além do escopo desta Recomendação. Se utilizados, os primeiros dois bits devem expressar qual é o cabeçalho, tal como especificado em 6.3.3.5.1.2.

ANEXO A

GLOSSÁRIO DE ACRÔNIMOS E TERMOS

(Este anexo **não** é parte da Recomendação.)

A1 PROPOSTA

Este anexo define os acrônimos chaves e termos que são usados nesta Recomendação para descrever a codificação original para compressão de dados.

A2 ACRÔNIMOS

<u>Acrônimos</u>	<u>Definição</u>
CDS	conjunto de dados codificados (“coded data set”)
CIP	pacote de identificação de compressão (“compression identification packet”)
CTI	identificação da técnica de compressão (“compression technique identification”)
FS	seqüência fundamental (“fundamental sequence”)
ICS	subcampo de configuração de instrumento (“instrument configuration subfield”)
LSB	bit menos significativo (“least significant bit”)
MSB	bit mais significativo (“most significant bit”)
ROS	resto do segmento (“remainder of segment”)

A3 TERMOS

CODIFICADOR COM ENTROPIA ADAPTATIVA: Um codificador de entropia codifica as amostras fonte com palavras-código decodificáveis de forma única que, após a decodificação, é capaz de reconstruir as amostras fonte. Com um Codificador por Entropia Adaptativa, o tamanho médio da palavra-código também segue de perto as informações contidas na fonte.

ENTROPIA: Entropia é uma medida quantitativa da quantidade média de informação por amostra original, expressa em bits/amostra.

SEQÜÊNCIA FUNDAMENTAL: O código Seqüência Fundamental (FS) representa o inteiro não negativo m com uma palavra de código binário de m zeros seguido por um dígito um. Aplicações do código FS a um bloco de J amostras produz uma seqüência de J palavras-código concatenadas denominadas por Seqüência Fundamental.

RECOMENDAÇÃO CCSDS PARA COMPRESSÃO DE DADOS LIVRE DE PERDAS

CODIGO ADAPTATIVO DE RICE: O algoritmo básico de codificação adaptativa de Rice escolhe o melhor dentre as várias opções de código para uso em um bloco de dados. Estas opções são destinadas para atuarem com eficiência em diferentes intervalos dos dados ativos. As opções são implementadas usando uma combinação de códigos FS e o particionamento de amostras pré-processadas em suas partes mais significativas e menos significativas de bits.

PARTICIONAMENTO DE AMOSTRA: Particionamento de amostra é um procedimento que separa a representação binária de uma amostra em dois grupos de bits adjacentes, um para a ordem inferior de bits e outro para a ordem superior de bits.

BITS DE PARTIÇÃO:

Bits de partição são aqueles de ordem mais baixa separados por particionamento de amostra, obtido a partir de uma representação binária da mesma.

ANEXO B

REFERÊNCIAS INFORMATIVAS

(Este anexo **não é** parte da Recomendação.)

- [B1] *Procedures Manual for the Consultative Committee for Space Data Systems*. CCSDS A00.0-Y-9. Livro Amarelo. Versão 9. Washington, D.C.: CCSDS, Novembro de 2003.
- [B2] *Lossless Data Compression*. Relatório Afeito a Padrões em Sistemas Espaciais de Dados, CCSDS 120.0-G-2. Livro Verde. Versão 2. Washington, D.C.: CCSDS, Dezembro de 2006.

NOTA: A última versão dos documentos CCSDS pode ser obtida no site www.ccsds.org.