

Universidade de Brasília - UnB
Faculdade UnB Gama - FGa
Engenharia Eletrônica

Implementação de Código Compressor de Vídeos Aplicado a Cenários de Comunicação TTE

Autor: Luíza Carneiro Cezário
Orientador: Prof. Dr. Leonardo Aguayo
Coorientador: Prof. Dr. Marcus Víncius Chaffim

Brasília, DF
2019



Luíza Carneiro Cezário

Implementação de Código Compressor de Vídeos Aplicado a Cenários de Comunicação TTE

Monografia submetida ao curso de graduação em (Engenharia Eletrônica) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia Eletrônica).

Universidade de Brasília - UnB

Faculdade UnB Gama - FGa

Orientador: Prof. Dr. Leonardo Aguayo

Coorientador: Prof. Dr. Marcus Vínicius Chaffim

Brasília, DF

2019

Luíza Carneiro Cezário Implementação de Código Compressor de Vídeos
Aplicado a Cenários de Comunicação TTE/ Luíza Carneiro Cezário. – Brasília,
DF, 2019- 66 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Leonardo Aguayo

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB

Faculdade UnB Gama - FGa , 2019.

1. . 2. . I. Prof. Dr. Leonardo Aguayo. II. Universidade de Brasília. III. Faculdade
UnB Gama. IV. Implementação de Código Compressor de Vídeos Aplicado a
Cenários de Comunicação TTE

CDU

Luíza Carneiro Cezário

Implementação de Código Compressor de Vídeos Aplicado a Cenários de Comunicação TTE

Monografia submetida ao curso de graduação em (Engenharia Eletrônica) da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em (Engenharia Eletrônica).

Trabalho aprovado. Brasília, DF, :

Prof. Dr. Leonardo Aguayo
Orientador

Prof. Dr. Diogo Caetano Garcia
Convidado 1

**Prof. Dr. Wellington Avelino do
Amaral**
Convidado 2

Brasília, DF
2019

*Este trabalho é dedicado a todos que foram importantes durante a realização da minha
graduação.*

Agradecimentos

Gostaria de agradecer à Universidade de Brasília pelo suporte oferecido para que pudesse me desenvolver profissionalmente e aos docentes da Universidade por todo o aprendizado e apoio que me proporcionaram.

Gostaria de agradecer ao Prof. Dr. Leonardo Aguayo e ao Prof. Dr. Prof. Dr. Marcus Vinícius Chaffim, cujas orientações e suporte foram essenciais para o desenvolvimento deste trabalho.

E por fim gostaria de agradecer aos amigos da FGA que me ajudaram durante a graduação e me proporcionaram momentos felizes na faculdade e a minha família pelo suporte durante todo o curso.

Resumo

A mineração é a atividade desempenhada para extrair minérios do solo. Em alguns casos os minérios se encontram em camadas profundas do solo, com isso são construídos túneis e galerias subterrâneas para possibilitar a exploração. O local onde essa exploração ocorre é chamado de mina subterrânea, ambiente muito propenso a ocorrência de acidentes e por essa razão é necessário um sistema de comunicação que possa ser usado em situações de emergência. Um dos sistemas de comunicação adequados para esse cenário é o sistema *Through-the-Earth* (TTE), adequado por ser resistente a explosões e acidentes e por ser capaz de transmitir informações nesse ambiente. A comunicação TTE é feita por indução magnética, utilizando frequências nas faixas *Ultra Low Frequency* (ULF) e *Very Low Frequency* (VLF). O canal TTE é um dos maiores desafios para a comunicação, sendo um canal passa-faixas com forte interferência de ruído e a banda disponível para transmissão é estreita. A banda do canal restringe também a taxa de dados da informação. Imagens e vídeos tem uma taxa de dados muito alta para esse tipo de cenário e para possibilitar a transmissão dessas informações é necessário o uso de técnicas de compressão. Nesse trabalho é proposto um codificador Wyner-Ziv de imagens e vídeos para aplicação no cenário TTE, pela característica de transferir a complexidade do codificador para o decodificador e por ter maior robustez a ruídos.

Palavras-chaves: Comunicação TTE. Compressão de Vídeos. Wyner-Ziv.

Abstract

Mining is the activity performed to extract minerals from the ground. Sometimes these minerals are found in deep layers in the ground, so that underground tunnels and galleries are built to enable exploration. The site where this exploration takes place is called underground mine, and they are very prone to accidents and therefore a communication system that can be used in emergency situations is required. One of the appropriate communication systems for this scenario is the Through-the-Earth (TTE) system, because it is resistant to explosions and accidents and is capable of transmitting information in this environment. TTE communication is made by magnetic induction, using frequencies in the Ultra Low Frequency (ULF) and Very Low Frequency (VLF) bands. The TTE channel is one of the major challenges for communication, it is a pass-band channel with strong noise interference and narrow bandwidth available for transmission. The channel band also restricts the data rate of information, images and videos have a very high data rate, requiring compression techniques. In this work, a Wyner-Ziv video encoder is proposed for application in the TTE scenario, due to its characteristic of transferring the complexity of the encoder to the decoder.

Key-words: TTE Communication. Video Compression. Wyner-Ziv.

Lista de ilustrações

Figura 1 – Representação de sistema TTE com antenas loop.	25
Figura 2 – Função de transferência do canal TTE quando $\sigma = 0,01S/m$ e $r = 200$ m.	29
Figura 3 – Construção do código de Huffman	35
Figura 4 – Procedimento para gerar a tag para a sequência ala2a3.	36
Figura 5 – Algoritmo básico de codificação diferencial.	38
Figura 6 – Bases da DCT.	39
Figura 7 – Matrizes da DCT.	40
Figura 8 – Teorema de Slepian-Wolf: Região onde a decodificação é perfeita. . . .	43
Figura 9 – Diagrama de Blocos do Algoritmo	46
Figura 10 – Matriz esparsa de um bloco linear	50
Figura 11 – Quadro original	50
Figura 12 – Informação lateral da sequência Akiyo, PSNR = 54.48	51
Figura 13 – Informação lateral da sequência Foreman, PSNR = 31.34 dB	52
Figura 14 – Quadro original	56
Figura 15 – Quadro chave decodificado sem a presença do ruído TTE, PSNR = 48.04 dB	56
Figura 16 – Quadro chave decodificado com a presença do ruído TTE, PSNR = 34.87 dB	56
Figura 17 – Quadro de Wyner-Ziv decodificado com a presença do ruído TTE, PSNR = 21.20 dB	57
Figura 18 – Quadro de Wyner-Ziv decodificado sem a presença do ruído TTE, PSNR = 30.97 dB	57
Figura 19 – Resíduo somado ao frame anterior após a aplicação da DCT, PSNR = 27.67 dB	58
Figura 20 – Imagem 1 filtrada por um filtro de média, PSNR = 32.78 dB	58
Figura 21 – Frame chave anterior decodificado, PSNR = 31.60 dB	58
Figura 22 – Frame chave posterior decodificado, PSNR = 31.57 dB	59
Figura 23 – Informação lateral, PSNR = 31.34 dB	59
Figura 24 – Imagem reconstruída, PSNR = 33.10 dB	59
Figura 25 – Quadro original	60
Figura 26 – (a) Taxa x PSNR - Akiyo, (b) Taxa x PSNR - Foreman	61
Figura 27 – (a) Taxa x PSNR - Salesman, (b) Taxa x PSNR - Grandma	62
Figura 28 – (a) Taxa x PSNR - Mobile, (b) Taxa x PSNR - Carphone	62

Lista de símbolos

σ	Condutividade do meio
ϵ	Permissividade do meio
μ	Permeabilidade do meio
δ	Profundidade Peculiar
v	Velocidade de Propagação
λ	Comprimento de Onda
ω	Frequência Angular
\hat{r}	Vetor na direção radial
$\hat{\theta}$	Vetor na direção elevacional
Ξ	Componente gaussiana do ruído atmosférico
Υ	Componente impulsiva do ruído atmosférico
Γ	Função Gamma
γ	Impulsividade

Sumário

1	INTRODUÇÃO	19
	Introdução	19
1.1	Aspectos Gerais	19
1.2	Objetivos	21
1.3	Metodologia	21
2	COMUNICAÇÕES EM MINAS SUBTERRÂNEAS	23
	Contextualização e análise teórica	23
2.1	Sistemas de comunicação <i>Through-the-Wire (TTW)</i>	24
2.2	Sistemas de comunicação <i>Through-the-Air (TTA)</i>	24
2.3	Sistemas de comunicação <i>Through-the-Earth (TTE)</i>	25
2.3.1	Propagação de Onda em Ambiente TTE	27
2.3.2	Modelo de Campo Magnético	28
2.3.3	Ruído em Ambiente TTE	29
3	CONCEITOS DE COMPRESSÃO DE SINAIS, IMAGENS E VÍDEOS	31
	Contextualização e análise teórica	31
3.1	Entropia	31
3.2	Codificadores Entrópicos	32
3.2.1	Códigos de Huffman	32
3.2.2	Código Aritmético	35
3.3	Codificadores com Perdas	36
3.3.1	Codificação Diferencial	36
3.3.2	Codificação por Transformadas	38
3.3.2.1	Transformada de Karhunen–Loève (KLT)	38
3.3.2.2	Transformada Discreta de Cosseno (DCT)	39
3.3.2.3	Transformada Discreta de Seno (DST)	40
3.3.2.4	Transformada Discreta de Walsh-Hadamard (DWHT)	40
3.3.3	Métricas de Qualidade para Compressão com Perdas	42
3.4	Teorema de Wyner-Ziv	42
4	METODOLOGIA	45
	Metodologia	45

4.1	Algoritmo Proposto	45
4.1.1	Funcionamento do Algoritmo	45
4.2	Materiais	46
4.2.1	Software usado no codificador e no decodificador	46
4.3	Etapas da Implementação	47
4.3.1	Separação em Grupos de Quadro	48
4.3.2	Aplicação da DCT	48
4.3.3	Quantização	48
4.3.4	Formação de Palavras Código	48
4.3.5	Codificador LDPCA e Decodificador LDPCA	49
4.3.6	Log de Verossimilhança	51
4.3.7	Geração da informação lateral	51
4.3.8	Codificador <i>intraframe</i> e decodificador <i>intraframe</i>	52
4.3.9	Aplicação da DCT inversa	52
4.3.10	Reconstrução	52
5	RESULTADOS	55
Resultados		55
5.1	Interferência do Ruído TTE nas Imagens	55
5.2	Tratamento das Informações para Reconstrução da Imagem	57
5.3	Taxas de Compressão	60
5.4	Curvas de Taxa x PSNR	61
6	CONCLUSÃO	63
Conclusão		63
	REFERÊNCIAS	65

1 Introdução

1.1 Aspectos Gerais

A mineração é uma atividade de extrema importância na sociedade moderna. Com ela são obtidos insumos usados na indústria, como alumínio, silício, carvão, dentre outros. Esses insumos são usados na produção de equipamentos eletrônicos, no setor automobilístico e diversos outros setores que movimentam a economia global. As minas de exploração mineral podem ser classificadas de duas maneiras: exploração a céu aberto e exploração subterrânea. As que são de exploração a céu aberto são minas onde as camadas superficiais do solo são retiradas para localizar e extrair os minérios, já as minas subterrâneas são minas onde os minérios se encontram em camadas mais profundas do solo, em algumas minas essa profundidade pode chegar a centenas de metros. Esse trabalho tem como foco o cenário de minas subterrâneas (1).

Nas minas subterrâneas a exploração é mais perigosa e mais complexa comparada a exploração nas minas a céu aberto, devido a isso ocorre um grande número de acidentes e mortes nesse tipo de mina. Existem diversos fatores que podem causar acidentes em minas subterrâneas, podendo estar presentes gases tóxicos, poeira e elementos corrosivos. A própria utilização de explosivos para facilitar o acesso ao minério é um fator de risco que pode levar ao desabamento da mina. Fatores climáticos também podem causar acidentes, como inundações da mina em casos de chuva e desabamentos em casos de abalos sísmicos (1).

Em caso de desabamento da mina o resgate dos trabalhadores além de ser demorado é arriscado. Para que a operação seja realizada com sucesso e as vítimas possam sair com vida, as informações do ambiente do acidente são essenciais. Através do conhecimento do cenário do desabamento a equipe de resgate pode definir um plano de ação consistente. Com isso a comunicação da superfície com as galerias subterrâneas é a principal fonte de informação pra equipe de resgate (2).

Para a comunicação em minas terrestres são utilizadas as seguintes técnicas: *Through-the-Earth* (TTE), *Through-the-Wire* (TTW) e *Through-the-Air* (TTA). A comunicação *Through-the-Wire* é feita através de fios e fornecem comunicação entre as galerias e entre o subsolo e a superfície. A *Through-the-Air* é feita através do ar, é um modelo bastante suscetível às mudanças no ambiente. A comunicação *Through-the-Earth* é feita através do solo, é a comunicação foco deste trabalho, a comunicação TTE é feita através de indução magnética utilizando geralmente antenas em formato de *loop* (3).

As pesquisas sobre comunicação TTE se intensificaram após acidentes ocorridos

em minas de carvão nos Estados Unidos, levantando a discussão sobre segurança em minas subterrâneas. A partir disso foi implementado em 2006 nos Estados Unidos o *Mine Improvement and New Emergency Response Act* (MINER Act). Uma de suas definições é o estabelecimento de um sistema de comunicação que possa ser usado em caso de emergência. A comunicação TTE é menos suscetível às mudanças no ambiente do que a comunicação TTA e é mais resistente a acidentes com relação a TTW, que por ser feita através de fios pode ser rompida em caso de desabamentos e explosões. Por essas características a comunicação TTE é ideal para uso em situações de emergência (3).

No mercado alguns equipamentos utilizam a tecnologia TTE para o uso em minas, onde a comunicação geralmente é feita através de mensagens de texto, que podem ser representadas em menor taxa comparadas a outros tipos de dados. Alguns equipamentos conseguem fazer a comunicação através de voz e imagem (4).

Para que a comunicação seja feita através de voz ou mensagem de texto é necessário que um usuário inicie essa comunicação, então é preciso que as vítimas estejam acordadas e em condições de se comunicar com a superfície, o que nem sempre é o que acontece. Já no caso da imagem o próprio equipamento pode fotografar ou filmar o local e transmitir para superfície, facilitando assim o trabalho do resgate. Outra vantagem da transmissão de imagens comparada a outros tipos de dados é que a informação que a imagem apresenta é mais confiável e mais completa.

A dificuldade de transmissão de imagens com relação a voz e texto é que o conjunto de dados exige uma taxa de transmissão maior, e a taxa de transmissão é um dos fatores limitantes do canal TTE. Quanto maior é a taxa, mais estreita é a banda disponível para transmissão. Dessa maneira para que se possa transmitir imagens através do canal TTE é necessário comprimir a imagem. A compressão de dados se trata de retirar as redundâncias do conjunto de dados de forma a diminuir a quantidade de bytes enviados.

Para comprimir a imagem é necessário um codificador que atendas as necessidades do canal e da própria imagem, além disso é necessário no receptor que se encontra na superfície um decodificador para que a informação contida na imagem possa ser recuperada. Em sistemas de transmissão são utilizados codificadores de fonte, que reduzem a representação em bits e comprimem a informação, ou codificadores de canal que aumentam a representação do sinal em bits adicionando redundância. Os codificadores de canal são utilizados geralmente quando o canal é muito suscetível a ruídos. O codificador de vídeo escolhido foi o de Wyner-Ziv foi escolhido por combinar os dois tipos de codificação e reduzir a representação final. O objetivo do trabalho é desenvolver um algoritmo de compressão de vídeos, com uso aplicado a situações de resgate de vítimas de acidentes em minas subterrâneas.

1.2 Objetivos

Objetivo geral: Implementar e simular através do software MATLAB o codificador e decodificador de Wyner-Ziv aplicado ao canal de comunicação TTE.

Objetivos específicos:

- Estudar os parâmetros e restrições do canal TTE e a influência do ambiente na transmissão e recepção de dados.
- Estudar as técnicas de codificação de imagens existentes, apresentar suas características e vantagens.
- Implementar os códigos do codificador e decodificador Wyner-Ziv
- Simular o funcionamento do algoritmo em ambiente TTE utilizando o simulador de canal.

1.3 Metodologia

Abaixo estão listados os passos da metodologia:

- Estudos teóricos e pesquisas bibliográficas sobre o assunto para definição do escopo do trabalho;
- Definição do diagrama de blocos do sistema;
- Definição dos passos de implementação e da estratégia de impletação de cada passo;
- Implementação do algoritmo;
- Testes do algoritmo implementado no simulador de canal

Este documento está organizado da seguinte maneira:

- O Capítulo 1: introduz o trabalho a ser desenvolvido, apresentando a contextualização, motivação, objetivos e metodologia do trabalho.
- O Capítulo 2: aborda a tecnologia de comunicação TTE e suas características.
- O Capítulo 3: aborda técnicas de compressão de imagens e codificação de dados.
- O Capítulo 4: aborda as etapas de implementação do algoritmo e técnicas utilizadas.
- O Capítulo 5: apresenta os resultados da implementação do algoritmo.
- O Capítulo 6: apresenta a conclusão do trabalho e considerações finais.

2 Comunicações em Minas Subterrâneas

A atividade de mineração é uma das mais antigas da humanidade, extraindo recursos usados em diversos setores da indústria, e gerando assim alto impacto na economia global. As minas podem ser classificadas em dois tipos de acordo com a exploração: a céu aberto ou subterrâneas. Nas minas a céu aberto o trabalho de extração dos minerais é feito inteiramente na superfície, enquanto nas minas subterrâneas são feitos túneis e galerias através do solo para extrair os minerais. As minas subterrâneas apresentam condições de trabalho e exploração mais hostis comparadas as minas a céu aberto, geralmente sendo muito úmidas e apresentando substâncias como poeira, gases tóxicos, explosivos, substâncias corrosivas, que afetam a saúde dos mineiros e o funcionamento das máquinas e aparelhos. Nesses ambientes os equipamentos de comunicação podem assumir diversas finalidades: resgate em caso de acidentes, automação de máquinas, acionamento de explosivos, comunicação entre as galerias e a superfície ou entre galerias, monitoramento da mina, dentre outros. As tecnologias de comunicação nesses ambientes podem ser sem fio, que é o caso das tecnologias TTE (*Through the Earth*) e TTA (*Through the Air*), e podem ser cabeadas, como a TTW (*Through the Wire*) (2).

Um fator importante a considerar é que o ambiente da mina subterrânea é dinâmico, devido a expansão da área de acordo com a extração dos minerais, expandindo consequentemente a área que o sistema de comunicação deve cobrir. Esse fator é um desafio quando comunicações por rádio são consideradas, os sinais de rádio são propensos às mudanças no ambiente físico, o que acarreta na mudança do comportamento de propagação das ondas de rádio. Características individuais de cada mina também influenciam nos sistemas de comunicação, como o mineral a ser explorado, o método de escavação utilizado e a estrutura do solo da mina. O mineral explorado e o método de escavação utilizados são fatores importantes nos sistemas de comunicação por ondas de rádio, já que as características eletromagnéticas de cada mineral, como condutividade e dieletricidade influenciam a comunicação. O método de escavação vai influenciar no comportamento da propagação das ondas nos ambientes da mina. Em sistemas de comunicação TTE a característica individual mais importante a ser considerada é a estrutura do solo da mina, que é o canal por onde é feita a comunicação, já que cada solo tem suas características, o alcance da transmissão e largura disponível para transmissão será diferente para cada caso (2).

2.1 Sistemas de comunicação *Through-the-Wire* (TTW)

As tecnologias TTW são utilizadas para comunicação interna nos ambientes do subsolo e também entre o subsolo e a superfície, alguns exemplos que utilizam essa tecnologia cabeada são os telefones de magneto, os *leaky feeders* e as fibras óticas. As comunicações através de fio podem ocorrer com apenas um fio conectando todos os terminais de comunicação, dessa maneira todos os sistemas conectados ao fio podem receber a informação transmitida por um dispositivo, esse sistema é chamado de *party line*. A comunicação também pode ser feita de forma privada com um sistema conhecido como *private line*, esse sistema permite com que dois dispositivos se comuniquem um com outro sem que os demais dispositivos recebam a informação, a comutação dessas linhas é processada por um computador (2).

Os telefones de magneto funcionam gerando uma corrente capaz de fazer com que os sinos dos outros telefones em uma linha privada toquem, a identificação do receptor é feita através de uma codificação usando toques curtos e longos. Quando a identificação for feita e outro dispositivo receber a ligação, a conversa passa a ser mantida pela energia da bateria. Outra tecnologia usada para comunicação em minas são os telefones alimentados por voz, esses telefones convertem a voz dos usuários em sinais elétricos usando transdutores eletromecânicos. Sua principal vantagem é que não precisam de energia externa para funcionar, eles usam a pressão do ar que a voz fornece como fonte de energia para o sistema. Eles conectam dois dispositivos através de um fio e transmitem a informação. A principal desvantagem desse sistema é seu alcance limitado (2).

Outro sistema conhecido são os *paging phones*, cada telefone tem uma bateria individual que fornece a energia necessária para que os amplificadores de áudio fortaleçam o sinal que será transmitido nos auto falantes. O uso da bateria individual faz com que o sinal de áudio tenha um nível alto, o que é uma vantagem em ambientes como minas subterrâneas, além disso são fáceis de instalar e são confiáveis. Os *paging phones* conseguem realizar a comunicação de forma individual, para identificar a pessoa de destino, o usuário anuncia o nome da pessoa com quem deseja falar e a ligação é feita (2).

2.2 Sistemas de comunicação *Through-the-Air* (TTA)

A comunicação TTA, assim como a TTW pode ser usada no cenário de mineração para diversos propósitos. Alguns sistemas são usados para comunicação subterrânea, entre as áreas da mina, outros são usados para comunicação entre o subsolo e a superfície. Quando as tecnologias TTA são usadas para comunicação no interior da mina subterrânea, um dos maiores desafios da é a modelagem da propagação das ondas de rádio. Como mencionado anteriormente o ambiente físico da mina (pilares, máquinas, pedras, entre outros) interfere consideravelmente na propagação, por isso ao fazer a modelagem da

propagação é necessário considerar esses fatores (2).

Um exemplo de sistema TTA usado em comunicação subterrânea são as redes sem fio. Algumas redes sem fio podem demandar infraestrutura, como é o caso da *wireless local area network* (WLAN), e algumas podem ser implementadas sem necessidade de infraestrutura, como as redes Ad Hoc. As redes sem fio podem ser usadas para comunicação entre os trabalhadores da mina e também para fins de localização. Uma desvantagem das redes sem fio é o congestionamento de informação no nó de saída, que é onde os dados dos demais nós são coletados, geralmente são usados protocolos de comunicação para gerenciar o fluxo de informações na rede. Sistemas UWB também estão sendo usados para fazer a comunicação em minas subterrâneas, são sistemas de curto alcance que conseguem fornecer alta taxa de dados, consumindo pouca energia. Uma aplicação para os sistemas UWB é localização e rastreamento, são sistemas rápidos e com boa precisão para este tipo de aplicação (2).

2.3 Sistemas de comunicação *Through-the-Earth* (TTE)

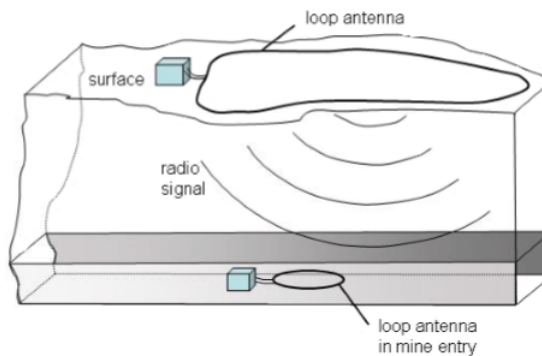


Figura 1 – Representação de sistema TTE com antenas loop.

Em 2006, após acidentes em minas subterrâneas causarem a morte de 19 trabalhadores, foi estabelecido nos Estados Unidos o *MINER Act* de 2006. Uma das resoluções do *MINER Act* era o estabelecimento de comunicação sem fio e bidirecional entre a superfície e as galerias subterrâneas, além de informação de localização. Antes a maioria das minas tinha comunicação através de *leaky feeders* ou *paging phones*. Algumas minas tinham sistemas de localização por áreas (*zone-based tracking*), mas localização por toda a mina não era comum. Desde o *MINER Act* os investimentos no desenvolvimento de sistemas de comunicação por rádio em minas subterrâneas aumentaram, trazendo ao mercado novos produtos para esta aplicação (4).

A maioria das ondas eletromagnéticas não consegue penetrar grandes profundidades no solo, mas ondas ULF (*Ultra Low Frequency*) conseguem penetrar centenas de

metros de profundidade no solo. Os sistemas que tem essa capacidade são chamados de sistemas TTE, eles conseguem fornecer comunicação sem fio entre a mina e a superfície, sem a necessidade de uma infraestrutura complexa, comparada aos outros sistemas de comunicação utilizados em minas, os sistemas TTE tem uma infraestrutura mais simples. Eles também são mais propensos a sobreviver a uma explosão na mina, fornecendo assim a comunicação bidirecional entre a equipe de resgate na superfície e os trabalhadores presos no interior da mina (4).

Os sistemas TTE geralmente operam em faixas de frequência abaixo de 30 kHz. Essas frequências resultam em comprimentos de onda muito longos, podendo chegar a quilômetros. As antenas TTE são pouco eficientes devido aos comprimentos de onda serem muito longos, dessa forma apenas uma fração da potência do transmissor é irradiada. Sistemas TTE operam geralmente em modo half-duplex, ou seja a comunicação é bidirecional, mas ocorre em uma direção por vez e não simultaneamente. Os dados transmitidos podem ser voz em tempo real, mensagens de voz ou texto, pode também ser um sinal de referência periódico, para que seja possível localizar o transmissor (4).

As taxas de transmissão de sinal é que caracterizam os tipos de dados transmitidos, para taxas de até 2,5 kbps é possível transmitir voz em tempo real, já taxas lentas como 10 bps limitam os dados à transmissão de texto com o toque de uma tecla por segundo. A faixa de transmissão TTE é bem limitada devido a alguns fatores da mina subterrânea. Esses fatores limitantes são a taxa de dados, a frequência, a potência do transmissor e a natureza dos estratos de sobrecarga, como condutividade elétrica e outras variações que influenciam as propriedades elétricas da sobrecarga. Ruídos também podem afetar a transmissão do sinal, eles podem ser provenientes da superfície ou do subsolo. A configuração da antena também é um fator importante para a transmissão e recepção do sinal, alguns tipos de antena utilizadas para a comunicação TTE são antenas loop, antenas de núcleo de ferrite e antenas de fio linear (4).

Os sinais de frequência mais alta sofrem maior atenuação através da terra, já os sinais de frequência mais baixa tem taxas de dados limitadas, sendo possível transmitir apenas mensagens de texto ou sinais de referência nessas frequências. Um equipamento com frequência de transmissão variável é interessante para a transmissão TTE, sendo possível configurar transmissão de dados como voz em profundidades menores e sendo capaz de transmitir sinais de referência em profundidades maiores. A potência do transmissor também é um fator limitante da transmissão, quanto maior é potência do transmissor, maior é o alcance da comunicação, mas para que o equipamento não coloque em risco a saúde dos usuários a potência de transmissão deve ser limitada (4).

As antenas TTE que ficam na superfície geralmente não tem seu tamanho limitado, o ideal é que ela abranja todas as áreas da mina que precisam de cobertura. As antenas de recepção e transmissão em formato de loop, devem ser orientadas na mesmas direção

para otimizar o acoplamento das antenas, otimizando consequentemente a transmissão (4).

Os transceptores podem ser fixos ou móveis e podem ser utilizadas para comunicação entre superfície e mina ou comunicação ponto a ponto entre as áreas da mina. Caso sejam transportáveis as antenas subterrâneas deverão ser menores para permitir a locomoção, no caso fixo o fator limitante da antena geralmente é a área da galeria e as características do ambiente físico (presença de pilastras, máquinas, pedras, entre outros) (4).

Os sistemas de comunicação TTE não são ideais para desempenhar um papel de comunicação primária da mina, como para operação remota de máquinas ou comunicação diária de ponto a ponto e de mina e superfície, devido às suas limitações de transmissão e a interferência de ruído. Entretanto, como são sistemas resistentes, são ideais para comunicação secundária, como situações de emergência causadas por acidentes na mina, ou oferecer suporte caso os outros sistemas de comunicação não estejam disponíveis, como pode acontecer em uma situação de falta de energia na mina (4). Algumas outras aplicações em mineração podem se beneficiar da resistência da comunicação TTE, como é o caso presente em (5), onde um equipamento TTE é utilizado para localizar os pontos onde o minério é melhor.

2.3.1 Propagação de Onda em Ambiente TTE

A penetração das ondas de rádio em meio condutor com condutividade σ , permissividade ε e permeabilidade μ obedece, geralmente, a equação da onda (6):

$$\nabla^2 H = \mu\varepsilon \frac{\partial^2 H}{\partial t^2}, \quad (2.1)$$

mas devido as propriedades intrínsecas do solo, canal da comunicação TTE, a propagação das ondas obedece a equação da difusão (6):

$$\nabla^2 H = \mu\sigma \frac{\partial H}{\partial t} \quad (2.2)$$

Devido a isso o campo decai exponencialmente em função da distância, frequência de operação ω e condutividade do meio. A profundidade peculiar é dada pela equação (6):

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}, \quad (2.3)$$

que defini em materiais bons condutores o grau de decaimento do meio. Materiais bons condutores são materiais onde

$$\sigma/(\varepsilon\omega) \gg 1 \quad (2.4)$$

A profundidade representa a distância na qual a intensidade da onda se reduz em 1/e de seu valor, por ser inversamente proporcional à raiz quadrada da frequência implica no uso

de frequências baixas em ambientes TTE. A propagação da onda em meio condutor altera as propriedades básicas da onda, como velocidade de propagação v_{rocha} e comprimento de onda λ_{rocha} . Com isso o comprimento de onda que atravessa as rochas da mina pode ser definido como (6):

$$\lambda_{rocha} = 2\pi\delta. \quad (2.5)$$

2.3.2 Modelo de Campo Magnético

Devido a atenuação do ambiente TTE e a suscetibilidade do meio ao ruído atmosférico e a ruídos gerados pelo homem, os sistemas TTE operam geralmente por meio de indução magnética. A forma mais eficiente de fazer a transmissão é utilizar antenas loop na superfície e no subsolo da mina, que trocam informação por meio de indução magnética. O momento magnético de uma antena loop indica a capacidade de transmissão indutiva da antena e pode ser definido como:

$$m_d = N_{tx} I_{tx} S_{tx}, \quad (2.6)$$

onde N_{tx} é o número de voltas do loop transmissor, I_{tx} é o valor eficaz da corrente elétrica tonal nos filamentos da antena de transmissão e S_{tx} é área do loop. O aumento do momento magnético da antena aumenta consequente a potência dissipada em calor:

$$P_{tx} = R_{tx} I_{tx}^2, \quad (2.7)$$

onde R_{tx} aumenta com o número de voltas do loop ou com a sua área, essa potência dissipada ao determinar a corrente I_{tx} se associa indiretamente a uma potência de transmissão (7) (8).

A aproximação mais simples para o campo magnético variante no tempo gerado por uma antena loop se faz considerando o vácuo como meio homogêneo infinito ignorando qualquer condição de contorno. A intensidade do campo magnético se assemelha a intensidade de um campo estático quando são consideradas distâncias próximas a fonte. Para uma antena loop, o campo magnético quase estático é aproximado por:

$$\mathbf{H}_{qc} = \frac{m_d}{4\pi r^3} \{2 \cos(\theta) \hat{\mathbf{r}} + \sin(\theta) \hat{\boldsymbol{\theta}}\}, \quad (2.8)$$

onde $\hat{\mathbf{r}}$ e $\hat{\boldsymbol{\theta}}$ são vetores em coordenadas esféricas nas direções radial e elevacional. A adaptação para um meio infinito condutivo (mic) para ambos os enlaces de subida e descida é feita modificando-se apenas o número de onda em função da distância pelicular δ . Em coordenadas esféricas, em sua forma fasorial é dada por:

$$\mathbf{H} = \frac{m_d}{4\pi r^3} e^{-jT} e^{-T} \{2 \cos \theta (1 + (1 + j)T) \hat{\mathbf{r}} + \sin \theta (1 + (1 + j)T + 2jT^2) \hat{\boldsymbol{\theta}}\}, \quad (2.9)$$

onde $T = r/\delta = r\sqrt{\mu\sigma\omega/2}$ representa o alcance normalizado pela profundidade peculiar. O ângulo θ é o ângulo entre o ponto que produz a indução e o ponto que sofre a indução.

O campo magnético induz uma tensão na antena receptora. Considerando as antenas receptora e transmissora coaxiais e que estão verticalmente alinhadas, o que implica $\theta = 0$, a tensão induzida é dada por :

$$V_{rx}(\omega) = -j\omega\mu N_{rx}S_{rx}H_r, \quad (2.10)$$

onde H_r é a componente do campo magnético na direção $\hat{\mathbf{r}}$, e N_{rx} e S_{rx} são, respectivamente, o número de voltas e a área do loop receptor (7) (8). Define-se a impedância de transferência como a razão entre a tensão induzida na antena receptora, V_{rx} , e a corrente elétrica na antena transmissora, I_{tx} :

$$Z(T) = \frac{V_{rx}}{I_{tx}}, \quad (2.11)$$

podendo também ser expressa por:

$$Z(T) = N_{tx}N_{rx}S_{tx}S_{rx}\mathbf{F}_r \quad (2.12)$$

em que a função de transferência \mathbf{F}_r é deduzida a partir de (2.9) e (2.10), é dada por:

$$\mathbf{F}_r = \frac{T^2}{\pi\sigma r^5} e^{-T} \sqrt{1 + 2T + 2T^2} \exp\left\{j\left[\arctan\left(\frac{T}{1+T}\right) - T - \frac{\pi}{2}\right]\right\} \quad (2.13)$$

A impedância de transferência é, portanto, uma combinação da função de transferência com os ganhos intrínsecos das antenas de transmissão e recepção. A partir da função de transferência \mathbf{F}_r é possível analisar a distorção causada pelo canal. O canal TTE pode ser caracterizado como um canal passa faixas, assimétrico, onde sua frequência ótima depende da distância entre as antenas e da condutividade do meio (7) (8).

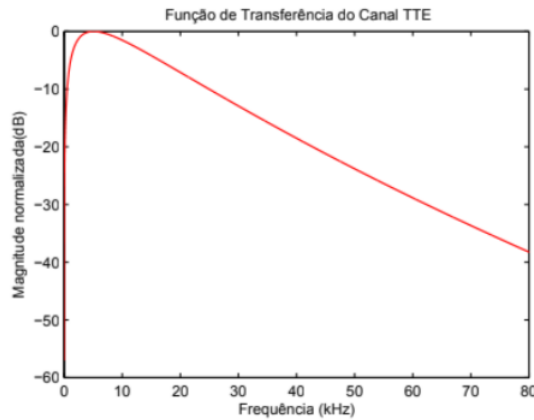


Figura 2 – Função de transferência do canal TTE quando $\sigma = 0,01S/m$ e $r = 200$ m.

2.3.3 Ruído em Ambiente TTE

A transmissão TTE ocorre geralmente nas faixas de frequência VLF e LF, nessas faixas existem dois tipos de ruídos que interferem na transmissão TTE, o ruído atmosférico e o ruído antrópico (*Man Made Noise*) (9).

O ruído antrópico é predominante nas baixas frequências, ele pode ter uma potência maior que a do ruído atmosférico e do sinal recebido pelo sistema de comunicação TTE. O ruído antrópico é gerado dos harmônicos das linhas de transmissão e das máquinas utilizadas na mineração. O ruído é dominante em frequências de até 1 khz e seus harmônicos não são sinais determinísticos, seus espectros ocupam uma certa banda e são centrados em suas frequências nominais (9).

O ruído atmosférico é composto por uma componente gaussiana e uma componente impulsiva. A componente Gaussiana é gerada por descargas distantes do receptor e a componente impulsiva é gerada por componentes próximas ao receptor. O ruído atmosférico P_N é a soma da componente gaussiana Ξ e da componente impulsiva Υ . A função densidade de probabilidade da magnitude da componente Gaussiana é a distribuição de Rayleigh, dada por :

$$p_{\Xi}(\xi) = \left(\frac{2\xi}{R_0^2}\right) \exp\left(-\frac{\xi^2}{R_0^2}\right), \xi \geq 0, \quad (2.14)$$

em que R_0^2 é a potência média de ruído. A distribuição de Weibull descreve a magnitude da componente impulsiva, dada por:

$$p_{\Upsilon}(v) = \left(\frac{av^{a-1}}{R^a}\right) \exp\left[-\left(\frac{v}{R}\right)^a\right], v \geq 0, \quad (2.15)$$

onde $0 \leq a \leq 2$ e R e a dependem da energia média contida em Υ . Uma vez que o ruído atmosférico é dado por $P_N = \Xi + \Upsilon$, a função densidade de probabilidade do ruído é dada pela convolução de (2.14) e (2.15):

$$p_{P_N}(\rho_n) = 2 \int_0^\infty \left(\frac{av^{a-1}}{R^a}\right) \exp\left[-\left(\frac{v}{R}\right)^a\right] \left(\frac{\rho_n - v}{R_0^2}\right) \exp\left(-\frac{(\rho_n - v)^2}{R_0^2}\right) dv, \quad (2.16)$$

a energia média e a potência média do ruído são dadas por:

$$E\{P_N\} = R\Gamma\left(1 + \frac{1}{a}\right) + R_0 \frac{\sqrt{\pi}}{2}, \quad (2.17)$$

$$E\{P_N^2\} = R^2\Gamma\left(1 + \frac{2}{a}\right) + R_0 R \sqrt{\pi} \Gamma\left(1 + \frac{1}{a}\right) + R_0^2, \quad (2.18)$$

em que $\Gamma(\cdot)$ é a função Gamma e $E\{\cdot\}$ é o operador esperança. O parâmetro γ^2 relaciona a razão da energia entre a componente impulsiva e a componente gaussiana do ruído. O parâmetro é dado por (9):

$$\gamma^2 = \frac{E\{\Upsilon^2\}}{E\{\Xi^2\}} = \frac{R^2\Gamma(1 + 2/a)}{R_0^2}. \quad (2.19)$$

3 Conceitos de Compressão de Sinais, Imagens e Vídeos

Falar ao celular, navegar na internet, ouvir música, assistir um filme no cinema ou na televisão, todas essas atividades cotidianas só são possíveis graças a compressão de dados. "Compressão de dados é a arte ou ciência de representar informação em uma forma compacta"(10). Dados digitais podem demandar uma enorme quantidade de bytes para que sejam representados e sem compressão ocupam grande espaço de armazenamento, além de levarem muito tempo para serem transmitidos. A compressão é feita ao identificar as redundâncias que existem nos dados e ao observar sua estrutura. Uma técnica de compressão simples usada em diversos algoritmos incluindo o código Morse, é identificar em um grupo de dados quais são os símbolos mais frequentes e os menos frequentes, ao codificar esses dados usam-se códigos mais curtos para representar os símbolos mais frequentes e os mais longos para codificar os menos frequentes. O contexto de utilização dos dados também pode ser usado como base da compressão, como por exemplo no caso de transmissão de áudio. Nesse caso, os dados transmitidos tem como usuário de destino um humano, considerando que os humanos são capazes de ouvir sons apenas em uma faixa de frequências, os dados presentes em outras frequências podem ser descartados (10).

As técnicas de compressão de dados podem ser divididas em duas classes: compressão sem perdas e compressão com perdas. Na compressão sem perdas não existe perda de informação, o arquivo de entrada comprimido é totalmente recuperado após a decodificação. Esse tipo de compressão é usado quando a aplicação considerada não tolera diferenças entre os dados recuperados e os dados originais, compressão de texto é um exemplo. Para dados que podem tolerar distorção entre os dados originais e os dados recuperados são usadas técnicas de compressão com perdas. A compressão com perdas implica na perda de informação, mas essa perda é recompensada com uma taxa de compressão maior, alguns dados como vídeos e imagens são fontes que geralmente são aplicadas técnicas de compressão com perdas (10).

3.1 Entropia

A entropia de uma fonte é definida como a quantidade média de símbolos binários necessários para codificar a saída de uma fonte. A técnica de compressão sem perdas mais eficiente para uma fonte é a técnica que consegue atingir o mesmo número de bits/símbolo que a entropia da fonte. A entropia pode ser calculada da seguinte forma:

$$H = \sum P(A_i) \log_2 \frac{1}{P(A_i)}, \quad (3.1)$$

onde H é a entropia da fonte, $P(A_i)$ é a probabilidade de ocorrência dos símbolos da fonte e $i(A_i)$ é informação contida na ocorrência de cada símbolo. A informação de um evento está associada com a probabilidade da sua ocorrência, quanto mais raro é esse evento, mais informação ele tem associado. A definição matemática de informação é feita da seguinte forma:

$$i(A_i) = \log_b \frac{1}{P(A_i)} = -\log_b P(A_i). \quad (3.2)$$

Substituindo (3.2) em (3.1) a entropia é definida da seguinte forma (10):

$$H = -\sum P(A_i) \log_b P(A_i). \quad (3.3)$$

3.2 Codificadores Entrópicos

Dois codificadores de entropia muito difundidos são o Código de Huffman e o Codificador Aritmético, nesta seção serão apresentadas as principais características dos dois codificadores.

3.2.1 Códigos de Huffman

O código de Huffman foi desenvolvido por David Huffman quando era aluno de Robert Fano em uma disciplina sobre teoria da informação no MIT (Michigan Institute of Technology). Os códigos de Huffman são códigos prefixos, o que significa que nenhum símbolo do alfabeto tem como prefixo os bits que representam outro símbolo, essa característica faz com que o código seja unicamente decodificável. Além disso os códigos de Huffman são códigos ótimos, os símbolos mais frequentes do alfabeto tem palavras códigos menores que os símbolos menos frequentes. Os dois símbolos menos frequentes do alfabeto terão palavras códigos de mesmo tamanho e elas se diferenciam apenas no bit menos significativo. Para construir um código de Huffman é necessário conhecer as probabilidades de ocorrência dos símbolos do alfabeto, a partir disso são feitos os seguintes passos (10):

1. Reorganizar os símbolos em ordem decrescente de probabilidade, ou seja da maior para menor.
2. Para os dois símbolos com probabilidades de ocorrência menor são atribuídas as palavras código $\alpha_1 * 1$ e $\alpha_1 * 0$, onde $*$ representa concatenação e α_1 é uma string binária a ser definida.
3. As probabilidades dos símbolos menos frequentes são somadas e denotadas a um novo símbolo a' que representa a junção dos dois símbolos menos frequentes, esse símbolo é adicionado ao alfabeto A' . O alfabeto A' contém além do símbolo a' os símbolos do alfabeto A , com exceção dos dois símbolos menos frequentes.

4. Os procedimentos de 1 a 3 são repetidos até que sobrem apenas dois símbolos no alfabeto final, a esses símbolos são atribuídos a palavras códigos 0 e 1.
5. As strings α_1 até α_n são definidas e com isso é possível construir as palavras código de cada símbolo do alfabeto original.

O exemplo abaixo ilustra os procedimentos do algoritmo:

Seja $A = \{a_1, a_2, a_3, a_4, a_5\}$ com probabilidades $P(a_1) = P(a_3) = 0,2, P(a_2) = 0,4$
 $P(a_4) = P(a_5) = 0,1$. Organizando os símbolos temos:

Tabela 1 – Símbolos Organizados

Símbolo	Probabilidade	Palavra Código
a_2	0,4	$c(a_2)$
a_1	0,2	$c(a_1)$
a_3	0,2	$c(a_3)$
a_4	0,1	$c(a_4)$
a_5	0,1	$c(a_2)$

Para os símbolos a_4 e a_5 são definidas as seguintes palavras código:

$$c(a_4) = \alpha_1 * 0$$

$$c(a_5) = \alpha_1 * 1$$

As probabilidades de a_4 e a_5 são somadas e definidas ao símbolo a'_4 . O novo alfabeto A' é composto pelos símbolos a_1, a_2, a_3 e a'_4 . Reorganiza-se o alfabeto:

Tabela 2 – Símbolos Organizados A'

Símbolo	Probabilidade	Palavra Código
a_2	0,4	$c(a_2)$
a_1	0,2	$c(a_1)$
a_3	0,2	$c(a_3)$
a'_4	0,2	$c(a_4) = \alpha_1$

Para os símbolos a_3 e a'_4 são definidas as seguintes palavras código:

$$c(a_3) = \alpha_2 * 0$$

$$c(a'_4) = \alpha_1 = \alpha_2 * 1$$

As probabilidades de a_3 e a'_4 são somadas e definidas ao símbolo a_3 . O novo alfabeto A'' é composto pelos símbolos a_1, a_2 e a_3 . Reorganiza-se o alfabeto:

Para os símbolos a'_3 e a_1 são definidas as seguintes palavras:

Tabela 3 – Símbolos Organizados A''

Símbolo	Probabilidade	Palavra Código
a_2	0,4	$c(a_2)$
a'_3	0,4	$c(a'_3) = \alpha_2$
a_1	0,2	$c(a_1)$

$$c(a'_3) = \alpha_2 = \alpha_3 * 0$$

$$c(a_1) = \alpha_3 * 1$$

As probabilidades de a_3 e a_1 são somadas e definidas ao símbolo a''_3 . O novo alfabeto A'' é composto pelos símbolos a_2 e a''_3 . Reorganiza-se o alfabeto:

Tabela 4 – Símbolos Organizados A'''

Símbolo	Probabilidade	Palavra Código
a''_3	0,6	$c(a''_3) = \alpha_3$
a_2	0,4	$c(a_2)$

O alfabeto A''' tem apenas dois símbolos e a eles são atribuídas as seguintes palavras código:

$$c(a''_3) = \alpha_3 = 0$$

$$c(a_2) = 1$$

Com a string α_3 definida é possível definir as strings α_2 e α_1 :

$$\alpha_2 = \alpha_3 * 0 = 00$$

$$\alpha_1 = \alpha_2 * 1 = 001$$

Com as strings definidas é possível definir as palavras código de cada símbolo e gerar a tabela com o código de Huffman:

Tabela 5 – Código de Huffman

Símbolo	Probabilidade	Palavra Código
a_1	0,2	01
a_2	0,4	1
a_3	0,2	000
a_4	0,1	0010
a_5	0,1	0011

A principal desvantagem dos códigos de Huffman é a necessidade de enviar o alfabeto e as palavras códigos dos símbolos junto com a mensagem, para alfabetos muito

grandes o *overhead* da mensagem pode ser enorme, reduzindo a eficiência da compressão (10).

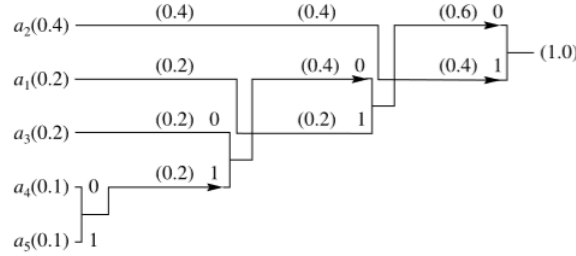


Figura 3 – Construção do código de Huffman

3.2.2 Código Aritmético

A principal ideia do código aritmético é gerar uma *tag* binária a partir das probabilidades acumuladas dos símbolos do alfabeto. Ao transmitir a *tag* não é necessário enviar as palavras códigos dos símbolos, como era necessário no código de Huffman, o que é interessante quando o alfabeto da fonte é muito grande. A *tag* transmitida está contida em um intervalo gerado a partir da sequência dos símbolos transmitidos, a cada símbolo transmitido o intervalo é subdividido, as subdivisões são feitas a partir das probabilidades acumuladas dos símbolos. Com o ultimo símbolo transmitido escolhe-se um valor de *tag* dentro do intervalo final, esse valor será convertido em uma fração binária para ser transmitido. O processo de gerar a *tag* ocorre da seguinte forma (10):

1. O intervalo inicial é $[0,1]$, esse intervalo é subdividido da seguinte forma $[F_x(i-1), F_x(i))$, com $i = 1, \dots, m$. $F_x(i)$ é a probabilidade acumulada dos símbolos a_i .
2. Ao receber um símbolo a_j os limites do intervalo são redefinidos. O limite inferior é calculado como $l(i) = F_x(k-1) + F_x(j-1)/(F_x(k) - F_x(k-1))$, onde $F_x(k-1)$ representa o limite inferior do intervalo anterior, $F_x(k)$ representa o limite superior do intervalo anterior e $F_x(j-1)$ representa a probabilidade acumulada do símbolo do alfabeto anterior a a_j . O limite superior é calculado como $l(s) = F_x(k-1) + F_x(j)/(F_x(k) - F_x(k-1))$, onde $F_x(j)$ é a probabilidade acumulada de a_j .
3. O processo é repetido para todos os símbolos da sequência que se deseja transmitir, no ultimo símbolo escolhe-se um valor dentro do intervalo para ser transmitido.

O procedimento de geração da *tag* é ilustrado na figura.

Para decodificar a *tag* o codificador precisa receber o número de símbolos que foram enviados e precisa conhecer as probabilidades acumuladas dos símbolos do alfabeto (10).

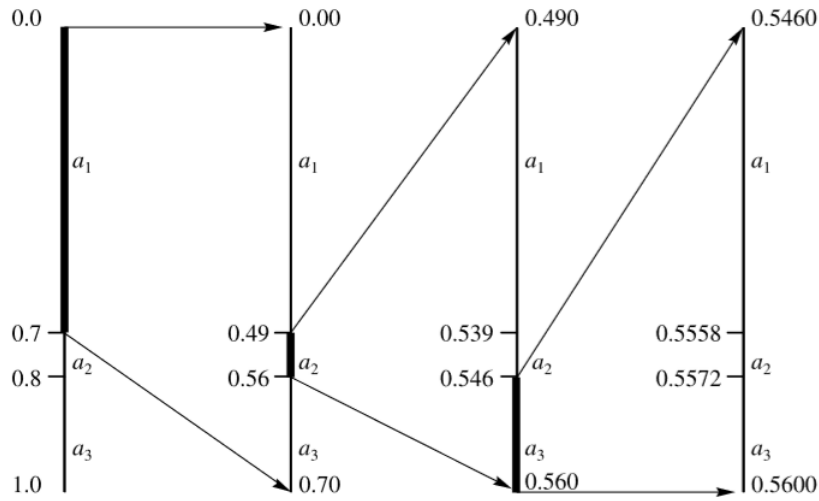


Figura 4 – Procedimento para gerar a tag para a sequência a1a2a3.

3.3 Codificadores com Perdas

3.3.1 Codificação Diferencial

As técnicas de codificação diferencial consistem em técnicas que codificam a diferença das amostras da entrada $\{d_n = x_n - x_{n-1}\}$ ao contrário das outras técnicas que codificam geralmente o próprio valor das amostras. Esse método de codificação pode ser interessante devido ao fato de que em muitas fontes as amostras de $\{X_n\}$ não variam muito de uma pra outra. Para fontes correlatas a codificação diferencial também diminui o erro de quantização. Funcionamento do algoritmo: considerando uma sequência $\{x_n\}$, a sequência de diferença $\{dn\}$ é obtida fazendo $x_n - x_{n-1}$, essa sequência é quantizada e gera a sequência $\{\hat{d}_n\}$:

$$\hat{d}_n = Q[d_n] = d_n + q_n, \quad (3.4)$$

onde q_n é o erro de quantização. A sequência reconstruída no receptor \hat{x}_n é obtida adicionando d_n ao valor anterior reconstruído (10):

$$\hat{x}_n = \hat{x}_{n-1} + \hat{d}_n. \quad (3.5)$$

O transmissor e o receptor começam a partir do valor x_0 , então $\hat{x}_0 = x_0$. Para as primeiras amostras os processos de quantização e reconstrução ocorrem da seguinte forma:

$$d_1 = x_1 - x_0 \quad (3.6)$$

$$\hat{d}_1 = Q[d_1] = d_1 + q_1 \quad (3.7)$$

$$\hat{x}_1 = x_0 + \hat{d}_1 = x_0 + d_1 + q_1 = x_1 + q_1 \quad (3.8)$$

$$d_2 = x_2 - x_1 \quad (3.9)$$

$$\hat{d}_2 = Q[d_2] = d_2 + q_2 \quad (3.10)$$

$$\hat{x}_2 = \hat{x}_1 + \hat{d}_2 = x_1 + q_1 + d_2 + q_2 = x_2 + q_1 + q_2. \quad (3.11)$$

Para a enésima iteração:

$$\hat{x}_n = x_n + \sum_{k=1}^n q_k \quad (3.12)$$

No processo acima o erro de quantização se acumula a partir das iterações, uma maneira de resolver esse problema é usando uma forma alternativa de gerar d_n . Ambos receptor e transmissor conhecem a sequência reconstruída $\{\hat{x}_n\}$, partindo desse princípio a sequência $\{d_n\}$ pode ser gerada da seguinte forma:

$$d_n = x_n - \hat{x}_{n-1}. \quad (3.13)$$

Dessa forma o processo de quantização e reconstrução para as primeiras amostras será o seguinte:

$$d_1 = x_1 - x_0 \quad (3.14)$$

$$d_1 = Q[d_1] = d_1 + q_1 \quad (3.15)$$

$$\hat{x}_1 = x_0 + \hat{d}_1 = x_0 + d_1 + q_1 = x_1 + q_1 \quad (3.16)$$

$$d_2 = x_2 - \hat{x}_1 \quad (3.17)$$

$$\hat{d}_2 = Q[d_2] = d_2 + q_2 \quad (3.18)$$

$$\hat{x}_2 = \hat{x}_1 + \hat{d}_2 = \hat{x}_1 + d_2 + q_2 = x_2 + q_2 \quad (3.19)$$

Para a enésima iteração (10):

$$\hat{x}_n = x_n + q_n, \quad (3.20)$$

agora cada amostra x_n reconstruída tem apenas o erro de quantização da enésima diferença acrescido, gerando um sinal reconstruído mais próximo do sinal original comparado ao processo anterior (10).

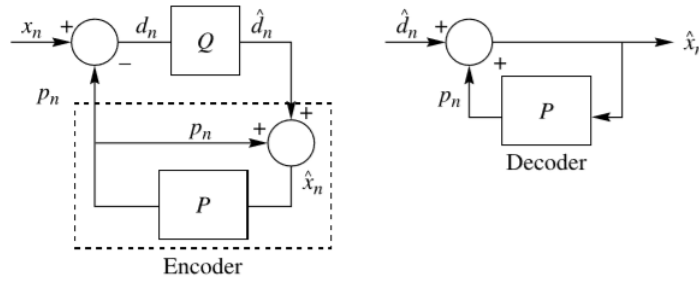


Figura 5 – Algoritmo básico de codificação diferencial.

Fonte: (10)

3.3.2 Codificação por Transformadas

A codificação por transformadas representa o sinal em outro domínio, essa representação mostra diferentes aspectos do sinal que podem ser usados para uma compressão mais eficiente. As transformadas permitem que a maior parte da energia do sinal seja representada em poucos componentes, dessa forma os componentes do sinal que apresentam baixa energia podem ser descartados. Além disso, os sinais de interesse para a percepção humana se concentram em bandas estreitas de frequência, representando o sinal no domínio da frequência é possível filtrar as componentes que não fazem parte da banda de interesse (10).

3.3.2.1 Transformada de Karhunen-Loève (KLT)

As linhas da transformada de Karhunen-Loève são os autovetores da matriz de autocorrelação. O elemento $[R]_{i,j}$ da matriz de autocorrelação de um processo X é dado por:

$$[R]_{i,j} = E[X_n X_{n+|i-j|}]. \quad (3.21)$$

O método de Karhunen-Loève minimiza a média da variância dos coeficientes transformados e é a técnica que promove o maior ganho de compressão de todos os métodos de codificação por transformada. As desvantagens da transformada de Karhunen-Loève é que para fontes não estacionárias a complexidade computacional é alta e é necessário mandar para o receptor a matriz de autocorrelação ou a própria transformada, o que gera um *overhead* significativo (10).

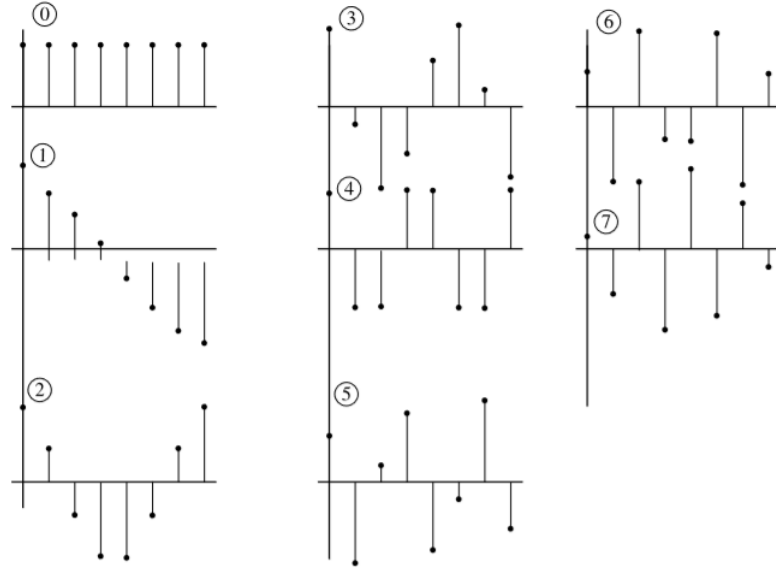


Figura 6 – Bases da DCT.

3.3.2.2 Transformada Discreta de Cosseno (DCT)

Na DCT as linhas da matriz $N \times N$ transformada são obtidas como uma função de cossenos. O elemento $[C]_{i,j}$ da matriz transformada C é dado por:

$$[C]_{i,j} = \begin{cases} \sqrt{\frac{1}{N}} \cos \frac{(2j+1)i\pi}{2N} & i = 0, j = 0, 1, \dots, N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{(2j+1)i\pi}{2N} & i = 1, 2, \dots, N-1, j = 0, 1, \dots, N-1 \end{cases} \quad (3.22)$$

A frequência das linhas da matriz transformada aumenta de cima para baixo, a Figura 6 mostra a variação de acordo com a linha, os números nos círculos correspondem ao número da linha (10).

As matrizes base da DCT são mostradas na Figura 7, o produto que corresponde ao coeficiente θ_{00} , localizado no canto superior esquerdo, é o componente de menor frequência da matriz, já o produto que corresponde ao componente $\theta_{(N-1)(N-1)}$, localizado no canto inferior direito, é o componente de maior frequência da matriz. Na Figura 7 é possível ver que a frequência das matrizes aumenta de cima para baixo e da esquerda para direita. A DCT é melhor em energia de compactação quando comparada a DFT (Transformada de Fourier), em alguns casos, como para fontes de Markov com alto coeficiente de correlação, a DCT se aproxima da transformada de Karhunen-Loève em termos de qualidade de compressão. Devido a sua capacidade de compressão a DCT é uma transformada muito usada, ela faz parte de algoritmos como JPEG, MPEG, dentre outros (10).

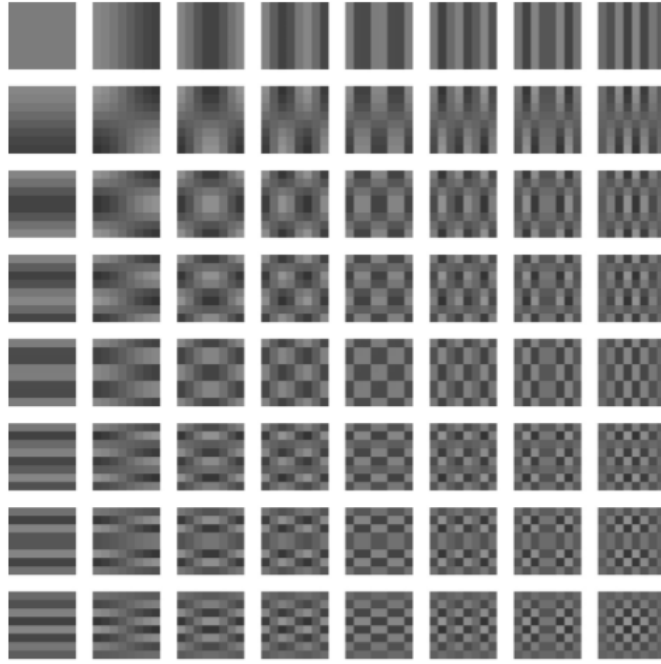


Figura 7 – Matrizes da DCT.

3.3.2.3 Transformada Discreta de Seno (DST)

A DST é a transformada complementar da DCT. A DCT proporciona uma qualidade de compressão próxima a KLT quando o coeficiente de correlação das fontes de Markov tem magnitude alta, já a DST proporciona essa qualidade quando o coeficiente de correlação tem magnitude pequena. O elemento $[S]_{ij}$ da matriz transformada $N \times N$ é obtido pela seguinte fórmula (10):

$$[S]_{ij} = \sqrt{\frac{2}{N+1}} \sin \frac{\pi(i+1)(j+1)}{N+1} \quad i, j = 0, 1, \dots, N-1. \quad (3.23)$$

3.3.2.4 Transformada Discreta de Walsh-Hadamard (DWHT)

A DWHT é uma transformada de simples implementação, as matrizes transformadas da DWHT são rearranjos das matrizes discretas de Hadamard. As matrizes $N \times N$ de Hadamard, onde as dimensões são potências de dois podem ser construídas da seguinte forma (10):

$$H_{2N} = \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix}$$

Se $H_1 = 1$, então:

$$\begin{aligned}
H_2 &= \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
H_4 &= \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \\
H_8 &= \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}
\end{aligned}$$

As matrizes de Hadamard tem a propriedade onde $HH^T = NI$, onde I é a matriz identidade NxN, a matriz transformada da DWHT pode ser obtida multiplicando a matriz de Hadamard por um fator de normalização, assim $HH^T = I$. Além disso, é necessário reordenar as linhas da matriz em ordem decrescente de sequência. A sequência da linha é definida dividindo o número de mudanças de sinal que ocorrem naquela linha por dois. Em H_8 a primeira linha não tem nenhuma mudança de sinal, assim sua sequência é 0, a segunda linha tem 7 mudanças de sinal, logo seu número na sequência é 7/2, o mesmo processo é feito para as outras linhas e ao final as linhas da matriz são reordenadas pela ordem da sequência. O fator de normalização é $1/\sqrt{N}$, assim a matriz H será (10):

$$H = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

A DWHT é uma transformada utilizada para casos onde a complexidade computacional necessita ser baixa, devido a matriz de base consistir em 1 e -1, mas sua qualidade de compactação é menor que a DCT, o que faz com que a DCT seja escolhida em casos onde a complexidade computacional não é um fator limitante (10).

3.3.3 Métricas de Qualidade para Compressão com Perdas

Algumas métricas foram definidas para avaliar a qualidade da compressão de um codificador com perdas, algumas dessas técnicas se baseiam na distorção, que é a diferença entre o sinal original e o sinal reconstruído. As métricas baseadas na distorção são a medida do erro quadrático e a medida da diferença absoluta. A medida do erro quadrático é calculada por:

$$d(x_n, y_n) = (x_n - y_n)^2, \quad (3.24)$$

onde x_n é saída da fonte e y_n é a sequência reconstruída. Para x_n e y_n a diferença absoluta é calculada por:

$$d(x_n, y_n) = |x_n - y_n|. \quad (3.25)$$

A partir da medida do erro quadrático é possível calcular a média do erro quadrático pela seguinte equação (10):

$$\sigma_d^2 = \frac{1}{N} \sum_{n=1}^N (x_n - y_n)^2. \quad (3.26)$$

Também foram definidas métricas para avaliar a qualidade da compressão a partir do tamanho do erro com relação ao sinal. Uma dessas métricas é a Relação Sinal Ruído, calculada por:

$$SNR = \frac{\sigma_x^2}{\sigma_d^2}, \quad (3.27)$$

onde σ_x^2 é a média quadrada da saída da fonte e σ_d^2 é a média do erro quadrático. A SNR também pode ser calculada em decibéis:

$$SNR(dB) = 10 \log_{10} \frac{\sigma_x^2}{\sigma_d^2} \quad (3.28)$$

A Relação Pico Sinal Ruído é uma métrica que relaciona o valor de pico do sinal com o tamanho do erro, ela é calculada por:

$$PSNR(dB) = 10 \log_{10} \frac{x_{peak}^2}{\sigma_d^2}. \quad (3.29)$$

Onde x_{peak} é o valor de pico do sinal (10).

3.4 Teorema de Wyner-Ziv

O teorema de Wyner-Ziv foi desenvolvido por Aaron Wyner e Jacob Ziv em 1976, ele parte de uma generalização do teorema de Slepian-Wolf para o caso com perdas. O teorema de Slepian-Wolf é a base para a codificação distribuída de fonte (*Distributed Source Coding* ou DSC). A teoria da DSC prova que realizar a codificação separadamente de múltiplas fontes estatisticamente dependentes e decodifica-las de forma conjunta, tem

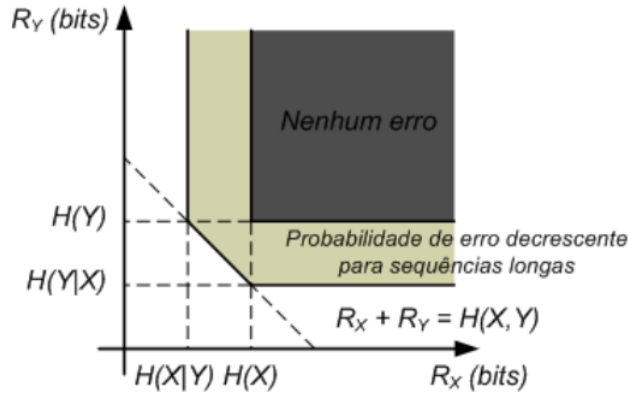


Figura 8 – Teorema de Slepian-Wolf: Região onde a decodificação é perfeita.

desempenho equivalente a fontes codificadas e decodificadas em conjunto (11). O teorema de Slepian-Wolf define que (12), (13):

Teorema de Slepian-Wolf: Para duas fontes X e Y , independentes e identicamente distribuídas (i.i.d) com probabilidade $p(x,y)$, a região onde a decodificação é perfeita é onde $R_x \geq H(X)$ e $R_y \geq H(Y)$, onde $H(X)$ é a entropia de X e $H(Y)$ é a entropia de Y .

O trabalho de Wyner e Ziv mostra que é possível codificar uma sequência X , sem o conhecimento de sua informação lateral Y , e decodificar a informação recebida a partir da codificação de X com o uso de Y , gera uma função taxa-distorção equivalente a função taxa-distorção gerada caso a informação lateral Y também fosse codificada. O teorema de Wyner-Ziv define que (13), (14):

Teorema de Wyner-Ziv: Para duas fontes X e Y , identicamente distribuídas, com probabilidade $p(x,y)$, e sendo $d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^n d(x_i, \hat{x}_i)$, onde x_i são as amostras de X e \hat{x}_i são as amostras reconstruídas de X , a função taxa-distorção com informação lateral é (11):

$$R_y(D) = \min_{p(w|x)} \min_f (I(X; W) - I(Y; W)), \quad (3.30)$$

onde $I(X; W)$ representa a informação mútua de X e W e a minimização é sobre as funções $f : Y \times W \rightarrow \hat{X}$ e funções de probabilidade condicional $p(w|x)$, $|W| \leq |X| + 1$, tal que:

$$\sum_x \sum_w \sum_y p(x, y) p(w|x) d(x, f(y, w)) \leq D. \quad (3.31)$$

4 Metodologia

4.1 Algoritmo Proposto

O algoritmo proposto para a solução do problema foi um algoritmo de Codificação Distribuída de Vídeo (*Distributed Video Coding ou DVC*), baseado no teorema de Wyner-Ziv. Uma das vantagens de utilizar codificação de vídeo distribuída é a transferência da complexidade do codificador de vídeo para o decodificador. Essa característica é importante para a aplicação em ambientes TTE, visto que os transceptores localizados nas galerias da mina subterrânea não dispõem de alta complexidade computacional, já os transceptores da superfície podem ter recursos suficiente para realizar uma decodificação com complexidade computacional maior. Além disso o Wyner-Ziv tem uma abordagem muito próxima da codificação de canal, onde são usados códigos corretores de erro para minimizar as distorções inseridas pelo canal. Essa característica também é importante para a aplicação, visto que o canal TTE sofre grande interferência de ruído, tanto ruído antrópico quanto ruído atmosférico, o que pode gerar grande distorção na informação recebida. Com o algoritmo de Wyner-Ziv é possível comprimir a informação sem degradar muito a robustez da informação transmitida, o que é necessário no cenário TTE.

4.1.1 Funcionamento do Algoritmo

O algoritmo implementado é uma adaptação dos algoritmos descritos (15), onde é apresentado um esquema de DVC usado para codificar vídeos capturados por sensores visuais de 1K-pixel, e (11), onde é desenvolvido um transcodificador de vídeo Winer-Ziv/H.263 para a comunicação entre dispositivos móveis. O algoritmo em (15) foi escolhido por ter um cenário próximo ao cenário de aplicação do trabalho, do algoritmo em (11) foi utilizada a ideia da geração do resíduo o que otimiza a compressão.

No codificador a primeira etapa é fazer o preenchimento dos quadros quando necessário, para que tenham a mesma dimensão da DCT, e separar os quadros em grupos de imagens (GOPs, do inglês *Group of Pictures*), esses GOPs são divididos em quadros-chave e quadros de Wyner-Ziv. Os quadros-chave são enviados ao codificador *intraframe*, a codificação *intraframe* é feita utilizando apenas as informações do próprio quadro, sem utilizar parâmetros e estatísticas de quadros anteriores ou posteriores. Posteriormente os quadros-chave são enviados ao decodificador *intraframe*. Os quadros-chave também são utilizados para gerar o resíduo. Na segunda etapa a DCT é aplicada aos quadros de Wyner-Ziv e aos quadros-chave, os coeficientes da DCT são quantizados por uma matriz de quantização Q e após a quantização o resíduo é gerado. Para gerar o resíduo subtrai-se

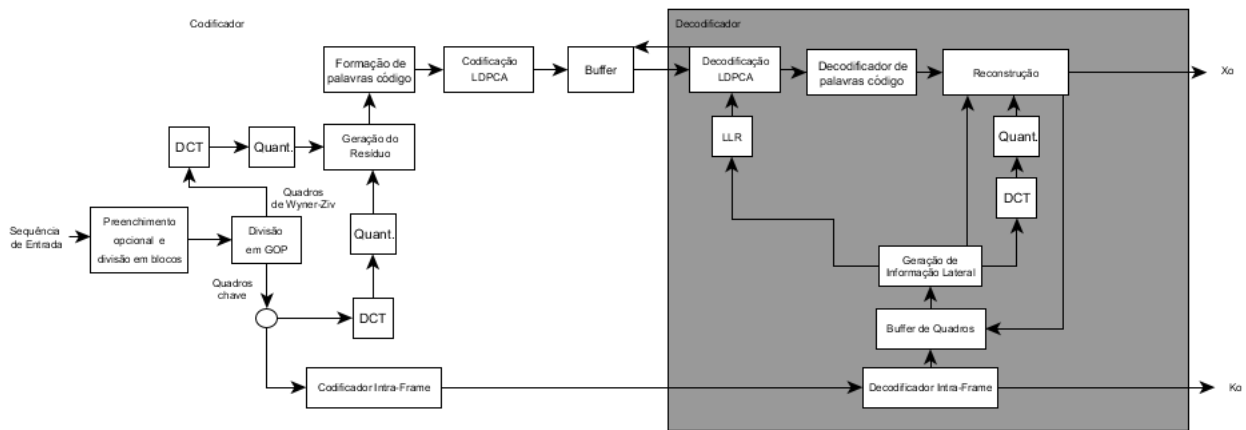


Figura 9 – Diagrama de Blocos do Algoritmo

os coeficientes do quadro de Wyner-Ziv dos coeficientes do quadro chave. Após a geração do resíduo as palavras código são formadas. As palavras código são enviadas para um *low-density parity-check accumulate* (LDPCA) e o resultado é enviado para o *buffer* e é transmitido para do decodificador.

No decodificador os quadros-chave são inicialmente decodificados pelo decodificador *intraframe* e são armazenados em um *buffer* de referência. Após isso o decodificador utiliza os quadros-chave decodificados para gerar a informação lateral. A informação lateral é enviada para o bloco onde é estimada a correlação do canal para expressar a correlação de um quadro de Wyner-Ziv e sua informação lateral. O modelo de correlação para o decodificador LDPCA é obtido através de taxas de log-verossimilhança (LLR, do inglês *log-likelihood rate*). O resultado da decodificação LDPCA é inserido no decodificador das palavras códigos pra que a informação seja recuperada, essa informação é enviada para a etapa de reconstrução. Na etapa de reconstrução a informação é processada e reconstruída e é onde será aplicada a DCT inversa. Ao final desse processo é produzido um quadro de Wyner-Ziv parcialmente decodificado, esse quadro é enviado para o *buffer* da informação lateral. O processo é feito recursivamente até que todas as informações tenham sido decodificadas. Após a aplicação da DCT inversa o quadro de Wyner-Ziv decodificado é armazenado no *buffer* de referência para ser usado futuramente.

4.2 Materiais

4.2.1 Software usado no codificador e no decodificador

O software escolhido para o desenvolvimento dos códigos do codificador e decodificador é o MATLAB. O MATLAB é desenvolvido pela empresa MathWorks, e ele pode ser usado para diversas soluções como visão computacional, sistemas de controle, processamento de sinais, dentre outros. O MATLAB também dispõe de uma ferramenta onde é

possível converter automaticamente o algoritmo construído para C, C++, HDL ou Cuda, e com isso é possível utilizar o código em algum sistema embarcado como FPGA, Raspberry e outros. O MATLAB conta com diversas funções em seu baco de dados, funções matemáticas como transformadas, funções lógicas, funções para gerar diversos tipos de gráfico, entre outras. (16)

4.3 Etapas da Implementação

Para facilitar a implementação do algoritmo, ele foi dividido em passos definidos a partir do diagrama de blocos. Cada passo consiste no desenvolvimento de um bloco principal da implementação. Os passos definidos foram os seguintes:

1. Separação em grupos de quadros
2. Aplicação da DCT
3. Quantização
4. Formação de palavras código
5. Codificador LDPCA
6. Decodificador LDPCA
7. Log de Verossimilhança
8. Geração da informação lateral
9. Codificador *intraframe*
10. Decodificador *intraframe*
11. Aplicação da DCT inversa
12. Reconstrução

Durante a implementação também foram necessários passos intermediários para tratar a informação entre cada um dos blocos. Nessa seção serão abordados apenas os passos principais de construção do algoritmo. Os códigos utilizados podem ser encontrados em (17).

4.3.1 Separação em Grupos de Quadro

O primeiro passo da codificação é a separação em grupos de quadro (GOP, do inglês Groups of Pictures), os quadros são divididos entre quadros-chave e quadros de Wyner-Ziv, os quadros de Wyner-Ziv são codificados pelo codificador LDPCA e os quadros-chave são codificados pelo codificador *intraframe*. Para um GOP de 2 serão codificados um quadro de Wyner-Ziv para cada quadro chave e para um GOP de 3 serão codificados dois quadros de Wyner-Ziv para cada quadro chave. Quanto maior o GOP, menor é a complexidade do codificador, mas a informação lateral gerada no decodificador para os quadros de Wyner-Ziv é pior. Para o algoritmo foi utilizado um GOP de 2.

4.3.2 Aplicação da DCT

As propriedades de compressão da DCT foram abordadas na subseção 3.3.2.2. No algoritmo foi implementada a DCT bidimensional de 8×8 , semelhante a implementação em JPEG. Para aplicação da DCT o frame foi dividido em blocos de 8×8 pixels, em cada bloco foi aplicada a DCT de ordem 8, as matrizes dos coeficientes de cada bloco foram agrupadas em um vetor para posteriormente serem utilizadas na etapa de quantização.

4.3.3 Quantização

As matrizes com os coeficientes da DCT foram quantizadas utilizando a matriz de quantização Q , que é uma matriz de quantização de coeficientes da DCT utilizada no padrão JPEG (10). Também foi definido um passo de quantização de 8 para cada coeficiente.

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 36 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

4.3.4 Formação de Palavras Código

Após a quantização dos coeficientes dos quadros-chave e dos quadros de Wyner-Ziv foi gerado o resíduo. O resíduo foi obtido subtraindo os coeficientes de Wyner-Ziv dos coeficientes chave. Foi definido para as palavras código um tamanho de 4 bits, dessa forma a informação transmitida terá metade do tamanho da informação original. O resíduo

apresentou em sua maioria valores entre -1.6 e 1.6, assim utilizando o intervalo de valores do resíduo R e a quantidade de bits das palavras código foi definido um intervalo de quantização de 0.2 para cada palavra código. Para a codificação dos resíduos foi utilizado o código de gray, devido a sua propriedade de modificar apenas um bit de um número para o outro, o que minimiza o erro em caso de irteferência de ruído. Para os números positivos foi utilizado o código de gray crescente de 1 a 8 e para os números negativos foi utilizado o código de gray decrescente de 15 a 9, dessa forma cada palavra código varia em apenas um bit dos intervalos anterior e posterior. A tabela abaixo mostra a definição das palavras código para cada intervalo de quantização.

Tabela 6 – Definição das Palavras Código

Intervalo de Quantização	Palavra Código
$-1.5 \leq R \leq -1.3$	1101
$-1.3 < R \leq -1.1$	1111
$-1.1 < R \leq -0.9$	1110
$-0.9 < R \leq -0.7$	1010
$-0.7 < R \leq -0.5$	1011
$-0.5 < R \leq -0.3$	1001
$-0.3 < R \leq -0.1$	1000
$-0.1 < R \leq 0.1$	0000
$0.1 < R \leq 0.3$	0001
$0.3 < R \leq 0.5$	0011
$0.5 < R \leq 0.7$	0010
$0.7 < R \leq 0.9$	0110
$0.9 < R \leq 1.1$	0111
$1.1 < R \leq 1.3$	0101
$1.3 < R \leq 1.5$	0100
$1.5 < R \leq 1.7$	1100

4.3.5 Codificador LDPCA e Decodificador LDPCA

Os códigos LDPCA são códigos derivados dos códigos Low-Density Parity-Check (LDPC), eles consistem em um código LDPC concatenado com um acumulador (18). Como a teoria dos códigos LDPC não foi abordada anteriormente, será discutida brevemente nesta seção.

Os códigos LDPC são códigos de bloco binários e lineares com matriz de paridade esparsa, ou seja a maioria dos elementos da matriz são nulos. Os códigos LDPC podem ser regulares ou irregulares dependendo dos seus pesos de linha w_r (quantidade de elementos não nulos em cada linha) e de coluna w_c (quantidade de elementos não nulos em cada coluna), caso tenham pesos constantes para as linhas e para as colunas são regulares, caso contrário são irregulares (19).

Nós de Verificação (NVs)

↓

0 1 2 3 4 5 6 7 8 9 ← Nós de Símbolo (NSs)

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Figura 10 – Matriz esparsa de um bloco linear

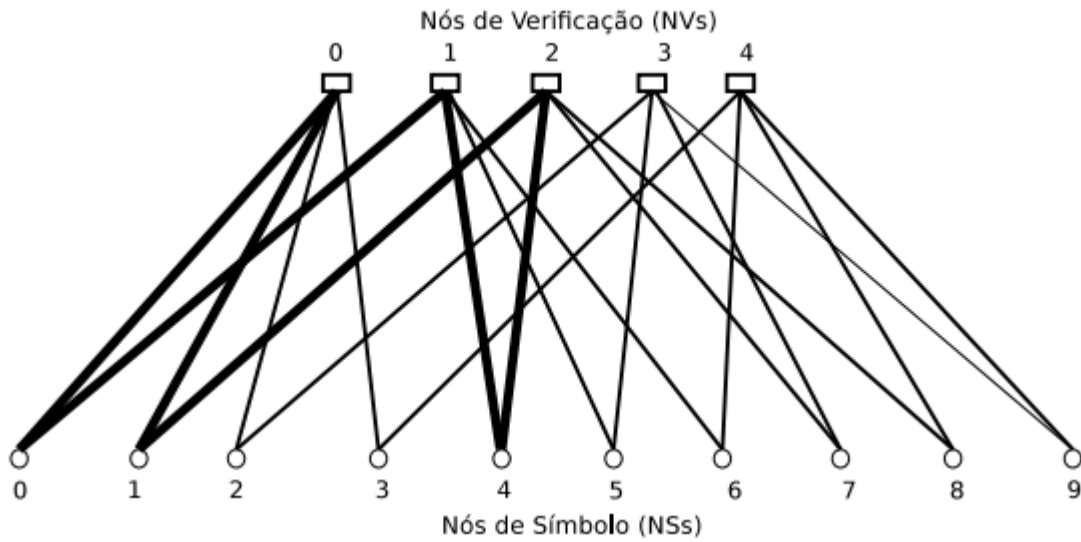


Figura 11 – Quadro original

A representação gráfica dos códigos LDPC é feita por meio de grafos bipartidos, onde os vértices de conjuntos diferentes se conectam. Os vértices são chamados de nós de símbolo (NS), que são os números das colunas da matriz, ou nós de verificação (NV) que são os números das linhas. O grafo bipartido é representado por um grafo de Tanner onde cada NV_i é conectado ao NS_j quando o elemento $H_{i,j}$ da matriz H é diferente de zero. As figuras abaixo mostram um exemplo de matriz esparsa e a representação do seu gráfico de Tanner (19).

Para o codificador e decodificador LDPCA foi utilizada a implementação da equipe da Universidade de Stanford disponível em (20), o algoritmo também implementa a função de log de verossimilhança, pois a mesma é necessária para a decodificação da informação, o decodificador também gera uma informação lateral para ser utilizada na decodificação, essa informação lateral não foi utilizada nos outros passos do algoritmo. O algoritmo funciona da seguinte forma: os bits de síndrome gerados pela codificação da fonte são



Figura 12 – Informação lateral da sequência Akiyo, PSNR = 54.48

acumulados por um buffer, parte desses bits são enviados para o decodificador. O decodificador utiliza a informação lateral Y em conjunto com os bits de síndrome para decodificar a informação, caso os bits de síndrome não sejam suficientes ele requisita mais bits ao codificador. Esse processo se repete até que a informação seja decodificada de maneira satisfatória. Mais informações sobre o funcionamento do algoritmo podem ser encontradas em (18).

4.3.6 Log de Verossimilhança

A log de verossimilhança é o log da função de verossimilhança, ela é computacionalmente mais simples de ser implementada comparada a máxima verossimilhança. A log de verossimilhança do algoritmo é dada por(15):

$$LLR = \log \frac{\int f_{X|Y}(x|y)}{\int 1 - f_{X|Y}(x|y)} \quad (4.1)$$

Onde X é a informação recebida do codificador, Y é a informação lateral e $f_{X|Y}(x|y)$ é a probabilidade condicional de X e Y , quando $X = 0$.

4.3.7 Geração da informação lateral

Para a geração da informação foram utilizados os quadros-chave decodificados pelo decodificador *intraframe* do quadro anterior e posterior ao quadro de Wyner-Ziv, ou seja para gerar a informação lateral do frame 2 foram utilizados os frames 1 e 3. Os frames chave são somados e divididos por 2. A geração da informação lateral apesar de simples, apresentou bom desempenho.



Figura 13 – Informação lateral da sequência Foreman, PSNR = 31.34 dB

4.3.8 Codificador *intraframe* e decodificador *intraframe*

Para o codificador e decodificador *intraframe* foram utilizados o codificação *intraframe* do padrão H264. Para isso foi adaptada a implementação disponível em (21), a implementação contempla as demais etapas do padrão H264, mas apenas a codificação e decodificação intra foi utilizada. Foi escolhido o padrão H264 por ser um padrão de alto desempenho, utilizado em diversas aplicações na indústria.

4.3.9 Aplicação da DCT inversa

Para a aplicação da DCT inversa as palavras código decodificadas foram novamente agrupadas em blocos de 8×8 , assim como no passo de aplicação da DCT no decodificador, e após o agrupamento foi aplicada a DCT bidimensional inversa de 8×8 em cada bloco. A aplicação da DCT inversa será feita na etapa de reconstrução.

4.3.10 Reconstrução

Para reconstruir a imagem o primeiro passo implementado foi filtrar a informação X recebida pelo decodificador LDPCA utilizando um filtro desenvolvido no matlab que filtra o ruído de alta frequência introduzido pelo ruído TTE. O resultado do filtro é somado aos valores anteriores e dividido por 2, isso é feito porque durante a filtragem alguns coeficientes de alta frequência que não são ruído são filtrados e essa informação é perdida. Como o resultado da decodificação consiste no resíduo, para que a imagem seja reconstruída é necessário que o resíduo seja somado ao frame anterior. Essa soma foi feita depois da aplicação da DCT inversa. O resíduo decodificado é multiplicado pela matriz Q e o passo de quantização 8 e posteriormente é aplicada a DCT inversa, o resultado é somado ao frame anterior, gerando assim o frame decodificado de wyner-ziv. Os frames chaves que são codificados pelo codificador intraframe e geram a informação lateral, precisam passar por um filtro de média 2-D, disponível no matlab, esse filtro funciona como um filtro

passa baixas e filtra o ruído adicionado pelo canal TTE. Cada pixel de saída contém o valor médio em uma vizinhança de 3 por 3 em torno do pixel correspondente na imagem de entrada. Os frames chaves decodificados e filtrados são usados para gerar a informação lateral. A informação lateral é somada ao frame decodificado de wyner-ziv e é feita uma média dos dois frames, o frame resultante desse processo é frame final reconstruído.

5 Resultados

Para testar o desempenho do algoritmo foram utilizadas as sequências dos vídeos *Foreman*, *Akiyo*, *Carphone*, *Mobile* e *Salesman* e *Grandma* de 144×176 pixels. Para simulação do canal TTE foi utilizado um simulador de canal desenvolvido pela equipe do Projeto TTE da Universidade de Brasília. O simulador realiza as seguintes funções: transmissão, passagem pelo canal e recepção. Na etapa de transmissão ele gera bits aleatórios, codifica os bits gerados, mapeia os bits codificados em uma constelação e filtra os bits mapeados com um pulso raiz quadrada de cosseno levantado. Na etapa de passagem pelo canal ele filtra o sinal pelo canal TTE e introduz ruído aditivo. Na etapa de recepção ele filtra o sinal recebido com um filtro casado, demodula o sinal, decodifica o sinal e calcula a taxa de erro de bit. Para os testes a geração de bits aleatórios foi substituída pelas palavras código do codificador LDPCA e o resultado dos bits de recepção do simulador foram enviados para o decodificador LDPCA.

5.1 Interferência do Ruído TTE nas Imagens

Para os testes foram utilizados os seguintes parâmetros no simulador: $f_b = 1K$, que define que a frequência de transmissão de bits é de 1 KHz, $BW = 2K$, que define a largura de banda de transmissão como 2 KHz e $E_b/N_0 = [1, 2, 3, 4, 5, 6, 7, 8, 9]$, que são as taxas de sinal ruído pré fixadas. As imagens abaixo ilustram a interferência do ruído TTE nas codificações dos quadros-chave e nos quadros de Wyner-Ziv, nenhuma das imagens abaixo passou por qualquer etapa de filtragem ou otimização, elas representam os bits decodificados recebidos após a passagem pelo simulador. Para visualização dos efeitos do ruído do canal TTE foram geradas 5 imagens: imagem do quadro original, imagem do quadro chave decodificado sem a presença do ruído TTE, imagem do quadro chave decodificado com a presença do ruído TTE, imagem do quadro de Wyner-Ziv decodificado com a presença do ruído TTE e imagem do quadro de Wyner-Ziv decodificado sem a presença do ruído TTE.

Nas imagens é possível visualizar a degradação causada pelo ruído TTE. Visualmente a figura 16 tem pouca diferença das imagens 14 e 16, mas a diferença da PSNR foi de 13.17 dB um valor maior do que a diferença em dB entre as imagens 17 e 18, que apresentaram 9.77 dB de diferença. Visualmente o ruído da figura 17 é bem mais perceptível e a imagem também apresentou o menor valor de PSNR, isso se deve aos passos de quantização, a aplicação da DCT e ao tamanho da palavra código, que por ser menor sofre mais com a interferência de ruído.



Figura 14 – Quadro original



Figura 15 – Quadro chave decodificado sem a presença do ruído TTE, PSNR = 48.04 dB



Figura 16 – Quadro chave decodificado com a presença do ruído TTE, PSNR = 34.87 dB



Figura 17 – Quadro de Wyner-Ziv decodificado com a presença do ruído TTE, PSNR = 21.20 dB



Figura 18 – Quadro de Wyner-Ziv decodificado sem a presença do ruído TTE, PSNR = 30.97 dB

5.2 Tratamento das Informações para Reconstrução da Imagem

Para gerar a imagem final reconstruída são utilizadas as seguintes imagens, cada imagem faz parte de uma etapa do processo de reconstrução:

- Imagem 1 - resíduo somado ao frame anterior após a aplicação da DCT.
- Imagem 2 - Imagem 1 filtrada por um filtro de média.
- Imagem 3 - Frame chave anterior decodificado.
- Imagem 4 - Frame chave posterior decodificado.
- Imagem 5 - Informação lateral, média das Imagens 3 e 4.
- Imagem 6 - Imagem reconstruída.

Os resultados experimentais do processo de reconstrução ilustram claramente a importância da informação lateral e do processamento de imagem no decodificador, com esse processo a imagem reconstruída teve um ganho próximo de 6 dB. Apenas o processo de



Figura 19 – Resíduo somado ao frame anterior após a aplicação da DCT, PSNR = 27.67 dB



Figura 20 – Imagem 1 filtrada por um filtro de média, PSNR = 32.78 dB



Figura 21 – Frame chave anterior decodificado, PSNR = 31.60 dB



Figura 22 – Frame chave posterior decodificado, PSNR = 31.57 dB



Figura 23 – Informação lateral, PSNR = 31.34 dB



Figura 24 – Imagem reconstruída, PSNR = 33.10 dB



Figura 25 – Quadro original

filtragem do ruído já proporcionou um ganho considerável para a reconstrução da imagem. A imagem reconstruída é visualmente bem próxima da imagem original, o que prova que mesmo com a presença do ruído TTE é possível recuperar a informação da imagem de forma satisfatória. O processo também ilustra a eficácia da codificação distribuída de fonte, a medida que a imagem reconstruída apresenta um bom desempenho.

5.3 Taxas de Compressão

Para calcular a taxa de compressão foi utilizada a seguinte equação:

$$C_r = \left(1 - \frac{l_c}{l_o}\right) 100\% \quad (5.1)$$

Onde l_c é o tamanho do código transmitido e l_o é o tamanho do código original. Foram calculadas as taxas de compressão para as sequências mencionadas no início do capítulo, tanto para o frame de entrada colorido quanto em preto e branco.

Tabela 7 – Taxas de Compressão

Vídeo	C_r para entrada colorida	C_r para entrada preta e branca
<i>Carphone</i>	91.30 %	73.62 %
<i>Foreman</i>	90.06 %	70.10 %
<i>Akiyo</i>	92.57 %	77.48 %
<i>Grandma</i>	91.00 %	72.71 %
<i>Salesman</i>	89.55 %	68.28 %
<i>Mobile</i>	85.83 %	65.89 %

As taxas de compressão dos vídeos ficaram bem próxima de 90% no caso do frame de entrada ser colorido, reduzindo grandemente a informação transmitida, já que o algoritmo não transmite informação de cores, dessa maneira a saída será preta e branca. No caso do frame de entrada já ser preto e branco as taxas de compressão ficaram em torno

de 70%, o valor das taxas de compressão do vídeo *Mobile* foi baixo comparado aos demais valores, o vídeo *Akiyo* teve as melhores taxas de compressão, ficando próxima de 93% no caso colorido e de 77% no caso preto e branco. Essas diferenças entre os valores se deve a correlação entre os quadros do vídeo, quanto menor for a diferença entre os quadros adjacentes do vídeo maior será a taxa de compressão, porque o resíduo transmitido será mais próximo de 0 e a informação lateral será melhor. Esse fato pode ser observado nos quadros do vídeo *Akiyo*, que variam pouco entre si e sua informação lateral apresenta valores de PSNR altos. Já o contrário pode ser observado no vídeo *Mobile*, onde existe uma grande variação entre os quadros.

5.4 Curvas de Taxa x PSNR

Uma das etapas do algoritmo de compressão é a quantização, para os resultados anteriores foi usado um fator de quantização $P = 8$. Para observar o comportamento do algoritmo ao mudar o passo de quantização foram feitos testes nas sequências de vídeo variando o passo de quantização entre os valores 2, 4, 8 e 16. Dessa maneira foi possível levantar as curvas de Taxa x PSNR.

Em todas as sequências é possível observar que a curva preta, que representa o passo de quantização 16, não foi a maior em nenhum ponto dos gráficos, não representando a escolha ideal de quantização em nenhum cenário. Esse passo de quantização diminui consideravelmente o resíduo transmitido, dessa maneira não passa quase nenhuma informação ao decodificador. Em quase todas as sequências o passo de quantização 8 foi melhor para os casos onde a taxa de bits/pixel era mais baixa e o passo de quantização 2 foi melhor para os casos onde a taxa de bits/pixel era mais alta. Nos gráficos é possível observar que a variação do valor de PSNR é baixa, isso acontece porque a informação transmitida é o resíduo de uma imagem para a outra, dessa maneira os valores transmitidos são baixos

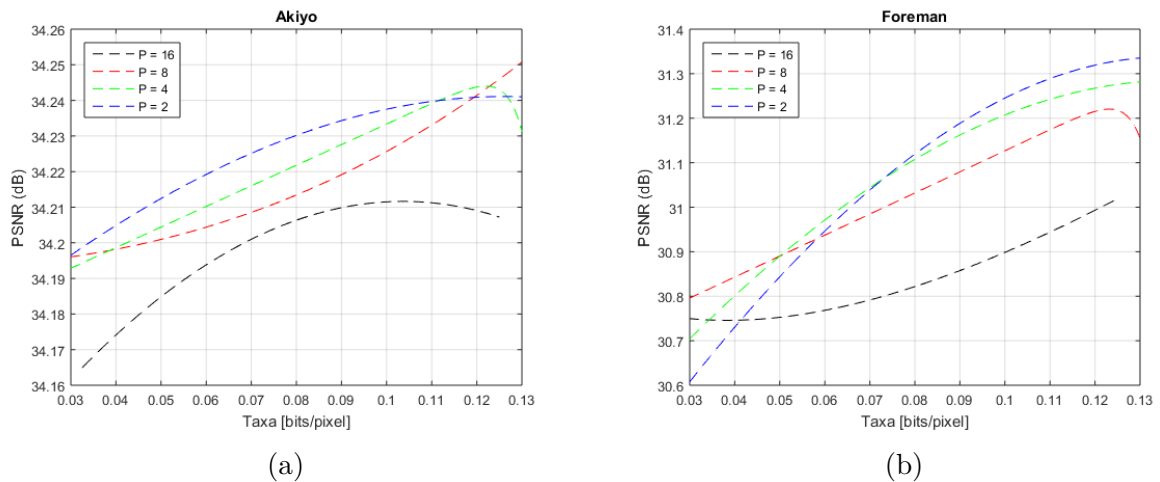


Figura 26 – (a) Taxa x PSNR - Akiyo, (b) Taxa x PSNR - Foreman

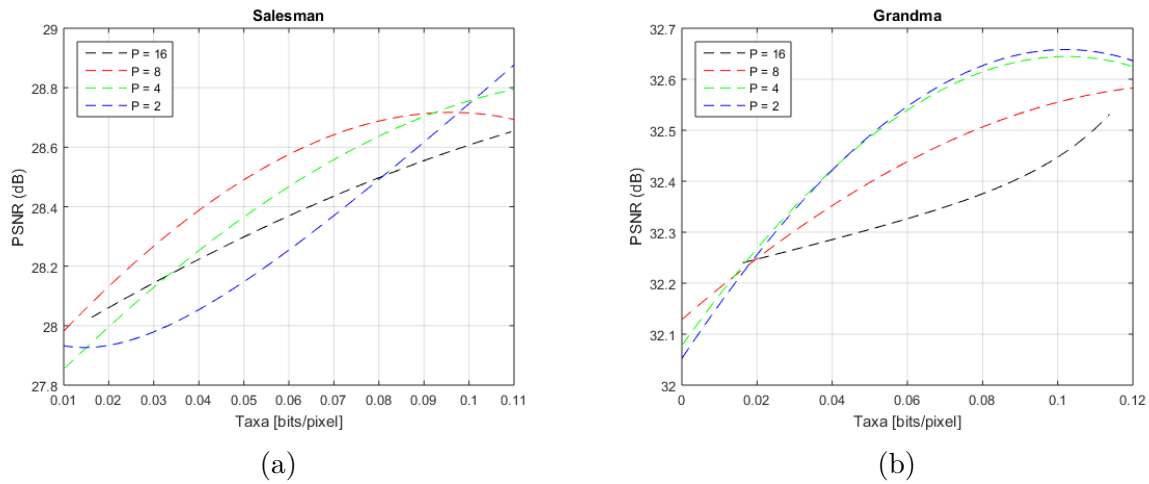


Figura 27 – (a) Taxa x PSNR - Salesman, (b) Taxa x PSNR - Grandma

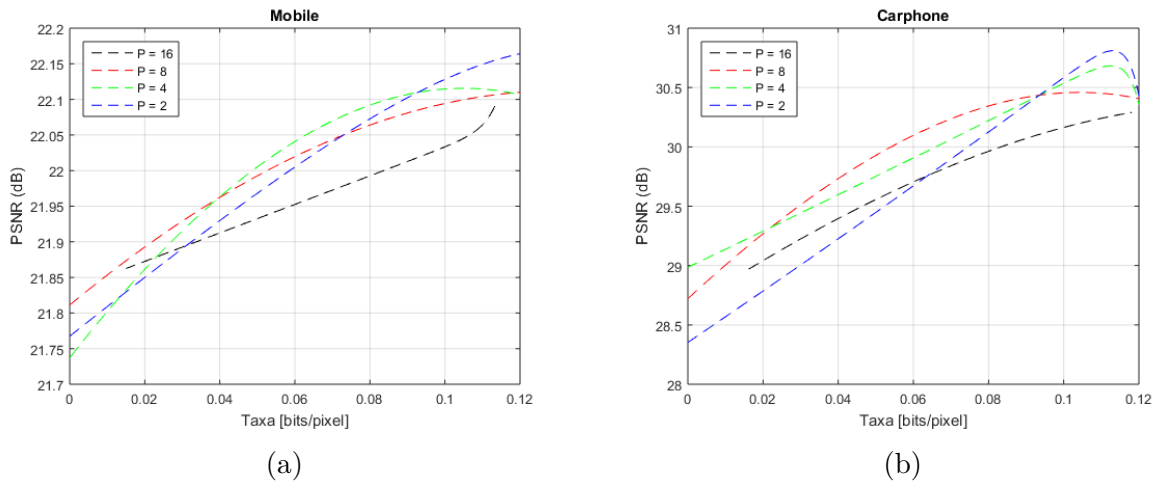


Figura 28 – (a) Taxa x PSNR - Mobile, (b) Taxa x PSNR - Carphone

para qualquer um dos passos de quantização. Assim como observado nos resultados das taxas de compressão a sequência de vídeo Akiyo apresentou o melhor desempenho, com as curvas de PSNR mais alto, enquanto o vídeo Mobile apresentou o pior desempenho, com os valores mais baixos, esse resultado também se deve a correlação entre os quadros dos vídeos.

6 Conclusão

O presente trabalho consiste na implementação de um algoritmo de Wyner-Ziv aplicado a cenários de mineração. O objetivo principal é que o codificador seja utilizado para possibilitar a transmissão de vídeos para o resgate de mineiros em casos de acidentes na mina. Para possibilitar a implementação foram estudadas técnicas de compressão de imagens e vídeos e foi estudado o cenário TTE para conhecimento das limitações do canal.

O codificador de Wyner-Ziv se adequa ao cenário TTE porque ele utiliza a técnica de codificação distribuída de fonte, que é uma técnica semelhante a codificação de canal, o que faz com que o algoritmo realize a compressão, mas continue sendo robusto a ruído. O que é necessário devido ao fato do canal TTE ser muito propenso a presença de ruído. O algoritmo de Wyner-Ziv também é um algoritmo de baixa complexidade no codificador o que se adequa aos requisitos do hardware utilizado em ambientes TTE.

Os códigos foram implementados em MATLAB, essa implementação foi dividida inicialmente em passos de blocos principais, para facilitar a organização do trabalho e as etapas intermediárias entre os blocos foram feitas posteriormente. Após a implementação o algoritmo foi testado em um simulador de canal, que simula a presença de ruído e demais características do canal TTE.

O algoritmo funcionou bem no cenário TTE, mesmo com a presença de ruído a imagem foi reconstruída de maneira satisfatória. O uso da informação lateral foi essencial para que esse resultado fosse atingido, também foram utilizados filtros digitais para filtrar o ruído. Outra característica observada no algoritmo foi que ele funciona melhor em fontes de correlação alta entre os quadros, para esses casos o algoritmo proporciona uma taxa de compressão mais alta e valores de PSNR mais altos também.

Para trabalhos futuros o algoritmo pode ser implementado em hardware para que possa ser testado em campo. Ele também pode ser otimizado utilizando técnicas para maximizar a qualidade da imagem recuperada ou aumentar a taxa de compressão. Outra sugestão é testar o algoritmo em outros cenários com restrições semelhantes ao cenário TTE.

Referências

- 1 FOROOSHANI, A. et al. A survey of wireless communications and propagation modeling in underground mines. *IEEE Communications Surveys and Tutorials*, v. 15, n. 4, 2013. Citado na página 19.
- 2 YARKAN S.; GUZELGOZ S. ARSLAN, H. M. R. Underground mine communications: A survey. in: *Ieee communications surveys and tutorials*. *IEEE Communications Surveys and Tutorials*, v. 1, n. 3, 2009. Citado 4 vezes nas páginas 19, 23, 24 e 25.
- 3 PITTMAN W.; CHURCH, R. M. J. Through-the-earth electromagnetic trapped miner location systems. a review. Bureau of Mines, U.S. Department of the Interior, and the University of Alabama. Citado 2 vezes nas páginas 19 e 20.
- 4 I.I, N. *Tutorial on Wireless Communications and Electronic Tracking Part 2: Advanced*. [S.l.]: Occupational Safety and Health, 2010. Citado 4 vezes nas páginas 20, 25, 26 e 27.
- 5 DANTAS, S. et al. Simulation of ieee 1902.1 (rube) protocol for communication with buried assets. 10 2017. Citado na página 27.
- 6 AGUAYO, L. et al. *Through-The-Earth (TTE) Communications for Underground Mines*. [S.l.]: Journal Of Communications And Information Systems, 2016. Citado 2 vezes nas páginas 27 e 28.
- 7 CARREÑO, J. P. et al. *Desafios para a transmissão de dados em um sistema de comunicação through-the-earth (TTE)*. Santarém, Pará: [s.n.], 2016. Citado 2 vezes nas páginas 28 e 29.
- 8 SILVA, L. S. e. *INTERFERÊNCIA INTERSIMBÓLICA, RUÍDO E CAPACIDADE DE CANAL EM UM SISTEMA DE COMUNICAÇÃO ATRAVÉS-DA-TERRA*. Dissertação (Mestrado em Engenharia Elétrica) — Universidade de Brasília, Brasília, 2017. Citado 2 vezes nas páginas 28 e 29.
- 9 NEVES, S. O. de A. *ANÁLISE DE RUÍDO E SIMULAÇÃO DE ENLACE EM COMUNICAÇÕES ATRAVÉS DA TERRA*. Dissertação (Mestrado em Engenharia Elétrica) — Universidade de Brasília, Brasília, 2017. Citado 2 vezes nas páginas 29 e 30.
- 10 SAYOOD, K. *Introduction to Data Compression*. 5a. ed. 50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States: Morgan Kaufmann, 2018. Citado 11 vezes nas páginas 31, 32, 35, 36, 37, 38, 39, 40, 41, 42 e 48.
- 11 SILVA, E. P. F. da. *TRANSCODIFICADOR DE VÍDEO WYNER-ZIV/H.263 PARA COMUNICAÇÃO ENTRE DISPOSITIVOS MÓVEIS*. Dissertação (Mestrado em Engenharia Elétrica) — Universidade de Brasília, Brasília, 2018. Citado 2 vezes nas páginas 43 e 45.
- 12 Slepian, D.; Wolf, J. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, v. 19, n. 4, p. 471–480, July 1973. ISSN 0018-9448. Citado na página 43.

- 13 COVER, J. A. T. T. M. *Elements of Information Theory*. [S.l.]: Wiley Interscience, 2006. Citado na página 43.
- 14 Wyner, A.; Ziv, J. The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on Information Theory*, v. 22, n. 1, p. 1–10, January 1976. ISSN 0018-9448. Citado na página 43.
- 15 VERBIST, F. et al. Transform-domain wyner-ziv video coding for 1k-pixel visual sensors. *2013 Seventh International Conference on Distributed Smart Cameras (ICDSC)*, p. 1–6, 2013. Citado 2 vezes nas páginas 45 e 51.
- 16 MATLAB. <<https://www.mathworks.com/products/matlab.html>>. Acessado em 03-07-2019. Citado na página 47.
- 17 REPOSITÓRIO do github com os códigos utilizados. <<https://github.com/LuizaCezario/TCC>>. Citado na página 47.
- 18 VARODAYAN, D.; AARON, A.; GIROD, B. Rate-adaptive codes for distributed source coding. *Signal Process.*, Elsevier North-Holland, Inc., Amsterdam, The Netherlands, The Netherlands, v. 86, n. 11, p. 3123–3130, nov. 2006. ISSN 0165-1684. Disponível em: <<http://dx.doi.org/10.1016/j.sigpro.2006.03.012>>. Citado 2 vezes nas páginas 49 e 51.
- 19 SANTOS, L. F. dos. *Decodificadores de baixa complexidade para códigos LDPC Q-ários*. Dissertação (Dissertação (mestrado)) — Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP, 2014. Citado 2 vezes nas páginas 49 e 50.
- 20 RATE-ADAPTIVE LDPC Accumulate Codes for Distributed Source Coding. <<http://ivms.stanford.edu/~varodayan/ldpca.html>>. Acessado em 11-11-2019. Citado na página 50.
- 21 H.264 Baseline Codec v2. <<https://www.mathworks.com/matlabcentral/fileexchange/40359-h-264-baseline-codec-v2>>. Acessado em 01-07-2019. Citado na página 52.