

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**QUALIDADE DE VÍDEO BASEADA EM
PERDA DE PACOTES UTILIZANDO O
PADRÃO *IEEE 802.11AC***

EDEMIR MARCUS CARVALHO DE MATOS

DM 13/2017

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2017

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

EDEMIR MARCUS CARVALHO DE MATOS

**QUALIDADE DE VÍDEO BASEADA EM
PERDA DE PACOTES UTILIZANDO O
PADRÃO *IEEE 802.11AC***

DM 13/2017

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2017

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

EDEMIR MARCUS CARVALHO DE MATOS

**QUALIDADE DE VÍDEO BASEADA EM PERDA DE
PACOTES UTILIZANDO O PADRÃO *IEEE 802.11AC***

Dissertação submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica da UFPA para a obtenção do Grau de Mestre em Engenharia Elétrica na área de Telecomunicações.

Orientador: Evaldo Gonçalves Pelaes
Coorientador: Bruno Souza Lyra Castro

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2017

Dados Internacionais de Catalogação - na – Publicação (CIP) Sistema de Bibliotecas da UFPA

Matos, Edemir Marcus Carvalho de. 1989 -

Qualidade de vídeo baseada em perda de pacotes utilizando o padrão IEEE 802.11AC / Edemir Marcus Carvalho de Matos .- 2017

Orientador: Evaldo Gonçalves Pelaes

Coorientador: Bruno Souza Lyra Castro

Dissertação (mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica , Belém, 2017.

1. Sistemas de comunicação sem fio. 2. Dispositivos de redes sem fio. 3. Vídeo digital – transmissão – qualidade . I. Título.

CDD 23. ed. 621.38456

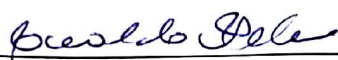
**“QUALIDADE DE VÍDEO BASEADA EM PERDA DE PACOTES
UTILIZANDO O PADRÃO IEEE 802.11AC”**

AUTOR: EDEMIR MARCUS CARVALHO DE MATOS

DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE TELECOMUNICAÇÕES.

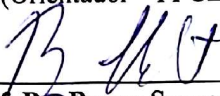
APROVADA EM: 10/03/2017

BANCA EXAMINADORA:



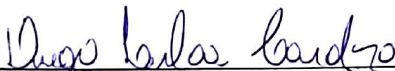
Prof. Dr. Evaldo Gonçalves Pelaes

(Orientador – PPGEE/UFPA)



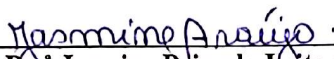
Prof. Dr. Bruno Souza Lyra Castro

(Co-Orientador – UFPA)



Prof. Dr. Diego Lisboa Cardoso

(Avaliador Interno – PPGEE/UFPA)



Prof.ª Dr.ª Jasmine Priscyla Leite de Araújo

(Avaliadora Externa ao Programa – ITEC/UFPA)

VISTO:

Prof. Dr. Evaldo Gonçalves Pelaes

(Coordenador do PPGEE/ITEC/UFPA)

Agradecimentos

Agradeço a Deus.

Agradeço à minha família pelo carinho e apoio dispensado durante todo o Mestrado em Engenharia Elétrica pela Universidade Federal do Pará.

Agradeço aos professores Dr. Evaldo Gonçalves Pelaes, Dr. Gervásio Protásio dos Santos Cavalcante e Dr. Bruno Souza Lyra Castro, pela imensa paciência, orientação, inspiração e pelo conhecimento repassado.

Agradeço aos amigos(as) que direta ou indiretamente deram apoio e motivação para a conclusão de mais uma etapa em minha vida.

Agradeço pela amizade e por todas as contribuições dos meus amigos do Laboratório de Computação e Telecomunicações (LCT¹) da Universidade Federal do Pará.

Agradeço a toda a comunidade de *software* livre, em especial ao *stackoverflow* pelo auxílio indireto. Sem os quais não seria possível a utilização adequada das ferramentas *open source* necessárias para o desenvolvimento deste trabalho.

Agradeço ao grupo de Redes do Centro de Tecnologia da Informação e Comunicação (CTIC) da UFPa por todo o apoio técnico para implementação dos *scripts* utilizados nessa dissertação.

Aos professores e funcionários do Programa de Pós-Graduação em Engenharia Elétrica (PPGEE²) que com seus ensinamentos contribuíram para a realização deste trabalho.

Agradeço ao CNPq³ pelo apoio financeiro em forma de bolsa.

A Universidade Federal do Pará (UFPa), pela oportunidade dada a mim para a realização desta dissertação.

¹<http://www.lct.ufpa.br/>

²<http://www.ppgEE.ufpa.br>

³<http://www.cnpq.br/>

Sumário

1	Introdução	1
1.1	Trabalhos correlatos	1
1.2	Motivações	3
1.3	Objetivos e contribuições	3
1.4	Organização da Dissertação	4
2	Conceitos	5
2.1	Considerações iniciais	5
2.2	Padrão IEEE 802.11	5
2.2.1	Arquitetura IEEE 802.11	6
2.2.2	Camada física	7
2.2.2.1	FHSS	7
2.2.2.2	DSSS	8
2.2.2.3	Infrared light	9
2.2.2.4	HR-DSSS	9
2.2.2.5	OFDM	9
2.2.3	IEEE 802.11ac	10
2.3	Vídeo digital	10
2.3.1	Codificação de vídeo	13
2.3.2	Resoluções de vídeos	14
2.3.3	Padrão H.264/AVC	15
2.3.3.1	Group of pictures	15
2.4	Avaliação objetiva	16
2.4.1	Peak Signal to Noise Ratio	17
2.4.2	Structural Similarity Index	17
2.5	Root Mean Square Error	18
2.6	Considerações finais	18
3	Ferramentas utilizadas	19
3.1	Considerações iniciais	19
3.2	Vídeos sem compressão	19
3.3	Compressão de vídeo	21
3.4	Etapas para transmissão do vídeo	22
3.5	Considerações finais	25
4	Metodologia	26

4.1	Considerações iniciais	26
4.2	Coleta dos dados	26
4.2.1	Transmissão ideal	28
4.2.2	Simulação de perda de pacotes	29
4.2.3	Reconstrução dos vídeos	30
4.2.3.1	Valores de PSNR nos vídeos	31
4.2.3.2	Valores de SSIM nas simulações	35
4.3	Modelagem	37
4.3.1	Perda média de PSNR simulada	39
4.4	Considerações finais	48
5	Análise dos resultados	49
5.1	Considerações iniciais	49
5.2	Coleta dos dados reais	49
5.2.1	Dados utilizados	49
5.2.2	Obtenção de novos dados	50
5.3	Tratamento dos dados reais	51
5.4	Análise comparativa entre os dados reais e as estimativas	53
5.5	Considerações finais	56
6	Conclusões	57
6.1	Trabalhos futuros	58
6.2	Produções científicas durante os estudos	58
	Referências	59

Lista de ilustrações

Figura 1	Arquitetura do padrão <i>IEEE 802.11</i>	6
Figura 2	Modo de operação do <i>IEEE 802.11</i> ad-hoc.	7
Figura 3	Modo de operação do <i>IEEE 802.11</i> infraestrutura.	8
Figura 4	Espectros FDM convencional 4a e OFDM 4b.	9
Figura 5	Exemplo de captura de imagem para o formato RGB.	12
Figura 6	Exemplos de subamostragem.	13
Figura 7	Comparativo entre as resoluções de vídeo.	14
Figura 8	Vídeos utilizados.	20
Figura 9	Diagrama de codificação, transmissão, recepção e reconstrução.	22
Figura 10	Roteador D-Link DIR 803.	27
Figura 11	Adaptador USB DWA-171.	27
Figura 12	Esquema de transmissão.	28
Figura 13	Sorteio das linhas.	30
Figura 14	Valores de PSNR da simulação 19 em 2160p.	32
Figura 15	Valores de PSNR da simulação 19 em 1080p.	33
Figura 16	Valores de PSNR da simulação 19 em 720p.	34
Figura 17	Valores de PSNR da simulação 23 Crowd Run 720p Frame 201 (I).	35
Figura 18	Valores de PSNR da simulação 23 Crowd Run 720p Frame 199 (P).	35
Figura 19	Valores de SSIM da simulação 19 em 2160p.	36
Figura 20	Valores de SSIM da simulação 19 em 1080p.	37
Figura 21	Média das perdas de PSNR para simulação 19.	38
Figura 22	Dados simulados.	40
Figura 23	Tratamento dos dados para perdas de pacote de 4%.	41
Figura 24	Quantidade de amostras aproveitadas.	42
Figura 25	Média das perdas de PSNR.	43
Figura 26	Média da Perda de PSNR em relação a perdas de Frames I e P.	44
Figura 27	Média dados tratados vs Estimativas.	45
Figura 28	Estimativas do comportamento de a e b	47
Figura 29	Dados gerados por Batalha (2016).	50
Figura 30	Medições reais.	51
Figura 31	Dados de Batalha (2016) tratados.	52
Figura 32	Medições reais tratadas.	52

Figura 33	Modelo proposto em relação aos dados de Batalha (2016).	53
Figura 34	Modelo proposto em relação aos dados reais.	54
Figura 35	Comportamento dos <i>Frames</i> I e P para os dados de Batalha (2016). . .	55
Figura 36	Comportamento dos <i>Frames</i> I e P para as medições.	56

Lista de códigos

Código 3.1 – Setup codificação de vídeo.	23
Código 3.2 – Servidor transmissão.	23
Código 3.3 – Cliente transmissão.	23
Código 3.4 – Enviar transmissão.	23
Código 3.5 – Reconstrução do vídeo.	24
Código 3.6 – Conversão mp4 para YUV.	24
Código 3.7 – Obtenção da PSNR.	24
Código 3.8 – Obtenção da SSIM.	24

Lista de tabelas

Tabela 1	Principais diferenças entre 802.11n e 802.11ac	11
Tabela 2	Resoluções	19
Tabela 3	Espaço para armazenamento dos arquivos YUV por resolução.	20
Tabela 4	Configuração do codificador de fonte.	21
Tabela 5	Tamanhos dos vídeos.	21
Tabela 6	Média de PSNR por vídeo.	22
Tabela 7	Descrição dos arquivos gerados na transmissão	23
Tabela 8	Descrição do log do etmp4.	24
Tabela 9	Dispositivos utilizados nas transmissões.	26
Tabela 10	Número de pacotes por transmissão.	29
Tabela 11	Comportamento de a e b	46
Tabela 12	RMSE e desvio padrão do modelo para a média dos dados simulados.	48
Tabela 13	RMSE e desvio padrão do modelo proposto em relação aos dados de Batalha (2016).	54
Tabela 14	RMSE e desvio padrão do modelo proposto em relação aos dados reais	55

Lista de abreviaturas e siglas

1080p Full HD

2160p 4K HD

720p High Definition

AC Access Point

BSS Basic Service Set

CODEC Coder-Decoder

CODEC Codificador/Decodificador

DS Distribution System

ESSS Extended Service Set

FPS Frames Per Second

GOP Group of Pictures

HD High Definition

IBSS Independent Basic Service Set

IEEE Institute of Electrical and Electronics Engineers

ITU-T International Telecommunication Union Telecommunication Standardization Sector

LTE Long Term Evolution

MAC Media Access Control

MIMO Multiple Input, Multiple Output

MOS Mean Opinion Score

MSE Mean Squared Error

PSNR Peak Signal to Noise Ratio

QoS Quality of Service

RMSE Root Mean Square Error

SSIM Structural Similarity Index

STA Wireless LAN Stations

UDP User Datagram Protocol

VQM Video Quality Measure

WLANs Wireless Local Area Networks

Resumo

O fluxo de informação multimídia digital em redes sem fio tem crescido de forma exponencial, devido à popularização dos padrões *IEEE 802.11* como tecnologia de acesso, bem como ao aumento de dispositivos (clientes) que operam neste padrão. O *streaming* de vídeo em altas resoluções também está se tornando comum nessas redes, impulsionando o desenvolvimento de *CODECs* mais eficientes. Neste contexto, algumas métricas de *QoE* e *QoS* devem ser atendidas para entregar um conteúdo com qualidade ao usuário final. Esta dissertação estuda o comportamento do *streaming* de vídeo em rede sem fio utilizando o padrão *IEEE 802.11ac* operando na faixa de $5,2\text{ GHz}$, o *CODEC* avaliado é o H.264/AVC para as resoluções 720p, 1080p e 2160p. A partir das simulações das perdas de pacotes durante a transmissão de vídeo foi desenvolvido uma estimativa da média da perda de PSNR para cada resolução, encontrando a relação da perda de qualidade de vídeo variando a resolução em função da perda de pacotes. O modelo proposto apresentou bons resultados quando comparado com os dados reais, obtendo RMSE de aproximadamente de 2,32 dB e desvio padrão de 2,2 dB. Esta modelagem pode auxiliar no planejamento da rede de comunicação bem como no aprimoramento de novos codificadores de fonte, resultando em uma melhor qualidade de experiência.

Palavras-chave: QoS, QoE, PSNR, IEEE 802.11ac, H.264.

Abstract

The flow of digital multimedia information in wireless networks has grown exponentially due to the popularization of the IEEE 802.11 Standards as access technology, as well as the increase of devices (clients) that operate in this standard. High-resolution video streaming is also becoming commonplace in these networks, driving the development of more efficient CODECs. In this context, some QoE and QoS metrics must be met to deliver quality content to the end user. This dissertation studies the behavior of wireless video streaming using the IEEE 802.11ac standard operating in the 5.2 GHz range, the CODEC rated is the H.264 / AVC for the 720p, 1080p and 2160p resolutions. From the packet loss simulations during video transmission an estimate of the average loss of PSNR for each resolution was developed, finding the relation of the loss of video quality varying the resolution in function of the loss of packets. The proposed model presented good results when compared to the real data, obtaining RMSE of 2.32 dB and standard deviation of 2.2 dB. This modeling can aid in communication network planning as well as enhancement of new source encoders, resulting in a better quality of experience.

Keywords: QoS, QoE, PSNR, IEEE 802.11ac, H.264.

1 Introdução

O número de dispositivos capazes de operar em redes sem fio de comunicação tem aumentado de forma significativa desde a padronização das tecnologias de acesso a redes sem fio, por exemplo *IEEE 802.11*. Como consequência, a quantidade de informação digital que trafega sobre essas redes crescem de forma exponencial, e nesse sentido as novas tecnologias de acesso têm provido uma largura de banda cada vez maior, minimizando a sobrecarga na rede. Dentre as informações (multimídias), as que têm ganhado destaque são os vídeos transmitidos sobre a rede de comunicação, conhecidos como *streaming* de vídeo, devido a popularização dos serviços que fornecem conteúdo de *streaming* de vídeo.

Os vídeos em altíssimas resoluções têm se tornado objeto de estudo tanto no sentido de codificação de fonte quanto para planejamento de redes. Em codificação de fonte, procura-se atender uma qualidade de vídeo aceitável, minimizando o espaço para armazenamento da informação. Enquanto para planejamento de redes procura-se otimizar o fornecimento do serviço. Contudo, em planejamento de redes não somente a qualidade do serviço deve ser atendida, mas também a qualidade de experiência do usuário.

Esta dissertação estuda o comportamento do vídeo sobre uma rede de comunicação sem fio, em ambiente inicialmente controlado (simulado). Utilizando-se quatro diferentes vídeos em três resoluções (720p, 1080p e 2160p) codificados com o H.264/AVC. A rede de comunicação escolhida opera no padrão 802.11ac na faixa de frequência de $5,2\text{ GHz}$. O modelo proposto foi desenvolvido dos resultados simulados e são comparados com os dados reais para fins de validação da proposta.

1.1 Trabalhos correlatos

Nesta seção, são apresentados os trabalhos que tem como objetivo investigar a perda de qualidade de vídeo, sejam por características do meio de comunicação ou do codificador de fonte.

Em Yim e Bovik (2011), investiga-se os efeitos que a perda de pacotes provoca na avaliação temporal da qualidade de vídeo em um ambiente sem fio simulado para o codificador de fonte H.264. Os autores analisam a perda de pacotes por erros do canal de comunicação, simulando três casos em que: sem perdas de pacotes e sem meio sem fio; com perdas de pacotes, com codificação de canal no meio sem fio; com perdas de pacotes, com codificação de canal no meio sem fio. Como resultado os autores propõem

um novo índice para avaliação da qualidade de vídeo, considerando a distorção temporal do vídeo em redes propensas a erros de perdas de pacotes na camada de enlace. Também é destacado a importância da codificação de canal quando trata-se da qualidade de vídeo no meio sem fio.

Em Zulpratita (2013), estuda-se a influência do tamanho do GOP utilizando o codificador H.264/AVC em transmissões de vídeos nas redes *Long Term Evolution* (LTE), em ambiente simulado. Diversos valores de GOP são utilizados a fim de determinar os melhores valores de PSNR e *Mean Opinion Score* (MOS) para cada vídeo estudado, constatando que quanto maior o valor do GOP melhor será a qualidade do vídeo. Também é constatado neste trabalho os efeitos do tamanho do GOP em *streaming* de vídeo em redes LTE. Concluindo que quanto maior o valor do GOP, apesar de oferecer melhor qualidade de vídeo, também aumenta os efeitos de distorção na transmissão do vídeo, definindo que a estrutura GOP influencia diretamente na propagação do erro na transmissão.

Em UHRINA et al. (2014), estuda-se a relação entre o *Group of Pictures* (GOP) no codificador H.264/AVC e a perda de pacotes para a qualidade do vídeo final entregue ao usuário. Os autores codificam dois vídeos, um com bastante movimento e outro com menos, utilizando o codificador H.264/AVC e variando o tamanho do GOP. Em seguida, realizam a transmissão pela rede simulando a perda de pacotes de 0,2% a 3% variando em 0,2%. Diante dos resultados experimentais é feito o comparativo com a métrica objetiva *Structural Similarity Index* (SSIM) em relação as perdas de pacotes. Os autores concluem que quanto maior o tamanho do GOP maior será o impacto na degradação do vídeo quando a análise é feita em perda de pacotes.

Em Frnda, Voznak e Sevcik (2016), propõem um modelo baseado em parâmetros de *Quality of Service* (QoS), perda de pacotes e *jitter*. Utilizando as métricas *Video Quality Measure* (VQM) e *Structural Similarity Index* (SSIM) para verificar a qualidade do vídeo. A análise é feita comparando dois *CODEC*, MPEG-2 e MPEG-4 H.264/AVC, em três resoluções (576p, 720p e 1080p) e três diferentes vídeos (estático, com movimentação e com muita movimentação). Com o objetivo de estudar os impactos que os parâmetros como *CODEC* e *bit rate* podem influenciar na qualidade do vídeo, simulando perdas de pacotes durante a transmissão do vídeo em uma rede IP, tendo como resultado um modelo matemático capaz de estimar a perda da qualidade de vídeo (SSIM) em um ambiente simulado. Outra contribuição deste trabalho diz respeito a utilização de *buffer de jitter* no receptor visto que foi constatado nas simulações uma degradação na qualidade do vídeo quando o valor do atraso é superior a 20 ms.

Em Bondzulic et al. (2016), analisa-se o desempenho da qualidade de vídeo sobre uma rede propensa a erros (perdas de pacotes). Para as simulações doze vídeos foram utilizados, codificados com o H.264/AVC em duas resoluções (seis em CIF e seis 4CIF). Seis porcentagens de perdas de pacotes são aplicadas nas simulações. Este trabalho procura

correlacionar a métrica objetiva (PSNR) com a subjetiva (MOS - Mean Opinion Score), utilizando duas métricas de correlação. O autor prova que a PSNR é uma métrica que tem alto grau de correlação com o MOS, desta forma demonstrando que a PSNR é uma métrica confiável para prever a qualidade subjetiva do vídeo.

A investigação da perda de qualidade de vídeo em redes de comunicações reflete o estado da arte. Contudo, pouco se foi explorado sobre a influência que a resolução tem na degradação do vídeo quando se tem perdas de pacotes durante *streaming* de vídeo. Assim como o um modelo matemático para predição de perda de PSNR para determinada resolução.

1.2 Motivações

O novo padrão de comunicação sem fio 802.11ac fornece altas taxas de transferências, ou largura de banda, possibilitando maior fluxo de dados na rede, e com o aumento no consumo de serviços de *streaming* de vídeo em altas resoluções é interessante entender o comportamento da transmissão de vídeo. Com o objetivo de medir a qualidade conforme ocorram perdas de informação. A proposta de modelagem de perda de qualidade de vídeo a partir de perdas de pacotes visa auxiliar no aprimoramento das técnicas de codificação de fonte bem como no planejamento de redes de comunicação.

1.3 Objetivos e contribuições

Como objetivo geral desta dissertação tem-se:

- Propor uma modelagem matemática de predição da média de perda de PSNR para três resoluções em função das perdas de pacotes na rede de comunicação baseada em simulações.

As principais contribuições desta dissertação são:

- O impacto da perda de pacotes na transmissão de vídeo em relação a qualidade de vídeo recebido;
- Mensurar o comportamento da perda de qualidade de vídeo relacionada as características do vídeo;
- Mensurar o comportamento médio da perda de qualidade de vídeo em relação a resolução;

1.4 Organização da Dissertação

A presente dissertação esta organizada da seguinte forma:

- Este capítulo busca contextualizar o objeto de estudo, apresenta alguns trabalhos correlatos e suas contribuições; a importância da pesquisa; o objetivo desta dissertação e suas contribuições;
- No Capítulo 2 são apresentados alguns conceitos para compreensão sobre os padrões em estudo;
- O Capítulo 3 explora as características dos vídeos e ferramentas utilizadas nesta dissertação;
- No Capítulo 4 descreve a metodologia utilizada para o desenvolvimento da proposta de modelagem;
- No Capítulo 5 discuti-se sobre os resultados obtidos em contraste com os dados medidos em ambiente real;
- As conclusões, dificuldades e trabalhos futuros serão expostos no Capítulo 6.

2 Conceitos

2.1 Considerações iniciais

Neste capítulo serão descritas algumas características das tecnologias envolvidas para o desenvolvimento desta pesquisa. Assim como as métricas para avaliação de desempenho relacionadas à qualidade de vídeo e ao modelo proposto.

2.2 Padrão IEEE 802.11

A necessidade por comunicação sem fio é sentida há muito tempo, devido as limitações impostas pelas redes cabeadas (mobilidade, alcance e custo). Além disso, o crescente número de usuários para acessar a Internet, fez com que se começasse a estudar alternativas para a distribuição do serviço entre os possíveis assinantes. A comunicação por banda estreita é bastante instável e lenta. O acesso à Internet e outros serviços de telecomunicação com qualidade começava a se tornar uma necessidade, e não mais uma possibilidade.

Quando trata-se de rede sem fio a tecnologia mais presente é o *IEEE 802.11*, responsável por fornecer acesso a rede de comunicação pelo meio sem fio, e aliado ao aumento de dispositivos móveis, fez com que melhorias sejam desenvolvidas a fim de proporcionar melhor qualidade de serviço e experiência ao usuário final.

O padrão *IEEE 802.11* original (IEEE 802.11, 1999) publicado em 1997 pelo *Institute of Electrical and Electronics Engineers* (IEEE), foi projetado para redes locais sem fio. Mais tarde, em 1999, uma organização sem fins lucrativos foi formada com o objetivo de certificar a interoperabilidade de produtos baseados no padrão *IEEE 802.11*, a *Wi-Fi Alliance*. Através de teste de interoperabilidade, o programa de certificação da *Wi-Fi Alliance* assegura que produtos de diferentes fabricantes funcionem sem problemas (JUNIOR, 2013).

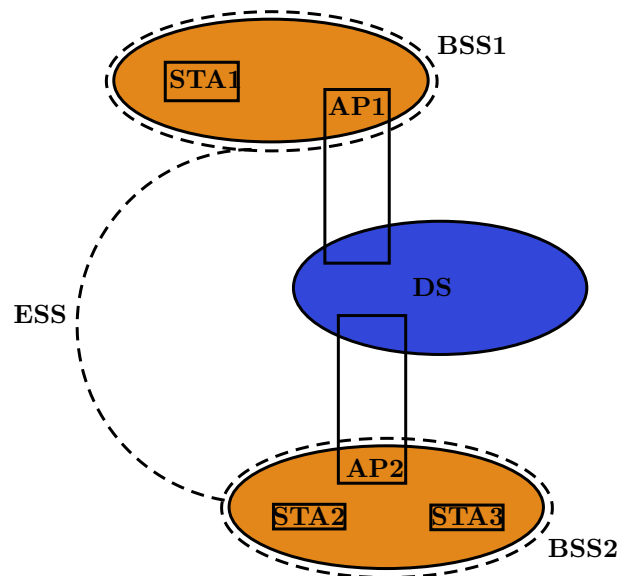
Os dispositivos do padrão *IEEE 802.11* operam nas faixas não licenciadas do espectro eletromagnético, ou seja, nas bandas *Industrial, Scientific, Medical* (ISM). Existem três bandas ISM em uso atualmente: a banda ISM de *902-908 MHz*, a banda ISM de *2,4-2,485 GHz* e a banda ISM de *5,734 - 5,860 GHz*, sendo a banda ISM de *2,4 GHz* mais comumente usada. Geralmente, os aparelhos que funcionam nas bandas ISM têm transmissores de baixa potência e não são capazes de cobrir mais do que alguns metros, como

é o caso das *Wireless Local Area Networks* (WLANs) (TANENBAUM, 2003).

2.2.1 Arquitetura IEEE 802.11

A arquitetura do *IEEE 802.11* é composta de alguns elementos que quando integrados formam a rede sem fio (TANENBAUM, 2003), conforme na Figura 1. Os elementos serão descritos a seguir.

Figura 1 – Arquitetura do padrão *IEEE 802.11*.



Fonte: Adaptação de Tanenbaum (2003)

O *Basic Service Set* (BSS), corresponde a uma célula de comunicação da rede. Esta célula é gerenciada por um *Access Point* (AP). De forma geral o BSS é composto por um AP e vários STA (*Wireless LAN Stations*). Os STAs são os diversos dispositivos clientes da rede com interface para conectar a um ambiente sem fio, devem possuir um *Media Access Control* (MAC), uma interface de camada física e tem que estar de acordo com o padrão *IEEE 802.11*.

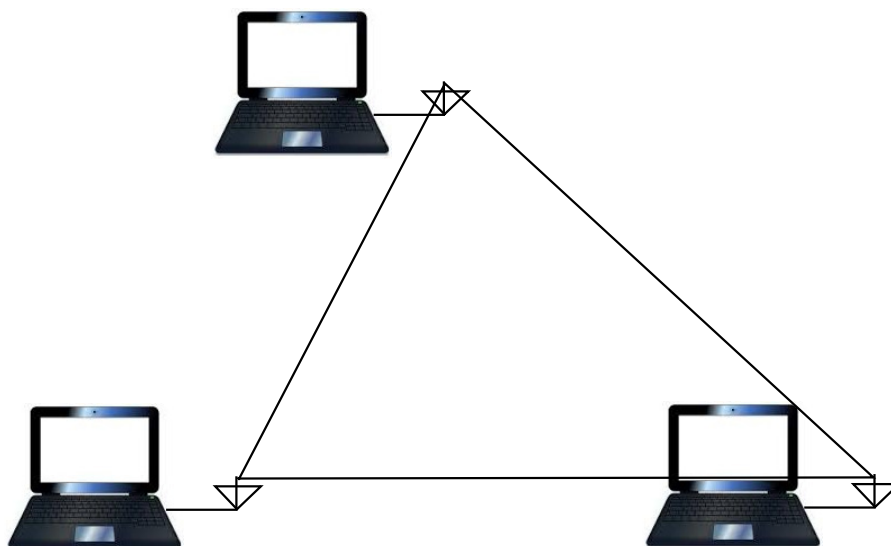
O *Distribution System* (DS), corresponde o *backbone* entre as BSSs, permitindo a comunicação entre as *WLANs*, ou seja, interliga as estações de diferentes BSSs.

O *Extended Service Set* (ESS) é o conjunto de células BSSs cujos APs estão conectados a uma mesma rede convencional permitindo uma maior área de cobertura. Nestas condições uma determinada STA pode se movimentar de uma célula BSS para outra permanecendo conectada à rede. Este processo é denominado de *Roaming*.

Para o padrão *IEEE 802.11* existem dois modos de operação, Ad-hoc e infraestrutura. Na rede ad-hoc, os clientes sem fio comunicam-se diretamente com os outros sem o uso de AP. Essa rede também é chamada de *peer-to-peer* (ponto-a-ponto). Na Figura 2 refere-se a uma rede Ad Hoc com um BSS independente, conhecida como *Independent*

Basic Service Set (IBSS). Na rede infraestruturada, é utilizado um ponto de acesso que é responsável por quase toda a funcionalidade de rede, como visto em Figura 3. Como observar-se, o que difere o modo de operação do *IEEE 802.11* é o gerenciamento da comunicação (presença de uma AP). O mais comum é termos redes sem fim operando em infraestrutura.

Figura 2 – Modo de operação do *IEEE 802.11* ad-hoc.



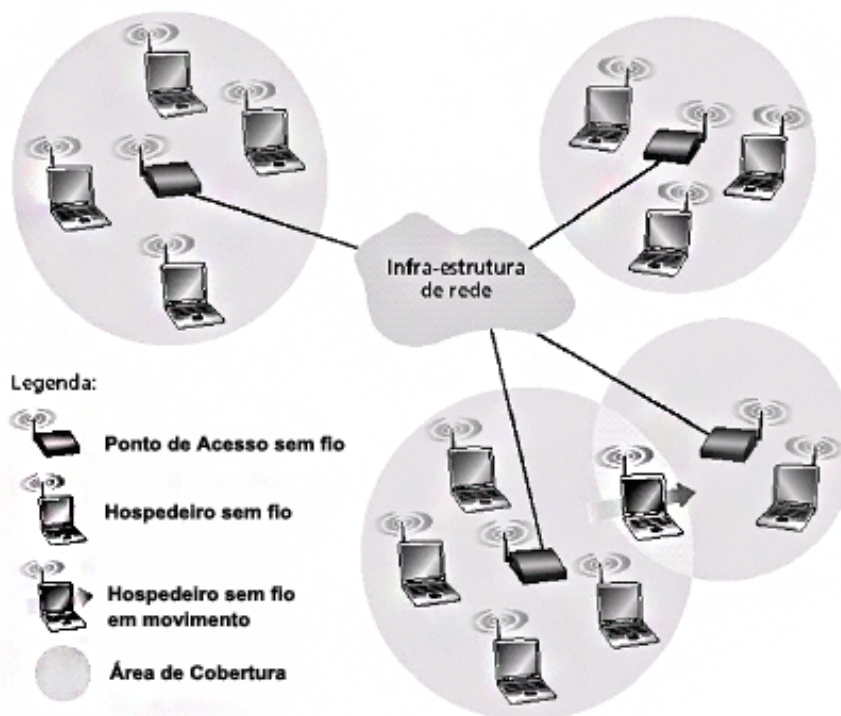
Fonte: Adaptação de ARAÚJO (2011)

2.2.2 Camada física

Originalmente foram definidas três técnicas de transmissão para o padrão *IEEE 802.11* na camada física: FHSS (*Frequency Hopping Spread*), DSSS (*Direct Sequence Spread Spectrum*) e IR (*Infrared light*). Mais tarde, em 1999, para alcançar maior largura de banda, surgiram duas novas técnicas de transmissão: OFDM (*Orthogonal Frequency Division Multiplexing*) usada primeiramente no padrão 802.11a e posteriormente no 802.11g e o HR-DSSS (*High Rate Direct Sequence Spread Spectrum*) usada no padrão 802.11b (TANENBAUM, 2003; JUNIOR, 2013).

2.2.2.1 FHSS

O FHSS é uma técnica que utiliza como meio de transmissão o rádio de alcance limitado, operando na banda ISM de $2,4\text{ GHz}$. A banda de frequência é dividida em 79 canais com 1 MHz de largura, sendo gerada uma sequência pseudo-aleatória destes canais, por onde o sinal é difundido. É necessário garantir o sincronismo de todas as estações para que elas mudem para as mesmas frequências de forma simultânea, utilizando igualmente os canais da sequência. Isso pode ser assegurado com a utilização de um mesmo gerador de números pseudo-aleatórios.

Figura 3 – Modo de operação do *IEEE 802.11* infraestrutura.

Fonte: ARAÚJO (2011)

Em um determinado momento, um canal desta sequência é utilizado por curto período de tempo para transmissão dos dados. Com o sincronismo entre receptor e o transmissor, considerando que a série de canais deste é conhecida pelo receptor, a informação será totalmente recuperada fornecendo, além disso, maior segurança. Assegurando assim que nenhum intruso poderá espionar as transmissões se não conhecer a sequência de saltos ou o tempo de parada.

Esta camada provê operações em *1 Mbps*, com *2 Mbps* opcional. A versão de *1 Mbps* utiliza 2 (dois) níveis da modulação GFSK (*Gaussian Frequency Shift Keying*), e a de *2 Mbps* utiliza quatro níveis da mesma modulação. O FHSS também é razoavelmente insensível a interferência de rádio, e tem como principal desvantagem baixa largura de banda.

2.2.2.2 DSSS

O DSSS é um método alternativo de espalhamento de espectro, onde os códigos são separados. O Método usa também o padrão ISM de *2,4 GHz* e a taxa básica de *1 Mbps* é gerada através de uma modulação diferencial binária por chaveamento de fase DBPSK (*Differential Binary Phase Shift Keying*), e a taxa de *2 Mbps* usa uma modulação DQPSK (*Differential Quadrature Phase Shift Keying*). O espalhamento é feito através da divisão do padrão disponível em 11 subcanais de *22 MHz* de largura de banda.

2.2.2.3 Infrared light

A terceira técnica para transmissão é o Infravermelho e possui um comprimento de onda pequeno e é recomendado para ambientes fechados e pequenos. Os sinais infravermelhos não ultrapassam paredes e têm alcance máximo de 10 metros dependendo do ambiente. A taxa de transmissão opera na faixa de 1 a 2 *Mbps*.

2.2.2.4 HR-DSSS

O HR-DSSS é uma técnica que apresenta melhorias se comparada com o DSSS. Com o objetivo de obter maiores taxas de transmissões, admite taxas dinâmicas que variam de 1, 2, 5 e 11 *Mbps*. Esse ajuste na velocidade é feito sob condições de carga e ruído de forma a fornecer velocidade ótima em condições adversas. Esta técnica também é empregada na banda 2,4 *GHz*, como por exemplo no 802.11b.

2.2.2.5 OFDM

O OFDM tem como principal vantagem a sua grande capacidade de minimizar efeitos negativos provocados por canais com desvanecimento seletivo em frequência, oferecendo maior facilidade na equalização se comparado a um sistema de uma única portadora. A técnica OFDM se baseia na sobreposição dos espectros das subportadoras, sendo possível através da ortogonalidade entre as mesmas, o que de fato reduz a banda ocupada pelo sinal.

Neste esquema de transmissão paralelo, um fluxo de dados em série com taxa de transmissão elevada é dividido em um conjunto de sub-fluxos paralelos de taxa de transmissão baixa que podem ser modulados, cada um, por diferentes técnicas de modulação. Essa redução na taxa de símbolos representa um estreitamento do pulso no domínio da frequência, tornando menos sensível a seletividade em frequência. De forma geral, o OFDM proporciona uma melhor eficiência espectral, para ilustrar o aproveitamento espectral a Figura 4 compara FDM (*Frequency Division Multiplexing*) com o OFDM.

Figura 4 – Espectros FDM convencional 4a e OFDM 4b.



Fonte: Prasad (2004).

2.2.3 IEEE 802.11ac

O padrão *IEEE 802.11ac* para WLANs teve início de seu desenvolvimento no fim de 2008 por um dos grupos do *IEEE*, instituição responsável por criar alguns padrões, e teve aprovação em meados de 2013 com melhorias nas camadas física e MAC. Este padrão opera apenas na faixa de 5 GHz , garantindo compatibilidade com suas versões anteriores *IEEE 802.11a* e *802.11n* (IEEE 802.11AC', 2013).

No geral, o *802.11ac* pode ser visto como uma extensão do *802.11n*, onde os dois conceitos básicos de *Multiple Input, Multiple Output* (MIMO) e maior largura de banda do canal, são reforçados. A ideia básica é que a taxa de dados máxima teórica possa ser linearmente aumentada por um fator do número de fluxos espaciais (antenas de transmissão/recepção) ou pela largura de banda de canal. Em outras palavras, a taxa de dados pode ser elevada duplicando-se o número de fluxos espaciais ou/e a largura de banda do canal (ONG et al., 2011).

As principais novidades observadas no novo padrão são:

- larguras de banda de 80 MHz e 160 MHz para comunicação;
- até 8 fluxos espaciais com MIMO;
- ordem de modulação de 256-QAM;
- multiusuários com MIMO, por meio de acesso múltiplo por divisão no espaço SDMA.

Devido às dificuldades de se alocar um canal de 160 MHz contínuo (há apenas dois disponíveis na faixa de 5 GHz), o padrão permite que dois canais de 80 MHz descontínuos sejam utilizados em uma transmissão que, ao todo, requeira 160 MHz de largura de banda.

O novo padrão prevê também a possibilidade de transmissão com múltiplos usuários. Um sistema MU-MIMO, permite que um determinado AP transmita dados para mais de um usuário ao mesmo tempo através de *beamforming*, técnica de múltiplas antenas com reuso espacial.

Como mencionado, o novo padrão pode ser visto como extensão do seu antecessor. Partindo para uma análise comparativa destaca-se a introdução de dois novos MCS 8 e 9 com base no esquema de modulação 256-QAM e taxas de códigos 3/4 e 5/6, atingindo acréscimos de 20% a 33% na taxa de transmissão dos dados (ONG et al., 2011). Algumas diferenças entre o padrão atual e o seu antecessor, *802.11n*, podem ser vistas na Tabela 1.

2.3 Vídeo digital

O vídeo digital é uma sequência de imagens codificadas e exibidas em rápida sucessão. Para o sistema de recepção visual humano é um movimento contínuo, devido

Tabela 1 – Principais diferenças entre 802.11n e 802.11ac

	IEEE 802.11n	IEEE802.11ac
Largura de Banda	20 MHz e 40 MHz	40 e 80 MHz (Obrigatório), 160 e 80 + 80 MHz, opcional
Multiusuários	Não	Opcional
Fluxos espaciais	Até 4	Até 8
Esquemas de Modulações	Até 4	Até 8
MCS desigual	BPSK, QPSK, 16-QAM 64-QAM	BPSK, QPSK, 16-QAM,64-QAM e 256-QAM (opcional)
Faixa de operação	2,4 GHz e 5 GHz	5 GHz
Preambulo Greenfield	Opcional	Não

Fonte: Batalha (2016).

a persistência da visão. A diferença de tempo entre as imagens pode ser de $0,1s$ para movimento muito lento, e menos de $0,05$ (mais de 20 *frames* por segundo) para movimentos rápidos, não apresenta uma apreciável diferença de qualidade. A maioria dos vídeos usa 15, 24, 25 ou 30 quadros por segundo (RICHARDSON, 2011). Pode-se dizer que o vídeo digital nada mais é que uma representação de varias imagens amostradas espacialmente e temporalmente.

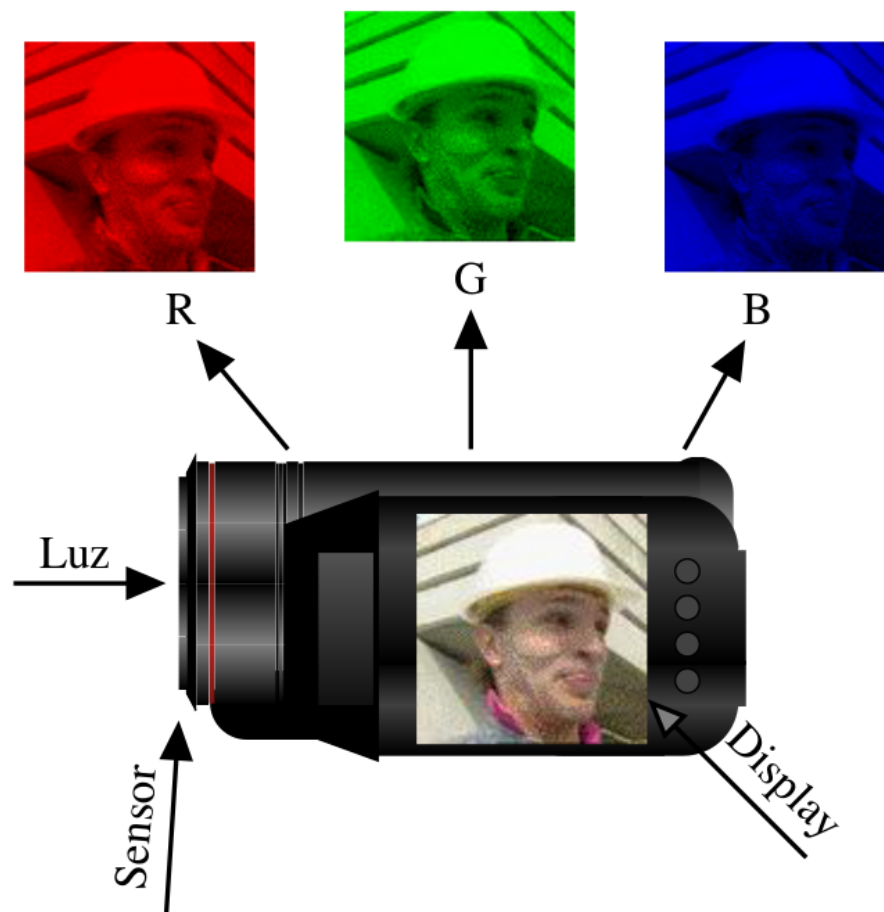
O vídeo é uma mídia que contém muita informação e conseqüentemente demanda mais espaço para armazenamento (*bits*), principalmente dentro das redes de comunicações. Os vídeos avançam para uma gama maior de dispositivos e atingem maiores resoluções, impulsionando a geração de novas tecnologias de vídeos. Nesse contexto, o usuário final está exigindo cada vez mais qualidade nos serviços, influenciando diretamente na resolução do vídeo que é entregue, neste caso, vídeos em altíssimas resoluções (4K e 8K).

A captura da imagem feita pela câmera, por exemplo, utiliza um Dispositivo de Carga-acoplada (*Charge-Coupled Device*, CCD), capaz de armazenar os fótons durante um intervalo de tempo (exposição). Este dispositivo é sensível a incidência de luz e é composto de uma matriz, em que cada elemento dessa matriz é responsável por capturar a informação de um *pixel* da imagem final. A medida que o tamanho da matriz aumenta, maior será o tamanho da imagem e quanto menor o tamanho do elemento do sensor, maior será a resolução da imagem (SEIDEL, 2014). De forma geral o CCD é responsável por capturar a informação do fóton que incide sobre a matriz e libera um elétron.

Costuma-se usar sobre o CCD um filtro de *Bayer*, comumente GRGB, permitindo a passagem de duas componentes verdes para uma componente vermelha e uma componente azul. Segundo Brown (2004), apesar de refletir a maior sensibilidade do olho humano ao verde, o uso do filtro de *Bayer* causa diminuição da resolução, uma vez que cada *pixel* do filtro deixa passar luz de apenas um dos canais Vermelho, Verde e Azul - Red, Green

and Blue (RGB). Outra alternativa utilizada para separar a representação das três cores é utilizar três CCDs por canal (WOOTTON, 2005). Existem outros tipos de técnicas a serem adotadas pelos fabricantes de câmeras. O processo de separação dos espectros visíveis ao olho humano é apresentado na Figura 5.

Figura 5 – Exemplo de captura de imagem para o formato RGB.



Fonte: Seidel (2014).

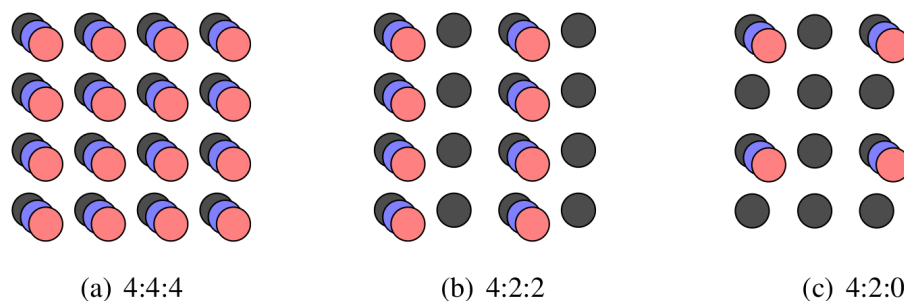
O espaço de cor mais utilizado durante o processo de codificação é o Y' (luma) Cb e Cr (“blue-difference” e “red-difference”) dos componentes de cor ($Y'CbCr$). Mas há diversas formas de codificar $Y'CbCr$ de acordo com recomendações da *International Telecommunication Union Telecommunication Standardization Sector* (ITU-T). Considerando imagem em resolução inferior à *High Definition* (HD), segue-se a recomendação BT.601-7 (2011). No caso de representação de alta definição, adota-se a recomendação BT.709-6 (2015). Atualmente, para os formatos de ultra e alta definição, a nova recomendação a ser adotada é a BT.2020-2 (2015).

No formato $Y'CbCr$, o canal Y' representa a luminosidade, ou seja, a intensidade de luz no ponto. O canal Cb representa a diferença da luminosidade com o canal azul do RGB e o Cr representa a diferença da luminosidade com o canal vermelho. Este es-

paço de cor é adotado, por exemplo, no padrão H.264/AVC, por conter menos correlação entre suas componentes e possibilitar a operação denominada subamostragem croma (RICHARDSON, 2004).

Como o sistema visual humano é mais sensível a luminosidade do que a cor, algumas características da amostragem do vídeo podem ser exploradas. (RICHARDSON, 2004). Assim, ao se adotar a codificação Y'CbCr é possível utilizar resoluções menores nas matrizes de croma (Cb e Cr), com pouco prejuízo para a qualidade subjetiva da imagem. De acordo com a subamostragem efetuada, linhas horizontais ou verticais são eliminadas das matrizes de croma e, quando decodificadas, as amostras mantidas serão interpoladas para gerar matrizes com a mesma resolução. Alguns exemplos de subamostragem croma comumente utilizadas em codificação de vídeo são 4:4:4, na qual não há subamostragem; 4:2:2, na qual amostra-se apenas uma dentre duas linhas ou dentre duas colunas dos canais de croma; e 4:2:0, na qual amostra-se apenas uma linha a cada duas e apenas uma coluna a cada duas dos canais de croma. A Figura 6 ilustra as subamostragens 4:4:4, 4:2:2 e 4:2:0; cada círculo representa um *pixel*, enquanto a “profundidade” representa cada canal de cor: o conjunto de círculos mais escuros (mais ao fundo) representa o canal Y, enquanto os demais representam os canais Cb e Cr.

Figura 6 – Exemplos de subamostragem.



Fonte: Seidel (2014).

Apesar da subamostragem representar uma perda inicial na representação da imagem, pois a reconstrução dos *pixels* removidos não será correta, para codificação de vídeo, dependendo da subamostragem aplicada, podem haver ganhos na compressão. Visto que terá menos *pixels* para representação.

2.3.1 Codificação de vídeo

Como mencionado, para armazenar o vídeo digital é necessário muitos *bits* de informação, ou seja, tem relação direta com o espaço digital para armazenamento. Uma imagem digitalizada é composta por uma matriz de *pixels* (px), cada *pixel* é representado por um número extraído de um conjunto finito predefinido constituído por diferentes

sequências de zeros e uns (*bits*). Quanto maior a resolução do vídeo maior será a quantidade de *bits* necessário para armazenamento. Por exemplo, um vídeo com duração de 10 segundos e resoluções de $3840 \times 1600px$, de largura e altura, necessitaria de aproximadamente *6,2 GB* (*Gigabytes*) para ser armazenado, este cálculo será melhor detalhado na Seção 3.2.

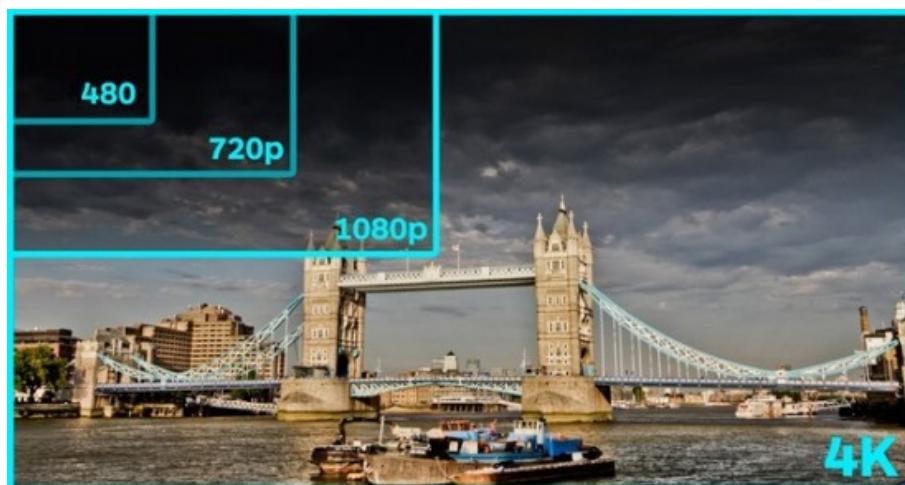
De acordo com Richardson (2011), codificação de vídeo é o processo de compressão e descompressão de um sinal de vídeo digital. Uma característica pertinente no vídeo digital é a presença de informações redundantes, esta característica é explorada pelos algoritmos de codificação de fonte, removendo as redundâncias e resultando em uma representação do vídeo utilizando menos espaço de armazenamento. Outras técnicas também podem ser aplicadas para auxiliar na compressão do vídeo, tal como a estimação de movimento.

2.3.2 Resoluções de vídeos

O vídeo digital é dimensionado conforme alguns grupos de resoluções. Geralmente a padronização das dimensões segue orientações para TV Digital. Contudo, existem resoluções que se adequam apenas para a indústria do cinema.

A resolução no geral define a largura e altura de exibição da imagem. As resoluções são subdivididas em categorias, das quais pode-se citar a *High Definition* (HD ou 720p), *Full HD* (1080p) *4K HD* (2160p) e *8K HD* (4320p). A seguir ilustra a diferença espacial entre algumas resoluções.

Figura 7 – Comparativo entre as resoluções de vídeo.



Fonte: Ynternix (2016).

2.3.3 Padrão H.264/AVC

H.264/AVC (*Advanced Video Coding*) é o padrão de codificação de vídeo do Grupo de Especialistas em Codificação de Vídeo ITU-T (H.264, 2016). Este padrão é otimizado para suportar diversas aplicações, tais como videoconferência, TV a cabo, *streaming* de vídeo, sistema interno de vigilância e etc.

Antes do padrão H.264/AVC surgir, outros padrões já haviam sido criados e já estavam consolidados servindo de base para o desenvolvimento do H.264/AVC. O primeiro padrão relevante para a construção do H.264/AVC foi o H.261. Este padrão serviu como base para os atuais padrões usados na compressão de vídeo: *Differential Pulse Code Modulation* (DPCM) com estimação de movimento na direção temporal, transformada discreta do cosseno aplicada no resíduo e quantização linear seguida de codificação por entropia. Após o padrão H.261, surgiu o padrão MPEG-1 da ISO/IEC seguido do padrão MPEG-2 da ISO/IEC, que também foi padronizado pela ITU-T como H.262. Este padrão se tornou popular e é muito usado até a atualidade em diversas aplicações. Apesar do grande sucesso do padrão MPEG-2, a evolução dos padrões de compressão de vídeo não parou. O padrão H.263 foi lançado e incorporou alguns avanços obtidos pelos padrões MPEG-1 e MPEG-2, bem como técnicas novas que vinham sendo pesquisadas intensamente, tanto pela indústria quanto pela academia (RICHARDSON, 2004).

A grande diferença do H.264/AVC em relação aos seus antecessores refere-se na sua eficiência em comprimir vídeos em altas resoluções, com valores baixos de *bit rate*. Minimizando desta forma o espaço em mídia para armazenamento ou transporte da informação.

O H.264/AVC é usado em HD-DVD e *Blu-ray*, além de ser o formato de vídeo adotado no Brasil para TV digital, MPEG-4 (H.264/AVC). É um dos padrões do estado da arte em compressão de vídeo, devido ao seu aumento em taxa de compressão, em comparação com padrões anteriores. O H.264/AVC é o antecessor do padrão de *High Efficiency Video Coding* (HEVC). De acordo com Seidel (2014), mesmo com o lançamento do HEVC, o H.264/AVC ainda pode ser considerado estado da arte, principalmente no que diz respeito ao consumo energético em plataformas embarcadas mantendo qualidade e baixo *bit rate*.

O estudo dos algoritmos e recursos do CODEC (Coder-Decoder) H.264/AVC não serão detalhados nesta dissertação. Contudo, é interessante ter o conhecimento como os *frames* são agrupados para exibição.

2.3.3.1 Group of pictures

Um fator muito importante que influencia na qualidade do vídeo é o tipo de *frame*. Existem três tipos definidos de *frames*: I, P e B.

Frame I (intra) são codificados sem referência a outros *frames* (sem qualquer compensação de movimento), de forma muito semelhante ao JPEG (*Joint Photographic Experts Group*), o que significa que eles contêm todas as informações necessárias para sua reconstrução pelo decodificador. Por este motivo, são o ponto de entrada essencial para o acesso a uma sequência de vídeo. Um *frame* I é usado como uma referência para *frames* previstos (P e B). A taxa de compressão de quadros I é relativamente baixa.

Os *frames* P (predito) são inter-codificados utilizando a predição compensada de movimento de um *frame* de referência (*frame* P ou *frame* I que precede o *frame* P atual). Assim, um *frame* P é prevista usando a previsão de avanço, e o mesmo pode ser utilizado como referência para outros *frames* preditos (*frame* P e B). A taxa de compressão dos *frames* P é significativamente maior do que de *frames* I.

Os *frames* B são inter-codificados usando a predição compensada por movimento a partir de dois *frames* de referência, os *frames* P e/ou I antes e depois da *frame* B atual. Dois vetores de movimento são gerados para cada macrobloco em um *frame* B - uma apontando para uma área de correspondência na imagem de referência anterior (um vetor de frente) e um apontando para uma área de correspondência na futura imagem de referência (um vetor de retrocesso). Um macrobloco de predição compensado por movimento pode ser formado de três maneiras: a predição direta usando o vetor direto, a predição inversa usando o vetor de retrocesso ou a predição bidirecional (onde a referência de predição é formada fazendo a média das referências de previsão para frente e para trás). Tipicamente, um codificador escolhe o modo de predição (para frente, para trás ou bidirecional) que fornece a energia mais baixa para o macrobloco de diferença. Os *frames* B oferecem a maior taxa de compressão.

Todos esses diferentes tipos de *frames* (I, P e B) são então agrupados em uma sequência (ordem de repetição específica), chamado *Group of pictures* (GOP). Um GOP deve sempre começar com um *frame* I e pode conter apenas I ou uma combinação de I e P ou I, P e B *frames*. A utilização e também um número de *frames* B ou P dentro de um GOP pode ser aumentado ou diminuído dependendo do conteúdo da imagem, taxa de compressão ou aplicação que o vídeo comprimido é destinado. Vários comprimentos de GOP e combinações de *frames* P e B podem ser codificados (BOVIK, 2010; UHRINA; HLUBIK; VACULIK, 2012a; UHRINA; HLUBIK; VACULIK, 2012b).

2.4 Avaliação objetiva

Pelo fato deste trabalho ser focado na qualidade de vídeo, algumas métricas serão abordadas. Tem-se duas formas de avaliar a qualidade de um vídeo, objetiva ou subjetiva. Uma avaliação subjetiva exige recursos humanos além de um treinamento especializado da mesma, tornando esta opção onerosa. Na avaliação objetiva, tem-se alguns modelos

matemáticos consolidados na literatura para atestar a qualidade de uma imagem.

2.4.1 Peak Signal to Noise Ratio

O *Peak Signal to Noise Ratio* (PSNR) é a métrica objetiva utilizada para avaliar a qualidade do vídeo recebido pelo usuário comparando com o vídeo original. O PSNR é mensurado em escala logarítmica e depende de outra métrica denominada de *Mean Squared Error* (MSE) entre o vídeo original e o vídeo estimado (DUBOLIA et al., 2011). O MSE e o PSNR são definidos pelas Equações 2.1 e 2.2, respectivamente:

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I_{(i,j)} - K_{(i,j)}]^2 \quad (2.1)$$

onde:

MSE - *Mean Squared Error*.

$I_{(i,j)}$ - Matriz que compõe um quadro transmitido.

$K_{(i,j)}$ - Matriz que compõe um quadro recebido.

M e N - Representam a dimensão do quadro.

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (2.2)$$

onde $(2^n - 1)^2$ representa o quadrado do maior valor possível de sinal na imagem, em que n é o número de *bits* necessários para representar um *pixel* da imagem.

Neste trabalho utiliza-se a média da PSNR para qualificar o vídeo, pois o vídeo é uma sucessão de imagens no domínio do tempo. Logo, calcula-se todos os valores de PSNR por *frame* e tira-se a média.

2.4.2 Structural Similarity Index

A métrica *Structural Similarity Index* (SSIM) é baseada na medição de três componentes similaridade de luminância, similaridade de contraste e similaridade estrutural. Esses três componentes são combinados, e então é gerado um único valor chamado de índice. Assim, pode-se dizer que a métrica representa as informações das imagens e vídeos, pois considera as características que possuem mais prioridade, assim como o sistema de visão humana. O índice varia de zero a um, onde quanto mais próximo de um, maior a

qualidade e índices próximos de 0, menor é a qualidade. A métrica SSIM é mostrada na Equação 2.3 dada a seguir:

$$SSIM_{(x,y)} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2.3)$$

onde:

μ_x e μ_y são as médias de x e y , respectivamente. Onde x e y são as dimensões do quadro.

σ_x^2 e σ_y^2 são as variâncias de x e y , respectivamente.

σ_{xy} é a variancia de x e y .

c_1 e c_2 são variáveis para estabilizar a divisão por um mínimo.

2.5 Root Mean Square Error

A *Root Mean Square Error* (RMSE) é dada pela Equação 2.4, é uma medida utilizada neste trabalho para avaliar o quão próximo dos pontos teóricos (estimados) os pontos medidos ficaram. Onde $x_{1,i}$ são os valores teóricos e $x_{2,i}$ são os pontos medidos.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{1,i} - x_{2,i})^2}{n}} \quad (2.4)$$

Nesta pesquisa também é utilizado o RMSE para avaliar a modelagem proposta para os dados simulados. Ou seja, identificar qual a melhor representação matemática para os dados simulados.

2.6 Considerações finais

O *streaming* de vídeo envolve diversas tecnologias, entre elas, a tecnologia de acesso a rede comunicação (802.11ac) e o CODEC (H.264/AVC) de vídeo. No próximo capítulo será detalhado todo o processo de *streaming* de vídeo, os recursos, as ferramentas e a metodologia para captura das informações geradas neste trabalho.

3 Ferramentas utilizadas

3.1 Considerações iniciais

Neste capítulo será descrito o *streaming* de vídeo ou transmissão de vídeo. Em que quatro diferentes vídeos são codificados com o H.264/AVC e enviados sobre a rede 802.11ac.

3.2 Vídeos sem compressão

Os vídeos utilizados nesta dissertação foram obtidos da Foundation (2016), organização sem fins lucrativos que disponibiliza uma base de multimídia para utilização no desenvolvimento livre de protocolos abertos e softwares para servir os mercados públicos e de negócios. Dentre os vídeos disponibilizados pela fundação escolheu-se os: “crowd_run”, “ducks_take_off”, “in_to_tree” e “park_joy”. Todos os vídeos estão no formato YUV (sem compressão); nas resoluções de 2160p, 1080p e 720p e; com taxa de amostragem de 4:2:0, conforme descrito por Foundation (2016). Possuem 500 *frames* com duração de 10 segundos, ou seja, com taxa de quadros igual a 50 *fps*. Na Tabela 2 apresenta algumas informações sobre as resoluções estudadas.

Tabela 2 – Resoluções

Resoluções	Largura (Pixel)	Altura (Pixel)
2160p	3840	2160
1080p	1920	1080
720p	1280	720

Fonte: Foundation (2016)

Para se determinar o tamanho em *bytes* para armazenamento de cada vídeo é preciso saber a taxa de amostragem, a quantidade de *bits* para representar e a resolução, conforme Equação 3.1. Esta equação é válida para YUV com taxa de amostragem igual a 4:2:0 (RICHARDSON, 2004). O tamanho de cada arquivo por resolução é apresentado na Tabela 3.

$$S = \frac{[(L \times A \times b) + (\frac{L}{2} \times \frac{A}{2} \times b \times 2)] \times N_f}{8} \quad (3.1)$$

onde:

L - Largura do vídeo em *pixels*.

A - Altura do vídeo em *pixels*.

N_f - Quantidade de *frames* do vídeo.

b - *bits* para representação.

Tabela 3 – Espaço para armazenamento dos arquivos YUV por resolução.

Resolução	Tamanho (<i>bytes</i>)
2160p	6220800000
1080p	1555200000
720p	691200000

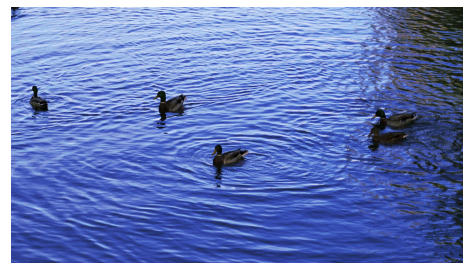
O tamanho do arquivo YUV é importante para avaliar a eficiência do CODEC após compressão do vídeo.

Os vídeos utilizados neste trabalho estão representados na Figura 8, onde apenas um *frame* dos 500 foi extraído. O vídeo da Figura 8a apresenta grande movimentação, as Figuras 8b e 8d apresentam movimentação leve, enquanto que a Figura 8c apresenta movimentação moderada. Esta classificação foi obtida após análise visual por meio da reprodução dos vídeos.

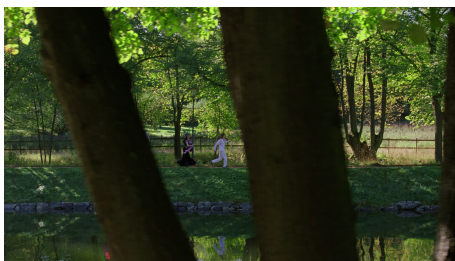
Figura 8 – Vídeos utilizados.



(a) Crowd Run



(b) Ducks Take Off



(c) Park Joy



(d) Into Tree

Fonte: Foundation (2016)

3.3 Compressão de vídeo

A compressão de vídeo é o processo muito importante para diminuir a quantidade de informação a ser representada. Os vídeos utilizados nesta dissertação foram codificados na configuração presente na Tabela 4. As extensões dos arquivos codificados até essa etapa tem o formato *m4v*. O CODEC utilizado foi o H.264/AVC com o perfil *High@L5*

Para condicionar os vídeos para transmissão sobre a rede é preciso adicionar algumas informações sobre seu comportamento na rede, ou seja, informar para a camada de rede como o vídeo deverá ser transmitido. Neste trabalho o protocolo escolhido foi o *User Datagram Protocol* (UDP), o tamanho dos pacotes é de *1024 bytes*. Mais detalhes sobre os processos de codificação de fonte, empacotamento, transmissão, reconstrução do vídeo serão abordados na Seção 3.4.

Tabela 4 – Configuração do codificador de fonte.

Codificador	H.264/AVC
Bit rate	32 Mbps
GOP	25
FPS	50

Fonte: o autor.

Nesta configuração o espaço necessário para armazenar os vídeos estão presentes na Tabela 5. Comparando com os tamanhos “brutos” (Tabela 3), os vídeos em 2160p obtiveram em média uma eficiência na compressão em torno de 138 vezes, já para 1080p de 36 vezes e para 720p de 16 vezes. Demonstrando-se a eficiência do H.264/AVC para ultra e alta resoluções sobre os vídeos estudados.

Tabela 5 – Tamanhos dos vídeos.

Nome do Vídeo	Resolução		
	720p	1080p	2160p
Crowd Run	43,0 MB	43,2 MB	43,5 MB
DucksTake Off	43,1 MB	43,3 MB	46,0 MB
Park Joy	42,0 MB	42,6 MB	42,6 MB
In to Tree	44,2 MB	45,1 MB	47,6 MB

Fonte: o autor.

Para cada vídeo a ser transmitido calcula-se a média de PSNR, sendo de grande importância para determinar quanto o vídeo foi degradado após a recepção no usuário final. Nesta dissertação, estes valores são considerados como os valores de PSNR de referências para o cálculo da perda média de PSNR. Os valores obtidos da média de PSNR por vídeo são apresentadas na Tabela 6.

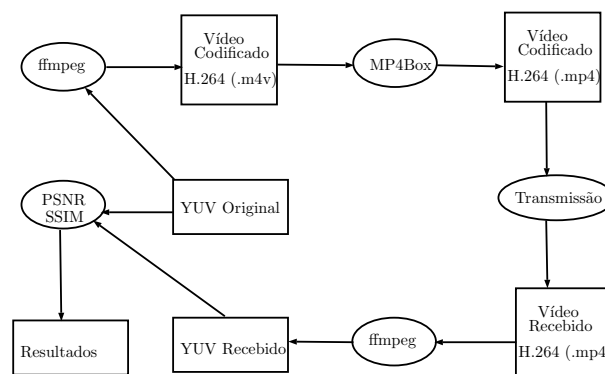
Tabela 6 – Média de PSNR por vídeo.

Nome do Vídeo	PSNR (dB)		
	720p	1080p	2160p
Crowd Run	38,45	35,82	31,66
DucksTake Off	36,22	30,69	27,4
Park Joy	37,17	31,58	30,13
In to Tree	40,97	36,93	34,3

Fonte: o autor.

Percebe-se que existe variação entre os valores da média de PSNR quando se analisa o vídeo em relação a resolução. Uma característica comum que pode inferir é que a resolução e PSNR são inversamente proporcionais, considerando que os vídeos são codificados em condições iguais (Tabela 4). Esta característica tem influência direta com o valor de *bit rate* escolhido, visto que quanto maior o seu valor, maior será a qualidade do vídeo (maiores valores da média de PSNR). É importante citar que neste momento as únicas perdas presentes são as perdas por codificação de fonte (compressão de vídeo). A seguir serão descritas as etapas para transmissão do vídeo, e de forma geral todo o processo pode ser descrito pela Figura 9.

Figura 9 – Diagrama de codificação, transmissão, recepção e reconstrução.



Fonte: o autor

3.4 Etapas para transmissão do vídeo

Como descrito anteriormente, para transmitir o vídeo é necessário primeiro condicionar a informação. Desta forma, os comandos usados para gerar o arquivo no formato *mp4* estão no Código 3.1, utilizando a ferramenta FFMPEG[®] (FFMPEG, 2016). O formato do arquivo “bruto” está na extensão *y4m*, o qual carrega algumas informações tais como resolução e a taxa de *Frames Per Second* (FPS) que o vídeo deve ser codificado, por esta razão o parâmetro de resolução não é informado ao FFMPEG. A ferramenta MP4Box empacota o vídeo para transmissão do vídeo sobre a rede (MP4BOX, 2016).

Código 3.1 – Setup codificação de vídeo.

```

1 ffmpeg -i YUV_Fonte.y4m -r 50 -b:v 32M -bt 4800 -c:v libx264 -g 25 destino.
   m4v
2 MP4Box -hint -mtu 1024 -fps 50 -add destino.m4v destino.mp4

```

Em uma análise geral da primeira linha do Código 3.1, a instrução basicamente está passando ao FFMPEG que ele crie um arquivo na extensão *m4v* com: 50 FPS, *bit rate* de 32 *Mbps* com tolerância de 4800 *bps*, no codificador H.264/AVC, com o tamanho do GOP de 25. Já a segunda linha, o MP4Box está retornando um arquivo na extensão *mp4*, adicionando um *hint*, definindo o tamanho do pacote para 1024 *bytes* e informando que o FPS é de 50.

Para realizar a transmissão do vídeo é preciso determinar quem estará enviando o vídeo e quem receberá, servidor e cliente respectivamente. Em ambos utiliza-se a ferramenta TCPDUMP (TCPDUMP, 2016). O servidor irá transmitir o vídeo utilizando uma das aplicações disponíveis no *framework* Evalvid, o mp4trace. Primeiramente é preciso executar no servidor e cliente as instruções conforme os Código 3.2 e Código 3.3, respectivamente. A instrução do Código 3.4 deve ser executada no servidor é a responsável por enviar o vídeo.

Código 3.2 – Servidor transmissão.

```

1 sudo tcpdump -n -tt -v udp port PORTA > sd.txt

```

Código 3.3 – Cliente transmissão.

```

1 nc -ul PORTA &
2 sudo tcpdump -n -tt -v udp port PORTA > rd.txt
3 sudo killall nc

```

Código 3.4 – Enviar transmissão.

```

1 mp4trace -f -t UDP -s IP_Destino PORTA destino.mp4 > st.txt

```

Após a transmissão de um determinado vídeo deverá ser feita a interrupção da captura de informação do TCPDUMP, esta interrupção deve ser feita pressionando-se as teclas *Ctrl+c*. Ao final deste processo será gerado três arquivos: “sd.txt”, “rd.txt” e “st.txt”. Estes três arquivos são utilizados para reconstruir o vídeo transmitido por meio do etmp4, outra aplicação do *framework* Evalvid, conforme tabela Tabela 7.

Tabela 7 – Descrição dos arquivos gerados na transmissão

sd.txt	Dump do remetente
rd.txt	Dump do receptor
st.txt	Trace do remetente

Fonte: Klaue, Rathke e Wolisz (2003).

Para reconstruir o vídeo recebido é preciso executar a instrução do Código 3.5. O `etmp4` gera após a reconstrução do vídeo quatro arquivos de *log*, os quais são descritos na Tabela 8.

Código 3.5 – Reconstrução do vídeo.

```
1 etmp4 -f -x sd.txt rd.txt st.txt destino.mp4 reconstruido.mp4
```

Tabela 8 – Descrição do log do `etmp4`.

loss_reconstruido.txt	Contém a perda de I, P, B e total em %
delay_reconstruido.txt	Contém frame-nr., lost-flag, end-to-end delay , inter-frame gap sender, inter-frame gap receiver, and cumulative jitter em segundos
rate_s_reconstruido.txt	Contém tempo, bytes por segundo (intervalo de tempo atual) e bytes por segundo (cumulativo) medido no remetente
rate_r_reconstruido.txt	Contém tempo, bytes por segundo (intervalo de tempo atual) e bytes por segundo (cumulativo) medidos no receptor

Fonte: Klaue, Rathke e Wolisz (2003)

A partir do arquivo reconstruído no formato *mp4* é feita a conversão para o formato YUV utilizando a instrução do Código 3.6. Esta conversão é necessária para obtenção dos valores de PSNR, visto que a aplicação `psnr` do *framework* Evalvid trabalha somente com os arquivos com extensão YUV.

Código 3.6 – Conversão *mp4* para YUV.

```
1 ffmpeg -i reconstruido.mp4 reconstruido.yuv
```

Com os arquivos no formato YUV (fonte e reconstruído) é possível calcular os valores de PSNR e SSIM dos vídeos recebidos (KLAUE; RATHKE; WOLISZ, 2003), conforme os Código 3.7 e Código 3.8, respectivamente. Os arquivos “`psnr_reconstruido.txt`” e “`ssim_reconstruido.txt`”, contém os valores de PSNR e SSIM por *frame* do vídeo recebido.

Código 3.7 – Obtenção da PSNR.

```
1 psnr Largura Altura 420 YUV_fonte.yuv reconstruido.yuv > psnr_reconstruido.txt
```

Código 3.8 – Obtenção da SSIM.

```
1 psnr Largura Altura 420 YUV_fonte.yuv reconstruido.yuv ssim > ssim_reconstruido.txt
```


3.5 Considerações finais

Neste capítulo foram apresentadas todas as fases para codificação de fonte e condicionamento do vídeo a ser transmitido sobre a rede de comunicação. Assim como uma descrição detalhada sobre o processo de transmissão do vídeo. A seguir será apresentada a metodologia adota para o desenvolvimento da proposta de modelagem.

4 Metodologia

4.1 Considerações iniciais

Para o desenvolvimento da proposta de modelagem é preciso primeiro entender o comportamento da variável alvo, neste caso, a média da perda de PSNR. O *streaming* de vídeo é caracterizado como uma comunicação unidirecional (FRNDA; VOZNAK; SEVCIK, 2016), desta forma, pretende-se modelar o comportamento médio da perda de PSNR em função da perda de pacotes na transmissão de vídeo. A seguir será explorada a metodologia aplicada neste trabalho.

4.2 Coleta dos dados

Neste trabalho, para a coleta dos dados utilizou-se um ponto de acesso (Figura 10) operando no padrão *IEEE 802.11ac*, dois *notebooks* e um adaptador *wireless* USB (Figura 11), cada componente está descrita na Tabela 9. Inicialmente, as transmissões dos vídeos ocorrem em condições ideais (sem perda). Para verificar se a informação coletada foi válida, a quantidade de pacotes enviados e recebidos são comparadas e caso haja igualdade nas informações pode-se garantir que a transmissão foi realizada sem perda (KLAUE; RATHKE; WOLISZ, 2003).

Tabela 9 – Dispositivos utilizados nas transmissões.

Dispositivo	Marca	Modelo	Descrição
Roteador	D-LINK	DIR-803	Wireless AC750 Dual Band
Adptador wireless USB	D-LINK	DWA-171	WirelessAC Dual-Band Nano USB Adapter
Notebook 1 (Servidor)	DELL	VOSTRO-3460	Ubuntu 16.04 LTS
Notebook 2 (Cliente)	DELL	INSPIRON-3421	Ubuntu 16.04 LTS

Figura 10 – Roteador D-Link DIR 803.



Fonte: o autor.

Figura 11 – Adaptador USB DWA-171.

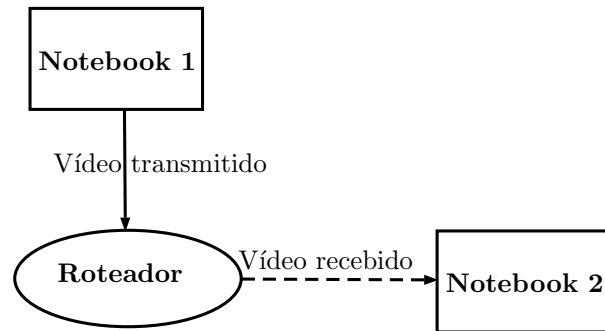


Fonte: o autor.

Para ilustrar o cenário em que os vídeos foram enviados, observe a Figura 12. Neste cenário a transmissão do vídeo é realizada por meio de uma conexão cabeada com o

roteador e a recepção do vídeo é pelo meio sem fio utilizando o padrão 802.11ac. A distância entre o *Roteador* e o *Notebook 2* depende do propósito da transmissão, inicialmente para evitar perda de pacotes (transmissão ideal), posiciona-se o *Notebook 2* relativamente próximo do *Roteador*.

Figura 12 – Esquema de transmissão.



Fonte: o autor

4.2.1 Transmissão ideal

Como mencionado anteriormente, em um cenário ideal a quantidade de pacotes enviados e recebidos deve ser a mesma. Esta etapa é importante, visto que para simular as perdas de pacotes no cliente é preciso se ter os arquivos não comprometidos. Neste momento foram realizadas doze transmissões ideais, obtendo-se para cada uma os três arquivos necessários para a reconstrução do vídeo recebido (conforme Seção 3.4). Destes três arquivos compara-se o “sd.txt” e o “rd.txt” de cada diferente transmissão, os quais devem ter o mesmo número de pacotes enviados e recebidos, respectivamente. Na Tabela 10 é apresentado os valores de pacotes enviados e recebido para cada transmissão.

Tabela 10 – Número de pacotes por transmissão.

Vídeo	Resolução	Número de Pacotes	
		Enviado	Recebido
Crowd Run	720p	40361	40361
	1080p	40562	40562
	2160p	40875	40875
DucksTake Off	720p	40488	40488
	1080p	40674	40674
	2160p	43199	43199
Park Joy	720p	39488	39488
	1080p	40003	40003
	2160p	40068	40068
In to Tree	720p	41544	41544
	1080p	42424	42424
	2160p	44754	44754

Fonte: o autor.

Foi verificado que durante algumas transmissões é possível que o número de pacotes enviados seja menor do que os valores da coluna “enviado” da Tabela 10, desta forma impossibilitando a reconstrução do vídeo por meio da aplicação *etmp4* do *framework* Evalvid. Portanto, é preciso primeiramente determinar qual o número correto de pacotes enviados, esta informação pode ser determinada através de algumas tentativas de transmissões e reconstruções do vídeo. Caso ocorra algum erro durante a reconstrução, o *etmp4* irá acusar que o arquivo “sd.txt” está incompleto, desta forma pode-se concluir que o arquivo que contém as informações sobre os pacotes enviados não pode ser utilizado para representar a quantidade de pacotes enviados. A informação do número de pacotes é exibida ao término da aplicação do *tcpdump* (pressionando Ctrl+c). Outra alternativa para determinar a quantidade de pacotes enviados seria contabilizar a quantidade de linhas que o arquivo “rd.txt” contém.

A partir da quantidade de pacotes enviados é possível determinar se a transmissão do vídeo foi ideal (sem perdas), simplesmente subtraindo a quantidade de pacotes enviados pelo o recebido, se o resultado for zero pode-se concluir que não houve perdas.

Ao final haverão 12 combinações diferentes dos arquivos: “sd.txt”, “rd.txt” e “st.txt”; os são necessários para a reconstrução dos vídeos recebidos. Para simular as perdas no receptor somente será necessário os arquivos do receptor (“rd.txt”).

4.2.2 Simulação de perda de pacotes

A partir das informações coletadas em condições ideais (sem perda). A perda de pacotes no receptor é simulada. Como mencionado por Evalvid (2016), uma das possibilidades de gerar erro durante uma transmissão seria excluindo manualmente algumas

linhas do arquivo “rd.txt”. Desta forma, adiciona-se perdas determinadas na recepção dos pacotes. Esta metodologia é utilizada para simular a perda de pacotes na recepção do vídeo.

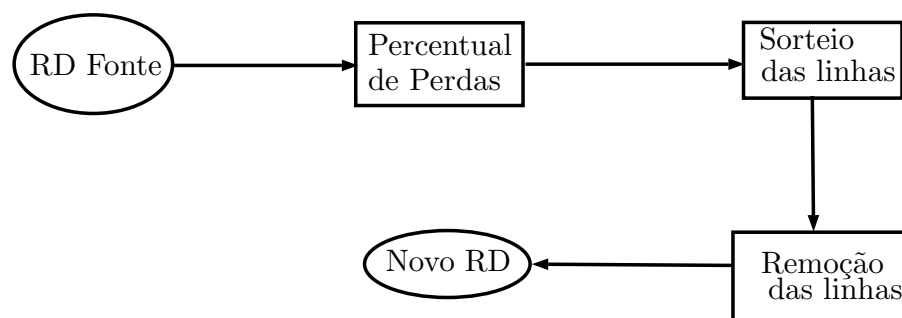
Para tornar as perdas de pacotes simuladas o mais próximo do ambiente real (aleatório), utilizou-se a função de distribuição de probabilidade uniforme discreta para sortear as linhas que serão removidas no arquivo do receptor, conforme Equação 4.1. Desta forma garantido que a probabilidade de uma linha ser removida é igual para todas as demais. Esta foi a função adotada para descrever a perda de pacotes nesta dissertação.

$$f(x) = \begin{cases} \frac{1}{n}, & \text{se } 1 \leq x \leq n; \\ 0, & \text{caso contrário} \end{cases} \quad (4.1)$$

Neste trabalho, foi estabelecido que a perda de pacotes inicial (em porcentagem) é de 2% variando-se de 2% até 30%. Estas perdas em porcentagem são simuladas para cada arquivo do receptor, ou seja, uma simulação corresponde a 15 valores de porcentagens de perdas de pacotes para cada “rd.txt” diferente. Desta forma gerando 15 arquivos do receptor para uma determinada configuração de vídeo (vídeo e resolução).

O software utilizado para realizar as simulações foi o MATLAB[®] (MATLAB, 2015), utilizando a função *randperm* para realizar o sorteio das linhas a serem removidas do arquivo fonte. Para garantir a aleatoriedade dos sorteios utilizou-se a função *rng(‘shuffle’)*, fazendo com que sempre que a simulação fosse executada a “semente” do sorteio fosse resetada (JAIN, 1991), evitando a repetição das linhas durante cada diferente sorteio. Este processo resultou no total de 50 simulações para cada configuração de vídeo e resolução. Para facilitar o entendimento da simulação veja a Figura 13

Figura 13 – Sorteio das linhas.



Fonte: o autor

4.2.3 Reconstrução dos vídeos

No total 9000 amostras foram geradas a partir das 50 simulações. Dais quais, 750 amostras correspondem a configuração de um vídeo para uma resolução, como se tem 12

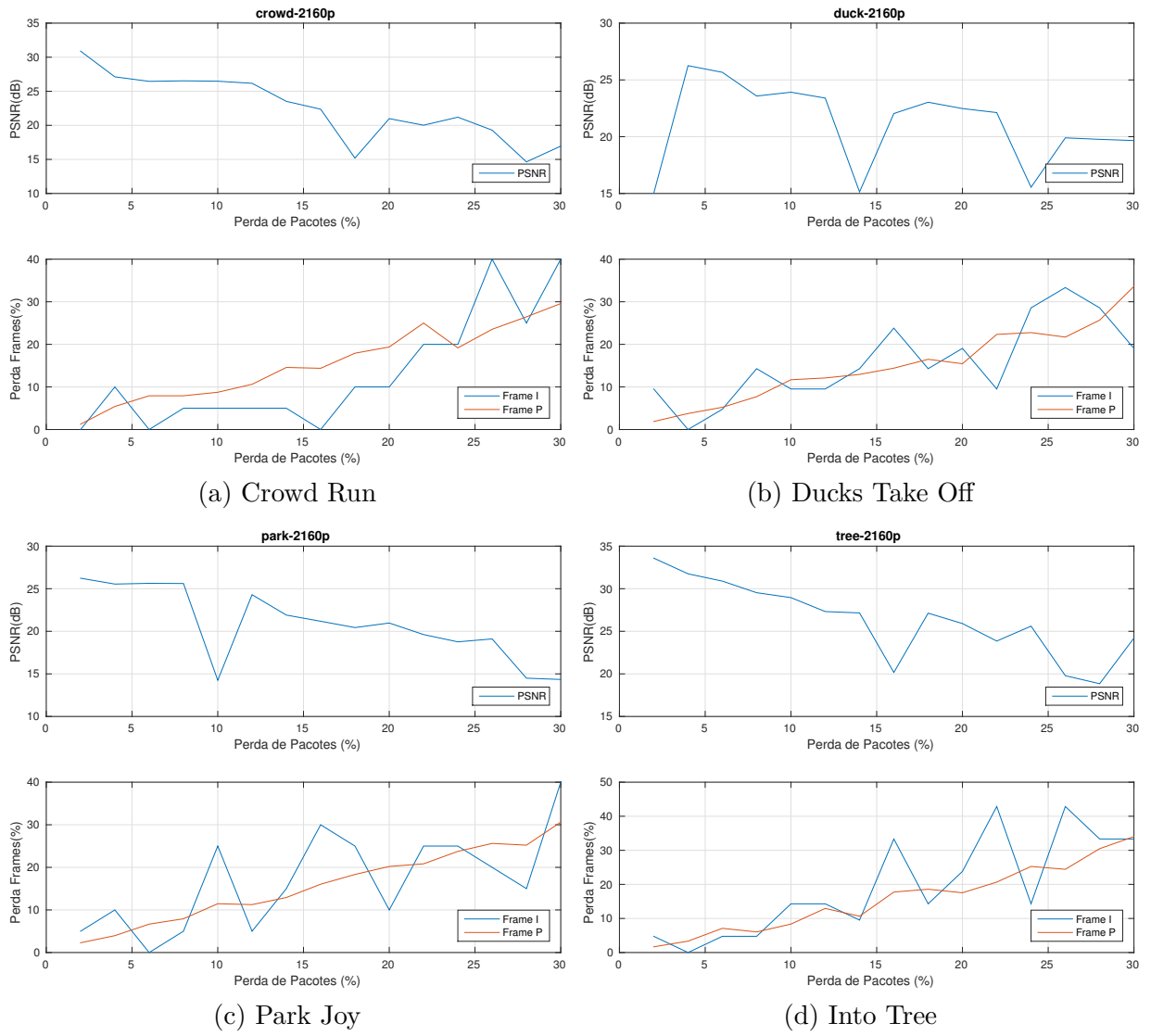
configurações (vídeo e resolução), logo se tem 9000 amostras.

Destas amostras foram extraídos os resultados a partir dos processos de reconstrução (Código 3.5), conversão para YUV (Código 3.6), PSNR (Código 3.7) e SSIM ((Código 3.7)) dos vídeos. Dos resultados destacam-se os valores da média de PSNR, SSIM, perdas de *frames* I e P, os quais serão abordados a seguir.

4.2.3.1 Valores de PSNR nos vídeos

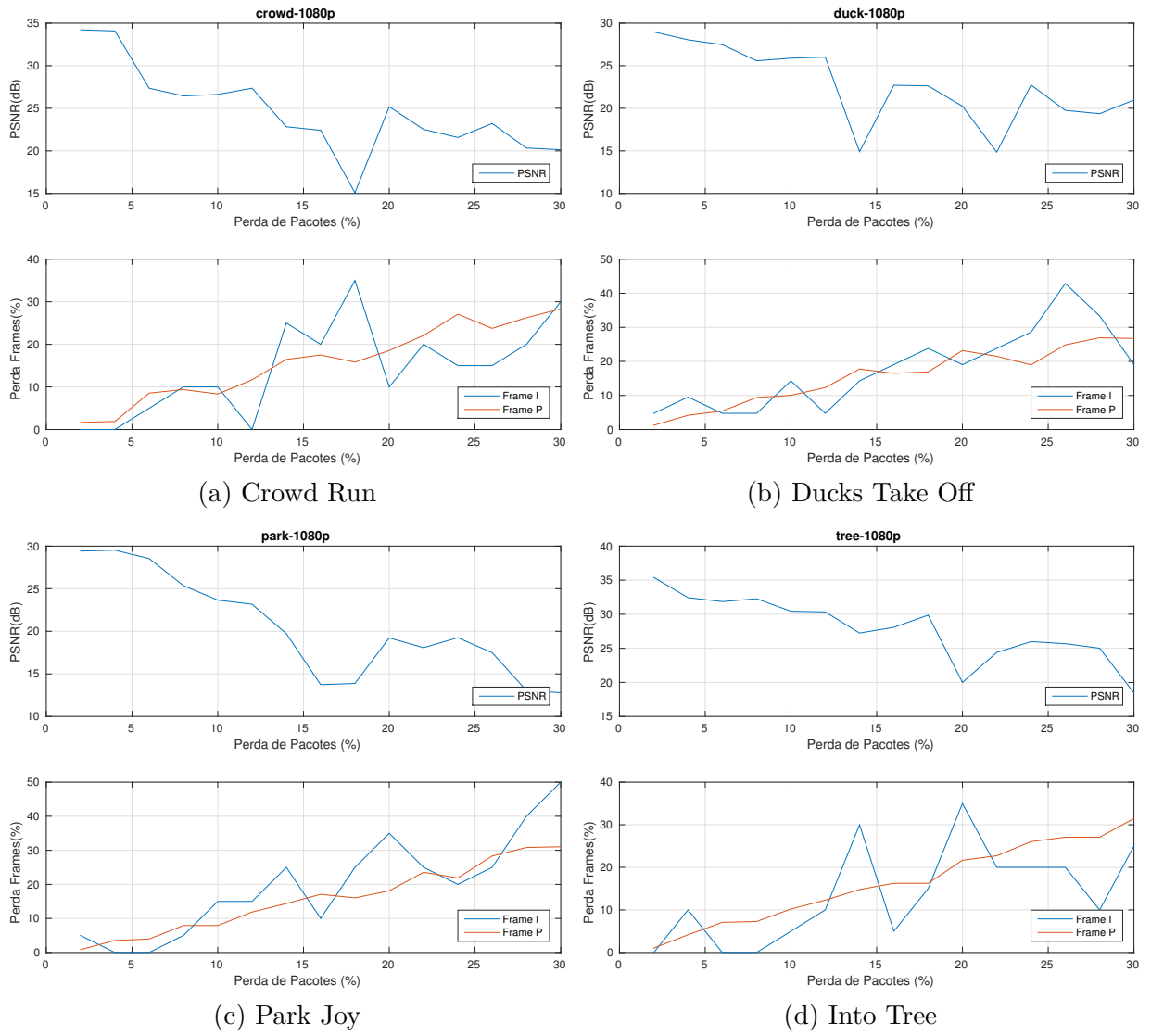
Para facilitar o entendimento de como as amostras estão distribuídas, a Figura 14a apresenta as informações relevantes. Esta figura contém a informação dos valores médios de PSNR para cada valor de perda de pacotes para a simulação 19, também é apresentada a porcentagem de perdas de frames I e P para o vídeo Crowd Run na resolução de 2160p. A mesma interpretação pode ser feita para as Figura 15 e Figura 16.

Figura 14 – Valores de PSNR da simulação 19 em 2160p.



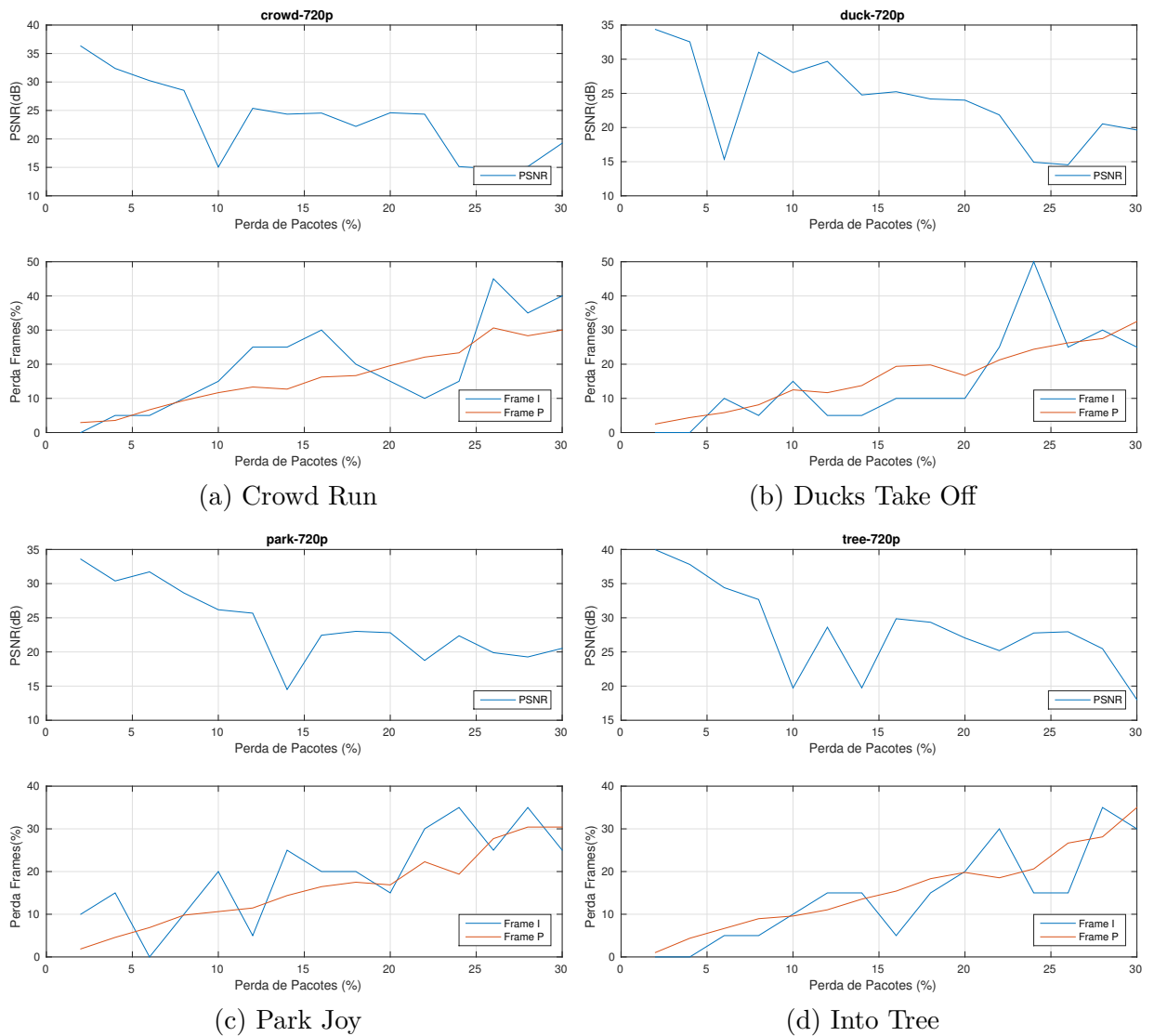
Fonte: o autor.

Figura 15 – Valores de PSNR da simulação 19 em 1080p.



Fonte: o autor.

Figura 16 – Valores de PSNR da simulação 19 em 720p.



Fonte: o autor.

No geral, pode-se inferir que a medida que a porcentagem de perda de pacotes aumenta, os valores da média de PSNR tendem a diminuir, ao contrário da porcentagem de perdas de I e P que tende a aumentar. Estes últimos são objetos de estudo para traçar um perfil de perda de *frames* que influência no valor da média de PSNR. Como exemplo veja as Figura 17 e Figura 18, onde apresenta-se a variação da PSNR para o *Frame I* e *P*, respectivamente. Contudo, o objeto de estudo desta dissertação é a influência da perda de pacotes na perda da qualidade de vídeo em diferentes resoluções.

Figura 17 – Valores de PSNR da simulação 23 Crowd Run 720p Frame 201 (I).



Fonte: o autor.

Figura 18 – Valores de PSNR da simulação 23 Crowd Run 720p Frame 199 (P).

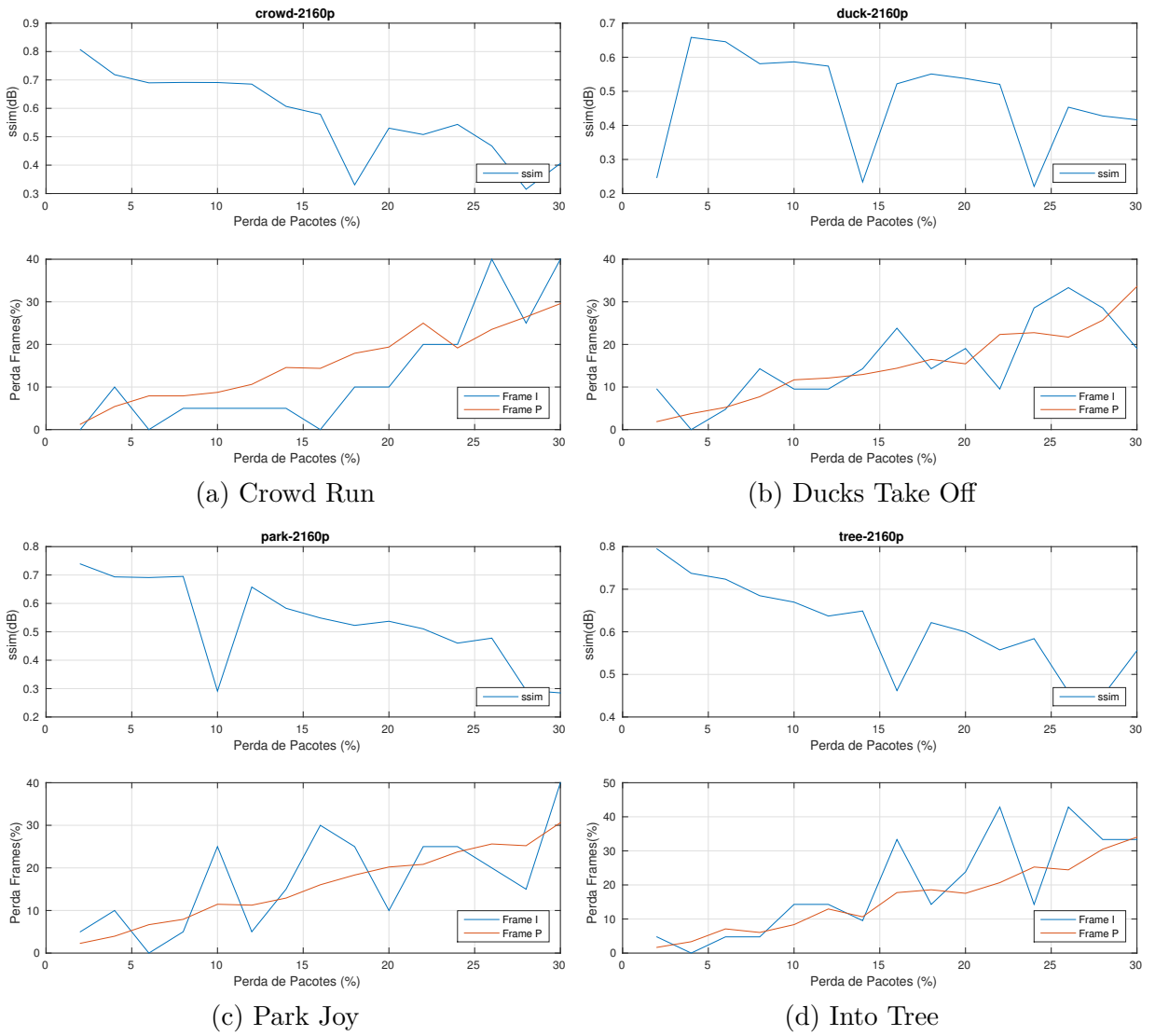


Fonte: o autor.

4.2.3.2 Valores de SSIM nas simulações

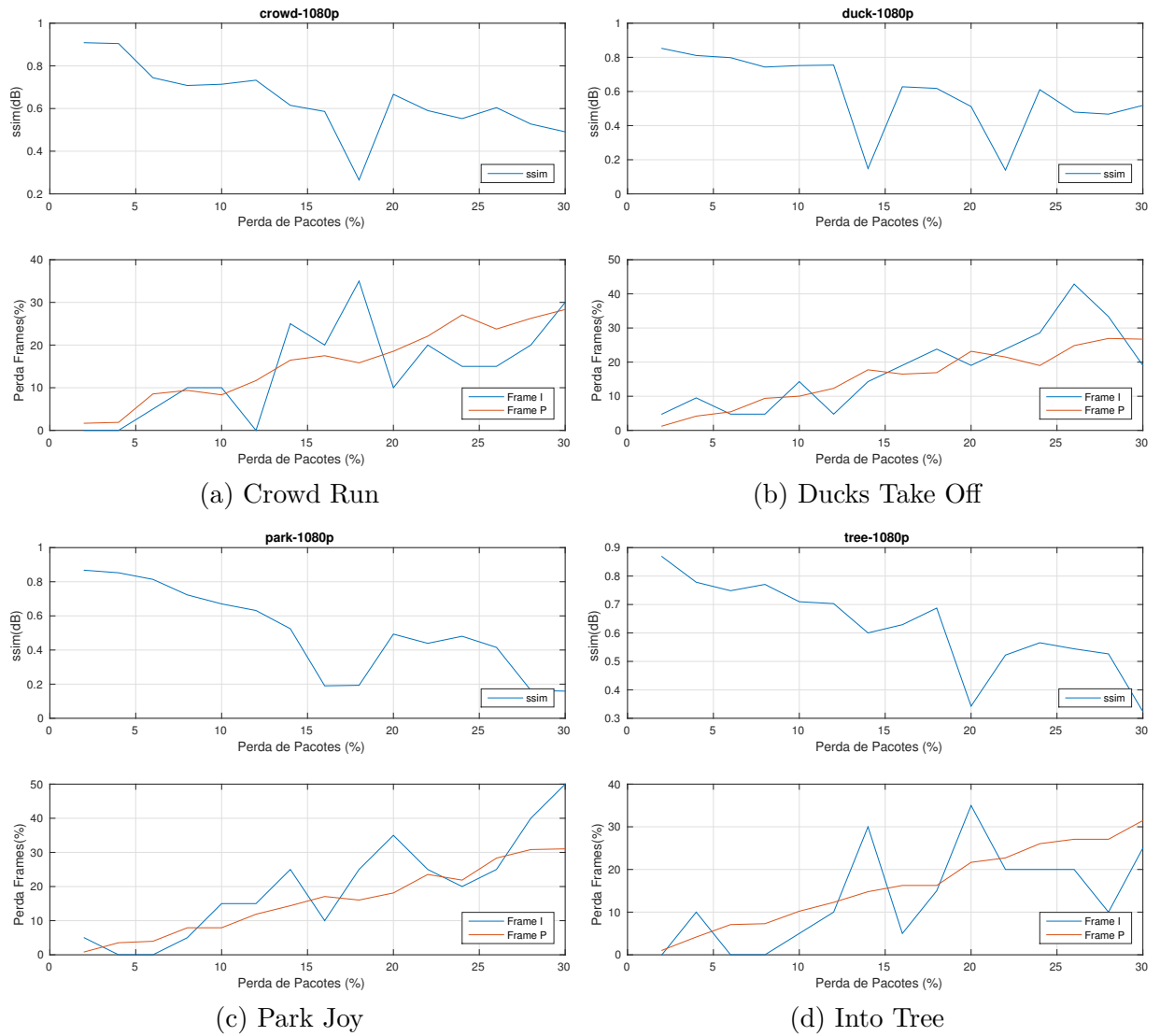
Outra métrica obtida dos resultados das simulações é o SSIM. Esta métrica não irá fazer parte dos parâmetros de estudo desta dissertação. Ela é mencionada somente para fins comparativos PSNR dos vídeos.

Figura 19 – Valores de SSIM da simulação 19 em 2160p.



Fonte: o autor.

Figura 20 – Valores de SSIM da simulação 19 em 1080p.



Fonte: o autor.

No geral, pode-se inferir que a medida que a porcentagem de perda de pacotes aumenta os valores da média de SSIM tendem a diminuir, ao contrário da porcentagem de perdas de I e P que tende a aumentar. Ou seja, o mesmo comportamento da PNSR.

4.3 Modelagem

A estratégia escolhida para dar início a modelagem foi obter a informação objeto de estudo desta dissertação, a perda média de PSNR. A qual é obtida utilizando-se a Equação 4.2.

$$L_{PSNR} = L_{ref} - L_{sim} \quad (4.2)$$

onde:

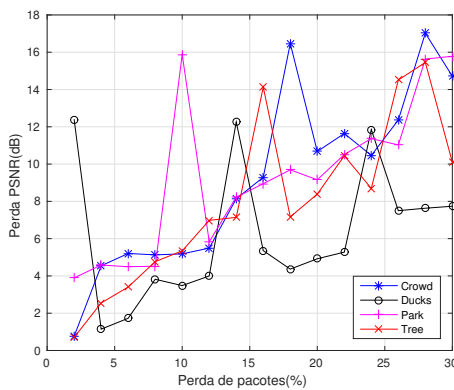
L_{PSNR} perda média de PSNR em dB ;

L_{ref} valor médio de PSNR do vídeo (transmitido) em dB ;

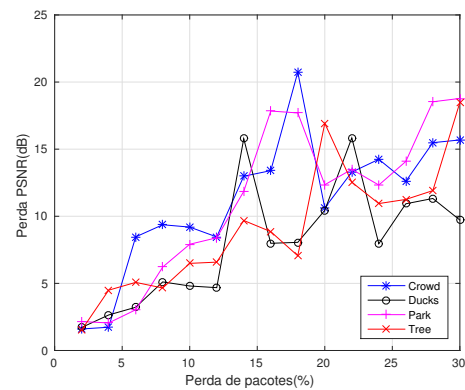
L_{sim} valor médio de PSNR do vídeo (recebido/simulado) em dB .

Os valores das L_{ref} de cada vídeo estudado segue na Tabela 6. Para cada configuração de vídeo e resolução obteve-se a L_{PSNR} referente a cada uma das 50 simulações, os valores obtidos para a simulação 19 são apresentados na Figura 21.

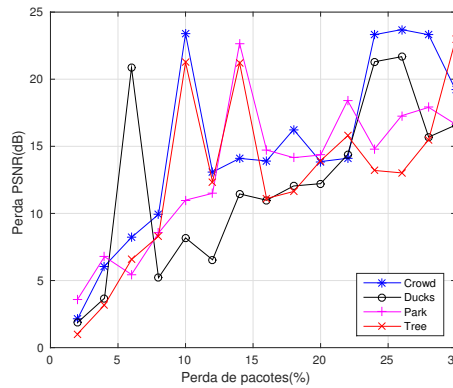
Figura 21 – Média das perdas de PSNR para simulação 19.



(a) 2160p



(b) 1080p



(c) 720p

Fonte: o autor.

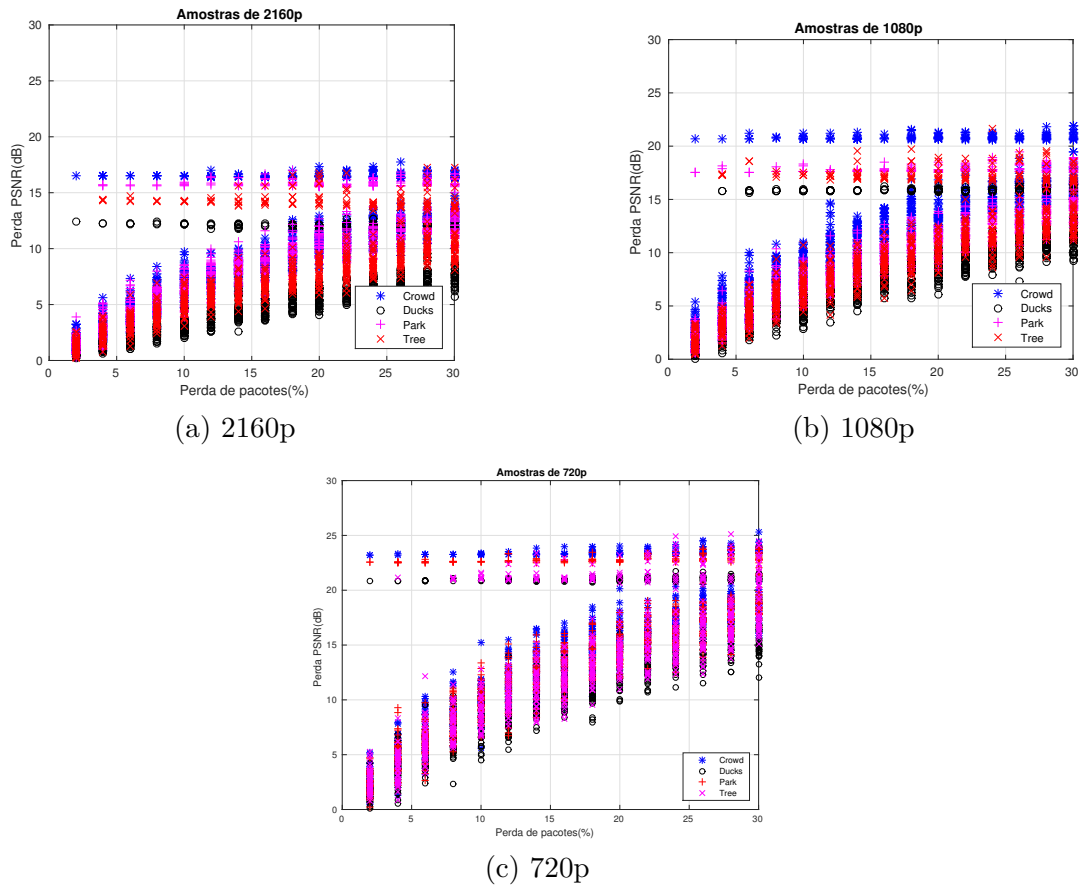
Como esperado, na medida que a perda de pacotes aumenta, também a um acréscimo na perda média de PSNR. É observado também que dentre os valores, existem pontos que não representam muito bem a perda média de PSNR para dada porcentagem de perda de pacotes, por exemplo, na Figura 21a para o vídeo *Crowd Run* no ponto correspondente a perda de pacotes de 2% tem-se o valor de aproximadamente 12 dB , comparando-se este valor com os demais perdas de PSNR para a mesma perda de pacotes pode-se notar

que 12 dB não pode representar a perda de PSNR para 2% de perda de pacotes. Essas amostras na literatura são conhecidas como *outliers* (BARNETT; LEWIS, 1964). Neste trabalho, o procedimento adotado para a filtragem dos *outliers* é baseado no corte a partir da informação do desvio padrão.

4.3.1 Perda média de PSNR simulada

Na Figura 22 são apresentados todos os dados simulados para cada diferente resolução estudada, por exemplo na Figura 22a contém todas as amostras das perdas médias de PSNR para os quatro vídeos. Analisando-se o comportamento dos dados simulados é possível identificar alguns *outliers*, estes valores não representam o comportamento da perda média de PSNR para dada perda de pacotes. Outro comportamento extraído da análise visual dos gráficos, é que existem tipos de vídeos mais sensíveis a perda de pacotes quando se analisa a qualidade do vídeo; na média este comportamento foi observado independente da resolução. Desta forma chegando a conclusão que as características do tipo de vídeo (movimentação) podem influenciar no desempenho do codificador durante a transmissão, provavelmente por conta da perda de *frames P*, os quais carregam informação da compensação de movimento (ZULPRATITA, 2013; ROZAS-RAMALLAL et al., 2013; UHRINA et al., 2014).

Figura 22 – Dados simulados.



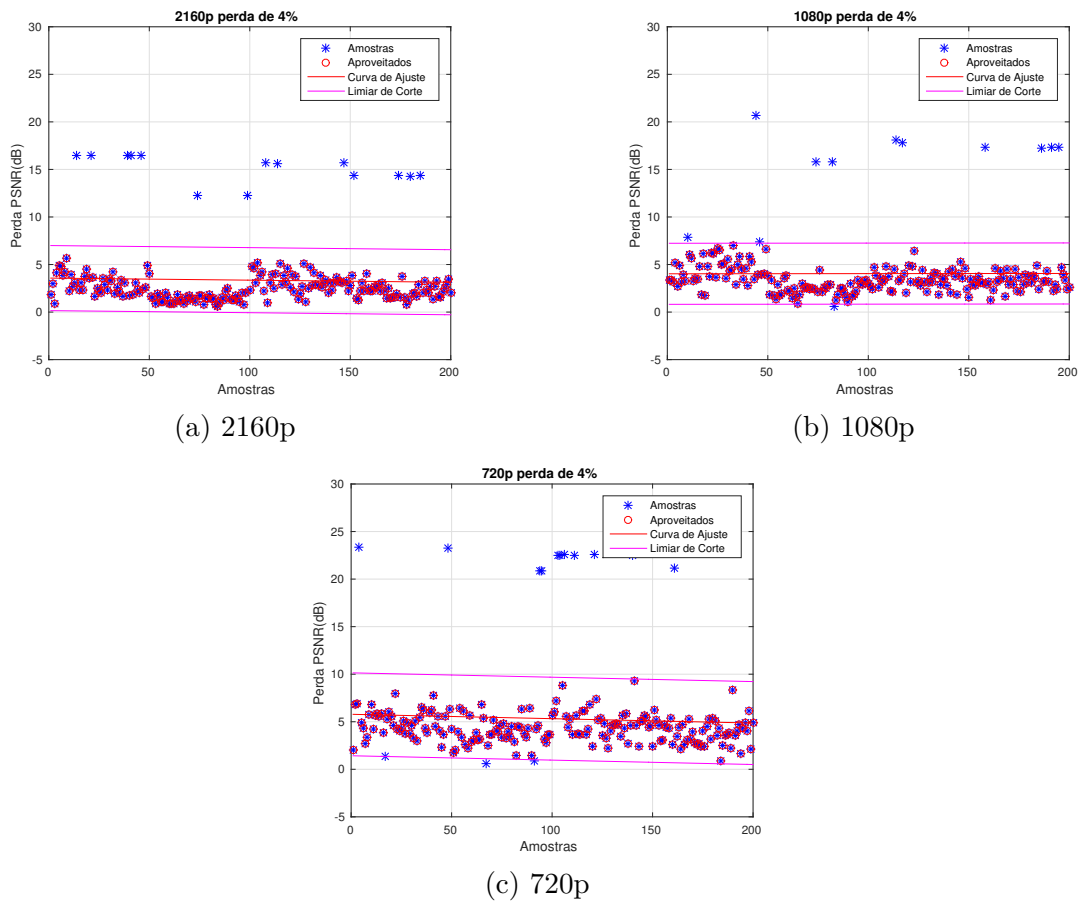
Fonte: o autor.

Como mencionado anteriormente, nos dados simulados existem pontos que não descrevem o comportamento da perda de qualidade do vídeo, logo esses pontos devem ser identificados e removidos. Nesta dissertação, o método escolhido para identificação destes pontos foi o corte pelo desvio padrão (CASTRO, 2014). Por meio dos dados simulados é traçado uma estimativa que melhor representa o comportamento das amostras e se faz um corte somando (limite superior) e subtraindo (limite inferior) o pelo desvio padrão das amostras, conforme a Equação 4.3.

$$L_i = \begin{cases} outlier, & L_{fit} - D_p > L_i > L_{fit} + D_p \\ não é, & caso contrário \end{cases} \quad (4.3)$$

Para tratar os resultados das simulações a análise inicial foi agrupar todos os dados dos quatro vídeos nas três resoluções para uma dada porcentagem de perda de pacotes, conforme apresentada na Figura 23. A melhor estimativa que descreve o comportamento dos dados é a função de primeiro grau, calculando-se o desvio padrão para essas amostras é definido o limite superior e inferior para a filtragem dos dados.

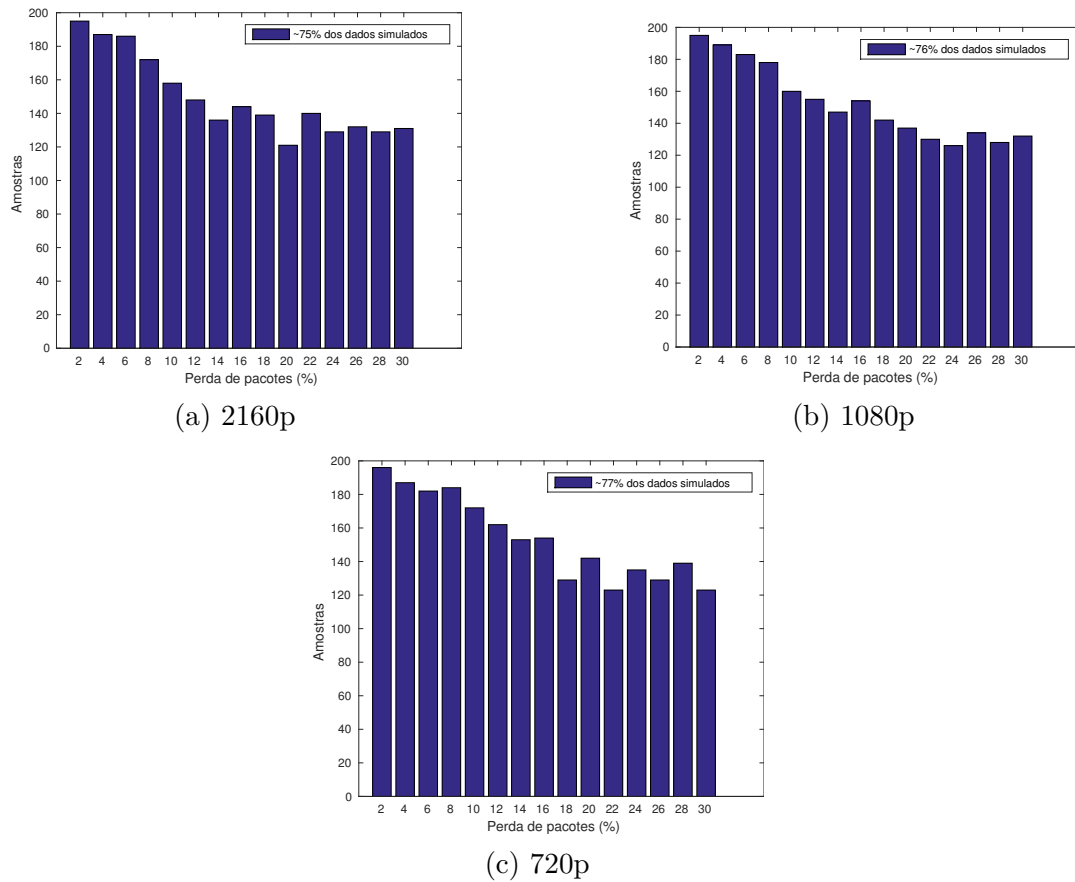
Figura 23 – Tratamento dos dados para perdas de pacote de 4%.



Fonte: o autor.

Realizando-se uma análise na Figura 23 pode-se verificar que após a filtragem dos dados, 13 amostras foram descartadas na Figura 23a; 11 amostras para a Figura 23b e 13 para a Figura 23c. Esta análise também foi aplicada para os demais valores de porcentagem de perda de pacotes. Para se ter uma visão geral do quanto dos dados simulados foram aproveitados, veja a Figura 24.

Figura 24 – Quantidade de amostras aproveitadas.



Fonte: o autor.

Na média se aproveitou de 76% dos dados simulados para o desenvolvimento da proposta de estimativa de perda média de PSNR desta dissertação. A partir dos dados tratados, encontra-se os valores médios da perda de PSNR a medida que ocorrem o aumento da perda de pacotes, obtido por meio da Equação 4.4.

$$L_{simu}(PP) = \frac{\sum_{a=1}^N L_r(a)}{N} \quad (4.4)$$

onde:

L_{simu} é a média das perdas simuladas;

PP porcentagem da perda de pacotes;

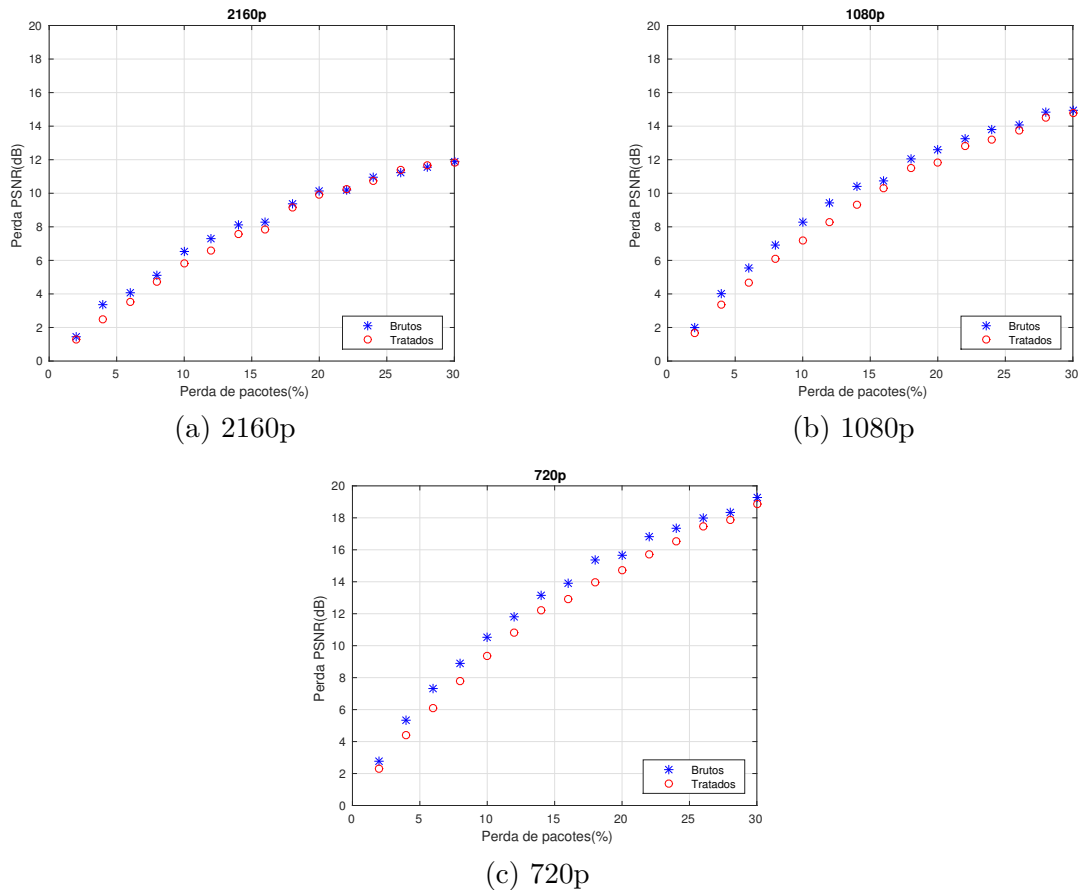
N o número máximo de amostras;

L_r perda média de PSNR por resolução;

Aplicando-se a Equação 4.4 para os dados tratados obtêm-se a média das perdas de PSNR de todos os vídeos para cada resolução. A fim de se ter um comparativo do

quanto os *outliers* interferem nos resultados, é apresentada na Figura 25 a média das perdas de PSNR dos dados considerados “brutos” (não tratados) com os aproveitados.

Figura 25 – Média das perdas de PSNR.



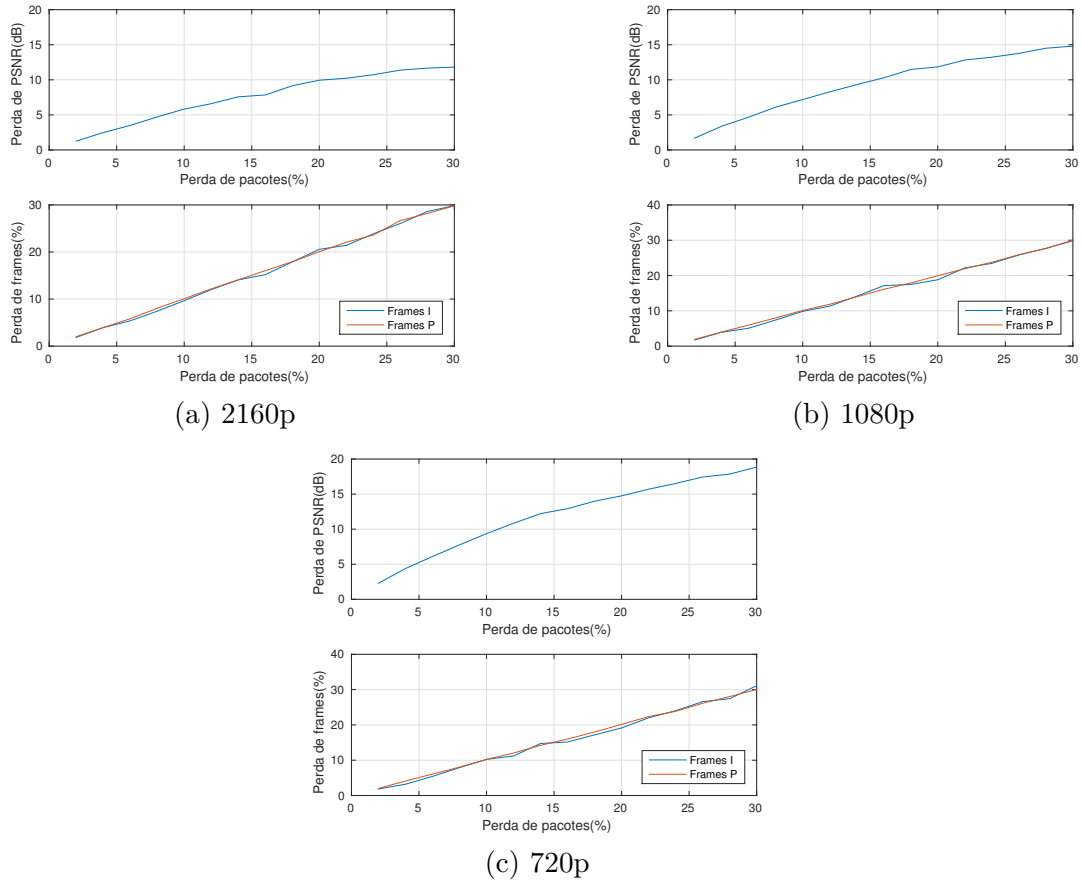
Fonte: o autor.

Como pode-se notar, apesar da remoção dos *outliers*, o comportamento das perdas é mantido, em que os dados tratados possuem um valor médio menor, o qual era esperado. Destes resultados, pode-se deduzir que vídeo em ultra resolução (2160p) degrada menos para pequenas variações de porcentagem de perdas de pacotes, assim como para 1080p, enquanto em 720p são mais sensíveis para a mesma faixa. Contudo, para altas variações de perdas de pacotes a perda de PSNR nas três variações de resoluções tendem a ser mais “comportadas”. Pode-se concluir que o H.264/AVC para os vídeos e resoluções estudados na média não obteve bom desempenho para 720p com pequenas variações de perda de pacotes.

Quando se analisa o comportamento da média da perda de PSNR dos dados simulados em relação às perdas dos *Frames* I e P obtem-se os resultados da Figura 26. Percebe-se uma distribuição proporcional na porcentagem de perdas dos *Frames*. Logo, a proposta apresentada por esta dissertação estuda a estimativa para pequenas variações

entre as perdas dos *Frames*.

Figura 26 – Média da Perda de PSNR em relação a perdas de Frames I e P.



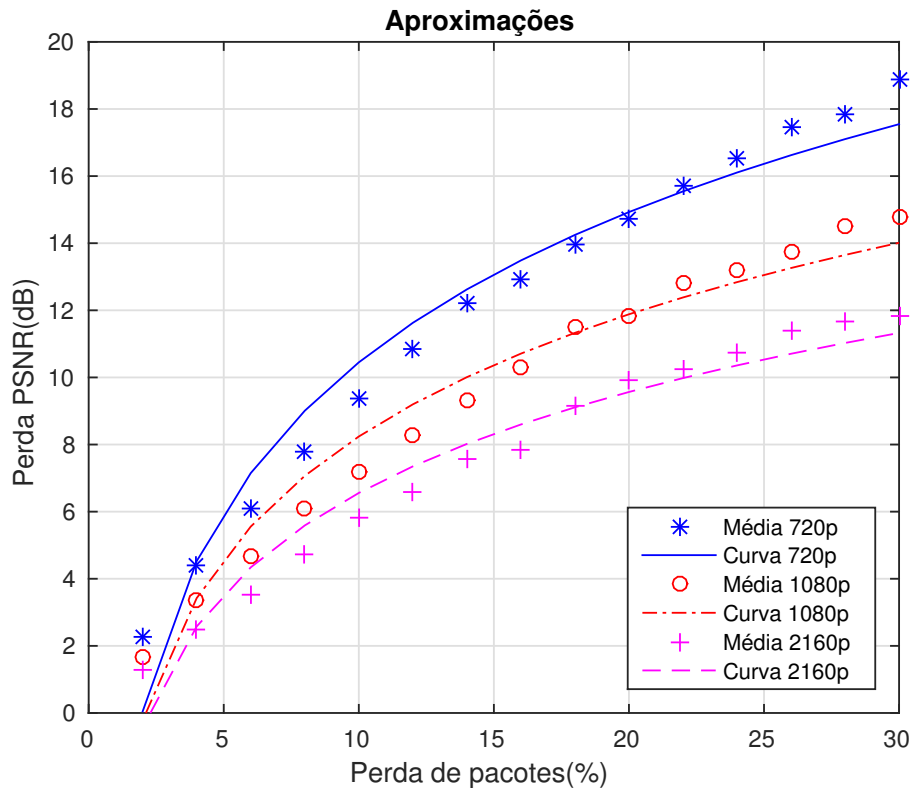
Fonte: o autor.

Com os dados tratados, encontrou-se a melhor expressão matemática para representar o comportamento médio das perdas de PSNR. Como pode-se observar, a perda deverá saturar em um determinado momento, logo utilizou-se a expressão logarítmica como função base, conforme na Equação 4.5 com $L_{PSNR_{fit}}$ sendo a média da perda de PSNR dada em *dB*, PP a porcentagem da perda de pacotes, e a e b são os coeficientes. Para determinar os coeficientes utilizou-se a técnica de minimização por mínimos quadrados, esta técnica de otimização tem sido largamente implementada na literatura, conforme apresentado por Matos et al. (2016) e CASTRO (2014). Para solução utilizou-se a Equação 4.6.

$$L_{PSNR_{fit}} = a + b \log_{10}(PP) \quad (4.5)$$

$$A = \begin{bmatrix} 1 & \log_{10}(p) \\ 1 & \log_{10}(p) \end{bmatrix}; \text{coef} = \begin{bmatrix} a \\ b \end{bmatrix}; B = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}; \text{coef} = (A'A)^{-1}A'B \quad (4.6)$$

Figura 27 – Média dados tratados vs Estimativas.



Como resultados encontram-se as Equações 4.7, 4.8 e 4.9

$$L_{720p} = 25.32 + 14.88 \log_{10}(PP\%) \quad (4.7)$$

$$L_{1080p} = 20.34 + 12.1 \log_{10}(PP\%) \quad (4.8)$$

$$L_{2160p} = 16.55 + 10 \log_{10}(PP\%) \quad (4.9)$$

Comparando-se as médias das três resoluções com suas respectivas estimativas na Figura 27, nota-se que as resoluções 2160p e 1080p alcançaram valores estimados que não representam a realidade (perda de PSNR negativas) para pequenas porcentagens de perda de pacotes. Concluindo que para perda de pacotes abaixo de 2,25% não se tem degradações significativas para os vídeos em 2160p e 1080p.

A partir das estimativas logarítmicas, procura-se determinar uma relação entre às três expressões (Equações 4.7, 4.8 e 4.9). Percebe-se que elas possuem comportamentos semelhantes, variando somente os coeficientes a e b . Desta forma, determina-se que estes coeficientes são dependentes da resolução do vídeo.

Neste caso, a expressão que melhor descreve o comportamento de a e b é uma exponencial decrescente, visto que ambas as variáveis decrescem a medida que há um aumento da resolução, na Tabela 11 esta observação fica evidente. A estratégia utilizada para determinar a função exponencial, foi a mesma utilizada na Equação 4.6, usando como função base a exponencial, conforme:

$$C = A + Be^{(-R)} \quad (4.10)$$

onde:

C é o comportamento a ser analisado;

A e B são constantes;

R é tipo de resolução.

Para esta dissertação adotou-se que o tipo de resolução será representado por R e poderá assumir os valores: 1 (um), 1,5 (um e meio) e 3 (três) para 720p, 1080p e 2160p, respectivamente.

Tabela 11 – Comportamento de a e b .

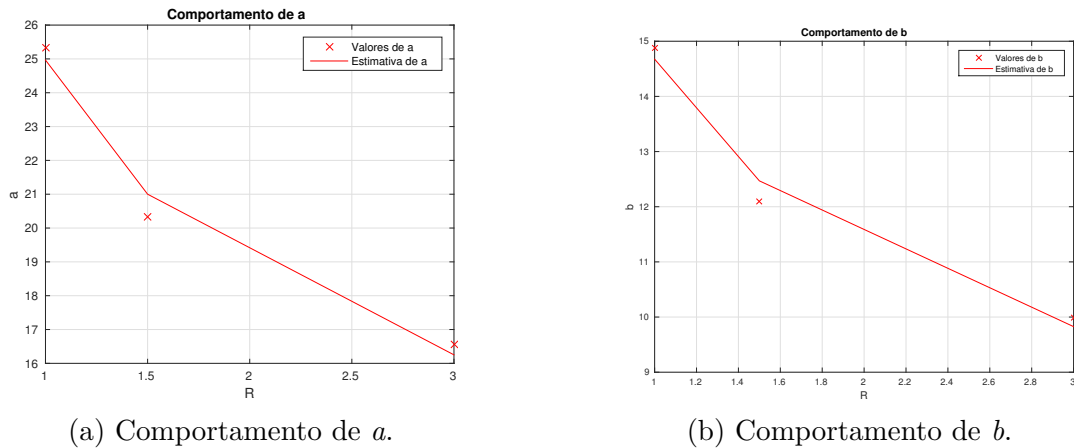
Variável	Resolução		
	720p	1080p	2160p
Comportamento de a	25,32	20,34	16,55
Comportamento de b	14,88	12,1	10

Fonte: o Autor

Aplicando-se a estratégia para obtenção dos novos coeficientes para A e B encontra-se as Equações 4.11 e 4.12 que estimam o comportamento das variáveis a e b , respectivamente. Na Figura 28 apresenta-se os valores dos coeficientes com suas estimativas.

$$C_a = 14,89 + 27,4e^{(-R)} \quad (4.11)$$

$$C_b = 9,07 + 15,24e^{(-R)} \quad (4.12)$$

Figura 28 – Estimativas do comportamento de a e b .

Fonte: o autor.

Neste momento tem-se uma expressão geral para estimar a média da perda de PSNR, visto que a partir dos dados simulados e tratados foi possível descrever a média do comportamento da perda de PSNR a medida que ocorram perdas de pacotes para cada resolução estudada. Esta expressão é apresentada na Equação 4.13.

$$L_{PSNR} = C_a + C_b \log_{10}(PP) \quad (4.13)$$

Expandindo a Equação 4.13 para às três resoluções estudadas nesta dissertação tem-se:

$$L_{PSNR}(R, PP) = (14,89 + 27,4e^{(-R)}) + (9,07 + 15,24e^{(-R)}) \log_{10}(PP) \quad (4.14)$$

onde:

$L_{PSNR}(R, PP)$ é a estimativa média da perda de PSNR dada em dB ;

R sendo a resolução;

PP sendo a porcentagem de perda de pacotes dada em %.

Para Equação 4.14 obteve-se os valores de RMSE e desvio padrão para a média dos dados simulados, conforme a Tabela 12.

Tabela 12 – RMSE e desvio padrão do modelo para a média dos dados simulados.

Métrica	Resolução		
	720p	1080p	2160p
RMSE em <i>dB</i>	1,3	0,89	0,73
Desvio Padrão em <i>dB</i>	0,97	0,85	0,74

Fonte: o autor.

Desta forma pode-se concluir que o modelo proposto aproximou-se da média dos dados simulados.

4.4 Considerações finais

Neste capítulo apresentou-se a metodologia adotada para o desenvolvimento da proposta de modelagem para predição de perda de qualidade de vídeo. No próximo capítulo será avaliada a modelagem proposta em relação aos dados reais obtidos a partir de medições.

5 Análise dos resultados

5.1 Considerações iniciais

No capítulo anterior, um modelo foi proposto para estimar a média da perda PSNR para três resoluções a medida que ocorram perdas de pacotes sobre uma determinada rede de comunicação sem fio. Contudo, em ambiente simulado as variáveis tendem a ser mais controladas. Para avaliar a proposta, dados reais são necessários. Logo, neste capítulo serão apresentados os dados reais coletados em contrastes com os valores estimados pelo modelo proposto.

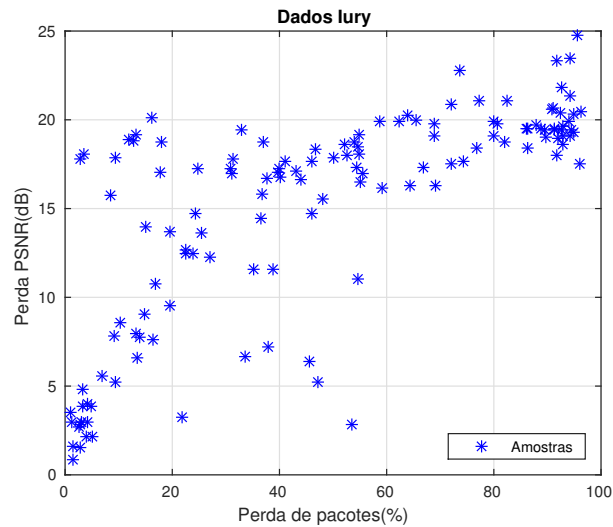
5.2 Coleta dos dados reais

O processo para coleta dos dados reais foi o mesmo seguido em Seção 3.4 excluindo-se a etapa de codificação e empacotamento do vídeo, visto que esta etapa necessita ser realizada uma única vez. Todos os vídeos são codificados utilizando o H.264/AVC e transmitidos sobre uma rede de comunicação sem fio no padrão *802.11ac*.

5.2.1 Dados utilizados

Uma abordagem inicial que se teve foi aproveitar os dados coletados por Batalha (2016), o qual realizou transmissões reais com o vídeo “Crowd Run” na resolução de 2160p codificado com o H.264/AVC no padrão *802.11ac*, seguindo as mesmas etapas para as transmissões desta dissertação. Diferenciando-se somente o *bit rate* utilizado para codificação do vídeo, que foi de aproximadamente *82 Mbps*. Analisando-se somente o valor da média de PSNR, isto quer dizer que o vídeo terá um valor maior em relação ao mesmo vídeo utilizado nesta dissertação, estudado por Auwera, David e Reisslein (2008). Para o vídeo transmitido por Batalha (2016) obteve-se média de PSNR de *33,74 dB*, se comparado o mesmo vídeo e resolução utilizados neste trabalho (ver Tabela 6), encontra-se uma variação de *2,87 dB*. Esta variação será importante para realização do ajuste na Equação 4.14. A Figura 29 apresenta as amostras obtidas das transmissões em ambiente real por Batalha (2016).

Figura 29 – Dados gerados por Batalha (2016).

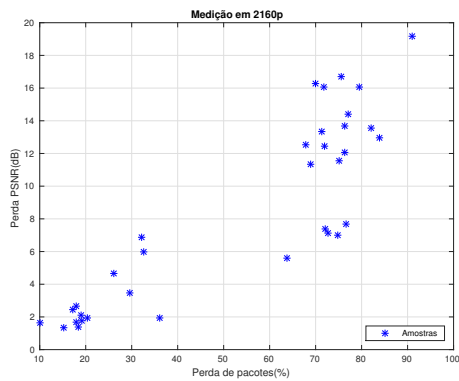


Fonte: Batalha (2016).

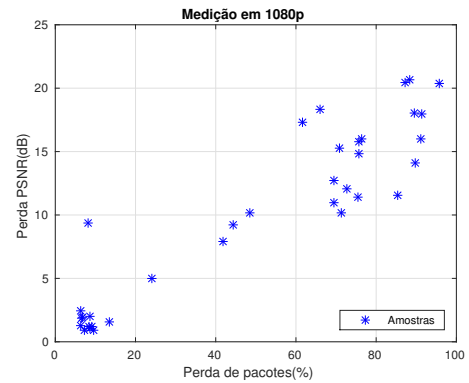
5.2.2 Obtenção de novos dados

Novas transmissões em ambiente real foram realizadas para os quatro vídeos em três resoluções seguindo as etapas descritas na Seção 3.4 no esquema de transmissão da Figura 12. Neste trabalho, define-se como medição a coleta dos dados em um ambiente real. As medições ocorreram no Anexo B do Laboratório de Engenharia Elétrica e Computação da Universidade Federal do Pará e com objetivo de reproduzir a perda de pacotes durante as transmissões dos vídeos, posiciona-se o receptor (*Notebook 2*) em determinados pontos distantes do *Roteador* até que seja notado uma perda significativa de pacotes. Para uma configuração de vídeo (tipo de vídeo e resolução) realiza-se três transmissões, ou seja, para o vídeo *Crowd Run* em 720p se terá 3 amostras. Logo para uma determinada medição se tem no total de 36 amostras. Realizou-se ao todo três medições em ambiente real. Na Figura 30 são apresentados os valores da média de perda de PSNR obtidos das 3 medições, organizou-se os dados por resolução.

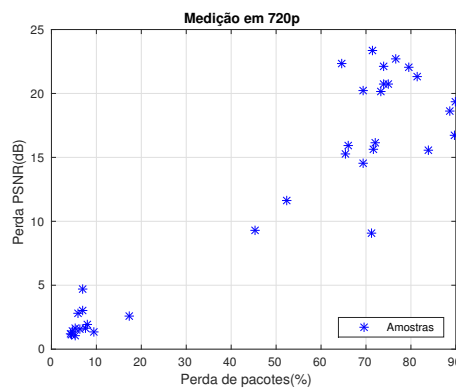
Figura 30 – Medições reais.



(a) 2160p



(b) 1080p



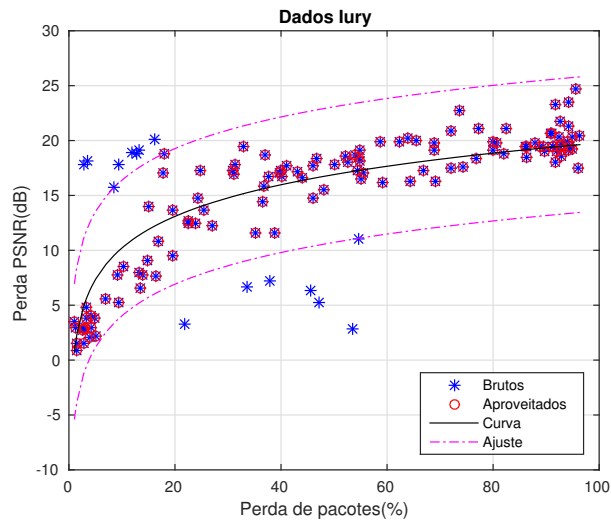
(c) 720p

Fonte: o autor.

5.3 Tratamento dos dados reais

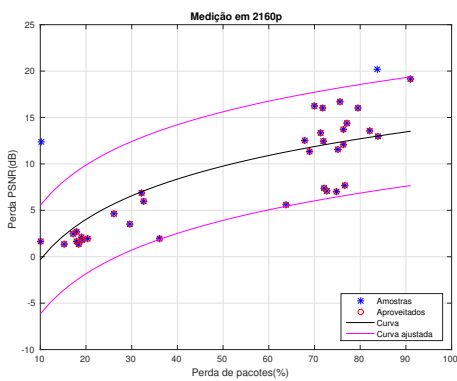
Como observado durante as simulações, foi possível identificar a presença dos *outliers* também nos dados coletados em ambiente real. Para remover as amostras que não representam o comportamento da média de perda de PSNR, utilizou-se o corte pela informação do desvio padrão, conforme mencionado na Seção 4.3. As amostras aproveitadas das medições podem ser visualizadas na Figura 31 e Figura 32.

Figura 31 – Dados de Batalha (2016) tratados.

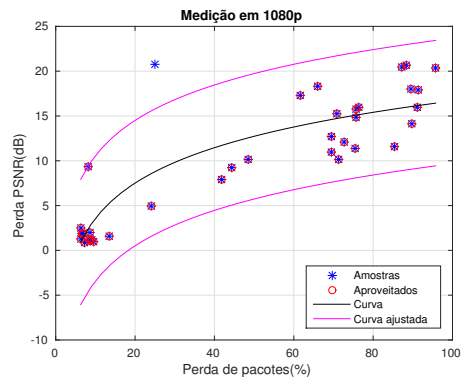


Fonte: o autor.

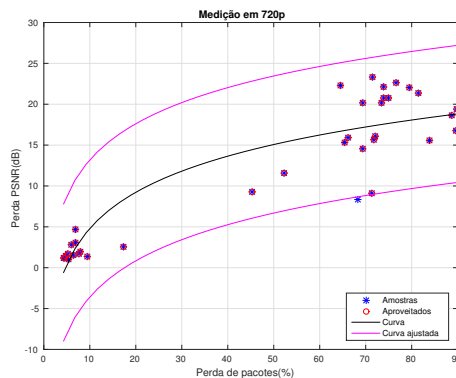
Figura 32 – Medições reais tratadas.



(a) 2160p



(b) 1080p



(c) 720p

Fonte: o autor.

No geral, nota-se que não houve grande perda na quantidade de amostras das medições.

5.4 Análise comparativa entre os dados reais e as estimativas

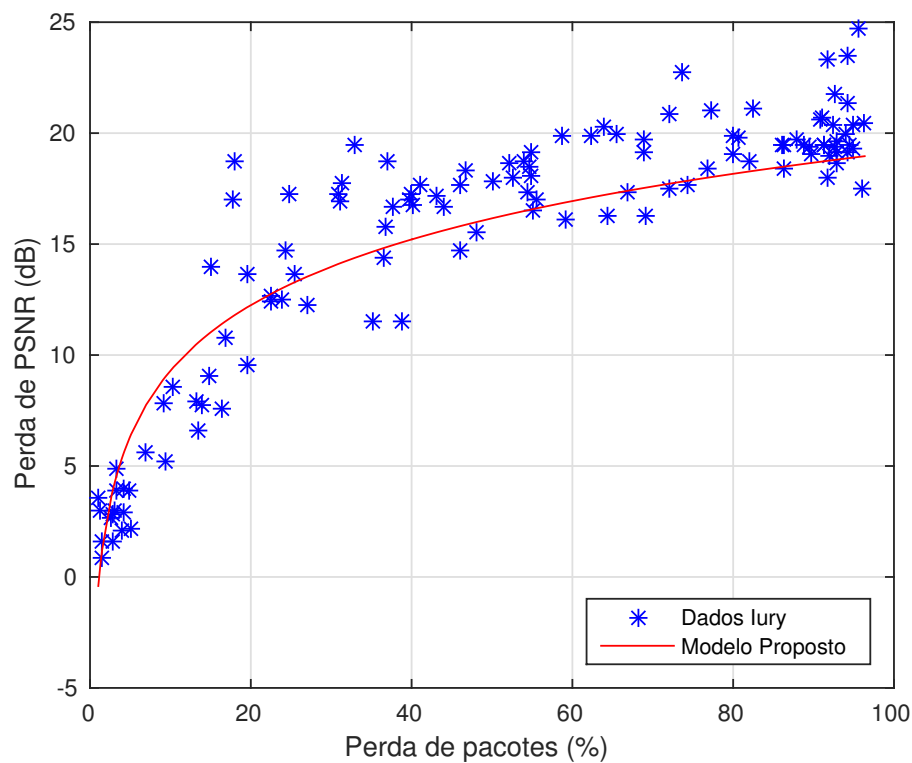
Com o objetivo de avaliar a precisão do modelo proposto por esta dissertação, pretende-se realizar uma análise comparativa entre os valores estimados (modelo proposto) com os dados reais sem os *outliers*. Esta análise será aplicada para às três resoluções em estudo. Sendo que para a resolução em 2160p se terá um conjunto de amostras a mais, conforme mencionado o aproveitamento dos dados fornecidos por Batalha (2016).

A partir da observação visual da média de perda de PSNR dos vídeos transmitidos, percebe-se uma similaridade entre as perdas de qualidades do vídeo. A partir desta observação e a análise da influência da variação do *bit rate* realizada por Auwera, David e Reisslein (2008). Propõem-se adicionar um ganho de $2,87 \text{ dB}$ ao modelo proposto, resultando na Equação 5.1.

$$L_{PSNR}(PP) = [14,89 + 27,4e^{(-R)}] + [(9,07 + 15,24e^{(-R)}) \log_{10}(PP)] + 2,87 \quad (5.1)$$

Os valores estimados pela Equação 5.1 são apresentados na Figura 33 em contraste com seus valores medidos.

Figura 33 – Modelo proposto em relação aos dados de Batalha (2016).



Fonte: o autor.

De acordo com a Figura 33 pode-se notar que o modelo proposto com o ajuste, por conta da diferença de *bit rate*, obteve valores satisfatórios, aproximando-se dos dados

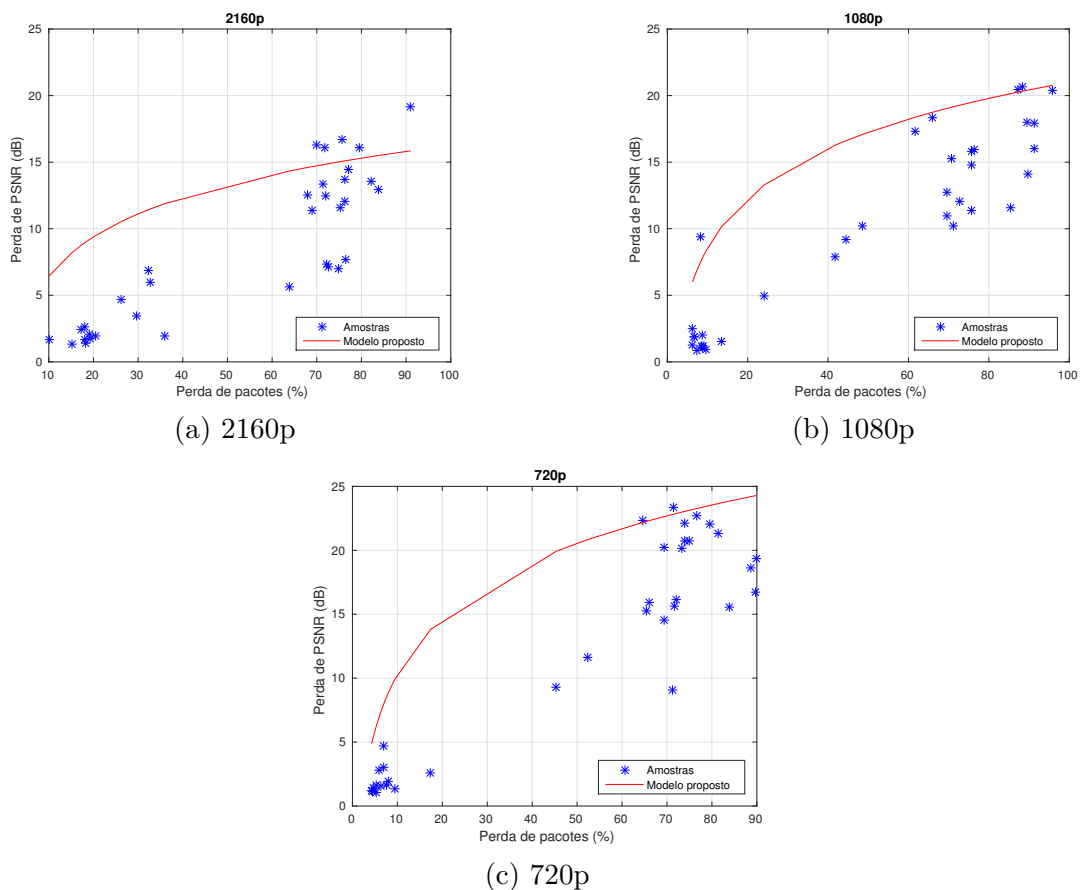
reais. Para se comprovar tal análise, calcula-se os valores de RMSE e desvio padrão para o modelo proposto, os quais são apresentados na Tabela 13.

Tabela 13 – RMSE e desvio padrão do modelo proposto em relação aos dados de Batalha (2016).

RMSE	2,32 dB
Desvio Padrão	2,20 dB

A mesma análise é realizada para os vídeos estudados nesta dissertação, resultando nos gráficos da Figura 34.

Figura 34 – Modelo proposto em relação aos dados reais.



Fonte: o autor.

De acordo com a Figura 34, o modelo proposto não obteve valores satisfatórios quando analisada para baixas porcentagens de perdas de pacotes nas três resoluções. Estes resultados são esperados, visto que o modelo proposto foi desenvolvido a partir da média das perdas de PSNR simuladas (visto na Seção 4.3). Para uma abordagem inicial que está levando em consideração somente a variação das perdas de pacotes para estimar a perda de qualidade do vídeo, pode-se concluir que no geral obteve-se bons resultados

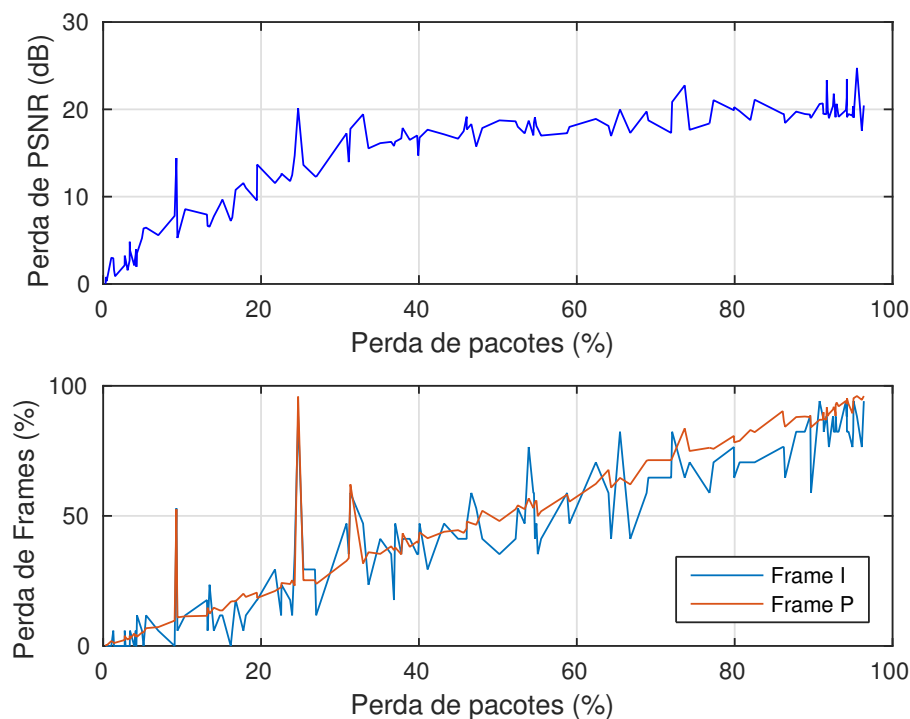
para um estudo com esse enfoque. Na Tabela 14 apresenta-se os valores obtidos de RMSE e desvio padrão.

Tabela 14 – RMSE e desvio padrão do modelo proposto em relação aos dados reais

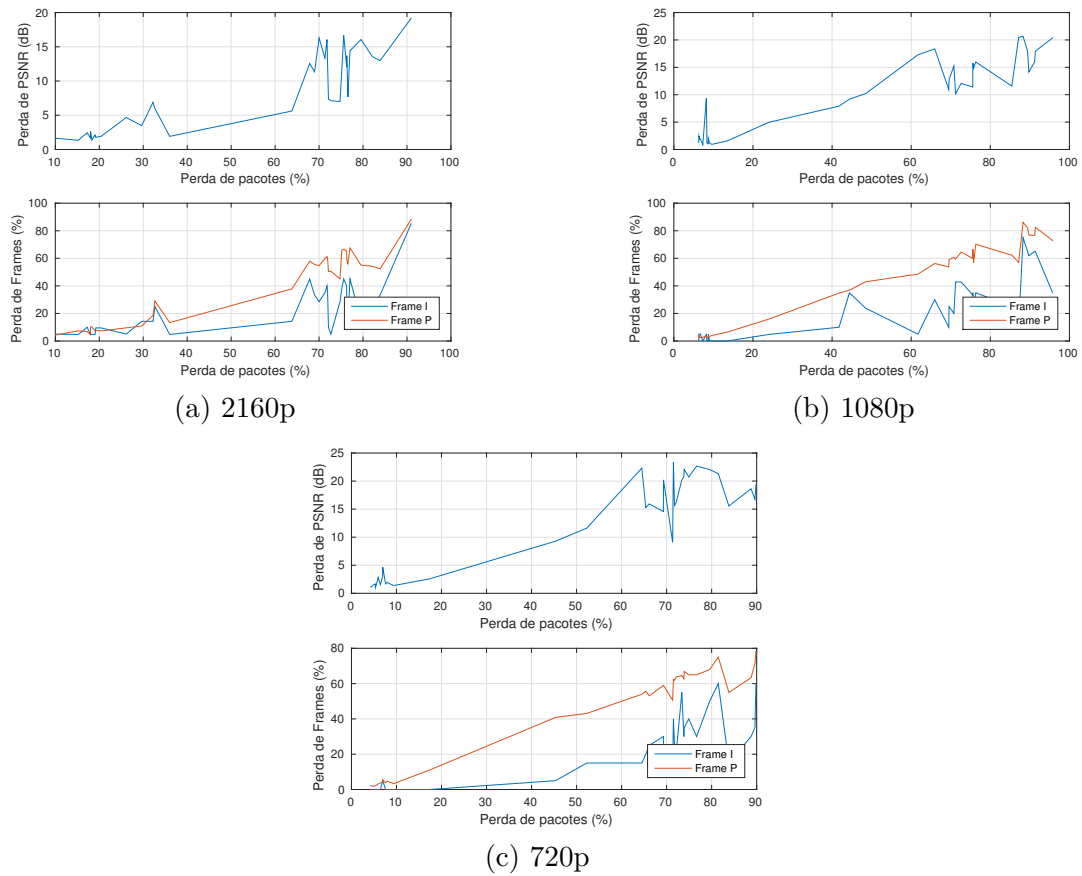
Métrica	Resoluções		
	720p	1080p	2160p
RMSE	6,16 dB	5,76 dB	5,56 dB
Desvio Padrão	3,29 dB	2,96 dB	3,52 dB

De acordo com a Tabela 14, o modelo proposto não apresentou bons resultados. Entretanto, quando se compara com os resultados da Tabela 13 percebe-se que não somente o *bit rate* se diferencia, mas também o comportamento das perdas de *Frames I* e *P*. O modelo proposto é baseado na média das perdas de PSNR dos quatro vídeos em três resoluções em ambiente simulado, como visto na Seção 4.3. Desta forma, os resultados tem distribuição proporcional das perdas dos *Frames*. Esta distribuição proporcional é notado nos dados de (BATALHA, 2016), como pode ser visto na Figura 35. Enquanto na Figura 36 nota-se que essa distribuição não é proporcional.

Figura 35 – Comportamento dos *Frames I* e *P* para os dados de Batalha (2016).



Fonte: o autor.

Figura 36 – Comportamento dos *Frames* I e P para as medições.

Fonte: o autor.

5.5 Considerações finais

A modelagem proposta apresentou bons resultados quando comparado com os dados medidos, representando a média de perda de PSNR em função da porcentagem de perdas de pacotes imposta na rede de comunicação no padrão 802.11ac. Contudo, o comportamento do vídeo sobre a rede de comunicação também depende de outros fatores de codificação de vídeo que não são considerados nesta dissertação.

6 Conclusões

Este trabalho apresentou uma abordagem para predição de perda de qualidade de vídeo, em uma rede de comunicação propensa a perda de pacotes durante a transmissão de vídeo. Quatro vídeos em três resoluções diferentes são codificados em iguais condições de codificação de fonte, no padrão H.264/AVC e posterior transmissão sobre a rede de comunicação (IEEE 802.11ac). Transmissões ideais, sem perda de pacotes, foram realizadas para se obter os arquivos necessários para reconstrução do vídeo. Simulando perdas de pacotes no receptor, reconstrução dos vídeos, obtenção das métricas de qualidade de vídeo, tratamentos dos dados simulados; é proposto o modelo matemático que relaciona a perda de qualidade de vídeo com a perda de pacotes para às três resoluções em estudo.

A modelagem proposta pelo presente trabalho tem como base, o comportamento médio da perda de PSNR dos vídeos estudados, agrupando o comportamento da perda de qualidade dos quatro vídeos por resolução e encontrando a média entre eles e; resultando em três médias. O método dos mínimos quadrados é aplicado às três médias obtendo-se suas respectivas estimativas não lineares por resolução. Por meio das estimativas é proposto a expressão matemática dependente da resolução do vídeo que descreve a estimativa da perda de qualidade de vídeo conforme se tem perdas de pacotes. No geral, a função logarítmica descreve o comportamento médio da perda de PSNR, visto que a PSNR tende a saturar em um determinado momento. Os coeficientes da função logarítmica são descritos por meio de uma função exponencial dependente do tipo de resolução.

O modelo proposto demonstrou ser bem representativo. Alcançando bons resultados quando comparado com a média dos dados simulados, chegando a ter valor médio de RMSE de $0,97 \text{ dB}$ e desvio padrão de $0,85 \text{ dB}$. Para verificar a eficácia da modelagem proposta, comparativos com dados reais foram utilizados.

Três medições foram realizadas para obtenção dos dados em ambiente real. Também aproveitou-se as medições de Batalha (2016), pois o trabalho utilizou um dos vídeos (*Crown Run* em 2160p) em estudo. Confrontando os dados medidos com suas estimativas do modelo proposto, obteve-se bons resultados quando analisado os dados de Batalha (2016), com média de RMSE de $2,32 \text{ dB}$ e desvio padrão de $2,2 \text{ dB}$. Contudo, o mesmo resultado não foi alcançado das medições desta dissertação, com média de RMSE e desvio padrão de $5,83 \text{ dB}$ e $3,25 \text{ dB}$, respectivamente.

As simulações apresentaram resultados satisfatórios, permitindo a análise dos dados simulados para propor a modelagem. Dentre os resultados observados nos dados si-

mulados, tem-se que o comportamento da perda de qualidade de vídeo é dependente da perda de pacotes, assim como da resolução do vídeo. Outra característica observada diz respeito a composição do vídeo, movimentação, o qual pode fazer com que determinado vídeo seja mais sensível a perda de pacotes; esta característica tem relação com a composição do GOP (*Frames I e P*). Contudo, a influência do GOP na perda de qualidade de vídeo foge ao estudo desta dissertação.

6.1 Trabalhos futuros

Como trabalhos futuros desta dissertação destacando-se:

- Modelar o impacto que as perdas de *Frames I e P* tem sobre a média de perda de PSNR;
- Realizar um ajuste ao modelo proposto a partir de dados medidos em condições reais de transmissão de vídeo, por exemplo usando ajuste com mínimos quadrados lineares.

6.2 Produções científicas durante os estudos

Os trabalhos publicados durante o mestrado tem-se:

- MATOS, E. M. C.; Costa, T. A. ; BATALHA, I. S. ; SILVA, D. K. N. ; CASTRO, B. S. L. ; CAVALCANTE, G. P. S. ; PELAES, E. G. . Modelo de extrapolação para perda de propagação para frequência de 10 GHz em ambiente indoor. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS, 2016, Santarém, PA. Anais do XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT 2016), 2016. p. 893-896.
- Costa, T. A. ; MATOS, E. M. C. ; BATALHA, I. S. ; CASTRO, B. S. L. ; CAVALCANTE, G. P. S. ; PELAES, E. G. . Abordagem Comparativa de Técnicas de Filtragem Espacial em Geometrias Linear e Circular de Arranjos de Antenas. In: Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 2016, Santarém, PA. Anais do XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT 2016), 2016. p. 134-138.

Referências

- ARAÚJO, J. P. L. de. *Estratégia de planejamento e otimização de sistemas sem fio, considerando redes interferentes: abordagem baseada em crosslayer*. Tese (Doutorado) — Universidade Federal Do Pará, Programa De Pós-graduação Em Engenharia Elétrica, 2011. Citado 2 vezes nas páginas 7 e 8.
- AUWERA, G. V. der; DAVID, P. T.; REISSLEIN, M. Traffic characteristics of H.264/AVC variable bit rate video. *IEEE Communications Magazine*, 2008. v. 46, n. 11, p. 164–174, nov. 2008. ISSN 0163-6804. Citado 2 vezes nas páginas 49 e 53.
- BARNETT, V.; LEWIS, T. *Outliers in statistical data*. [S.l.]: Chichester: John Wiley, 1995. 584p, 1964. Citado na página 39.
- BATALHA, I. da S. *Estudo da tecnologia IEEE 802.11ac para o desenvolvimento de modelos empírico e cross-layer*. Dissertação (Mestrado) — Universidade Federal Do Pará, Programa De Pós-graduação Em Engenharia Elétrica, 2016. Citado 11 vezes nas páginas VII, VIII, X, 11, 49, 50, 52, 53, 54, 55 e 57.
- BONDZULIC, B. P. et al. Performance of peak signal-to-noise ratio quality assessment in video streaming with packet losses. *Electronics Letters*, 2016. v. 52, n. 6, p. 454–456, 2016. ISSN 0013-5194. Citado na página 2.
- BOVIK, A. C. *Handbook of image and video processing*. [S.l.]: Academic press, 2010. Citado na página 16.
- BROWN, M. *Advanced Digital Photography*. [S.l.]: Media Publishing, 2004. Citado na página 11.
- BT.2020-2. *Parameter values for ultra-high definition television systems for production and international programme exchange*. [S.l.], 2015. Citado na página 12.
- BT.601-7. *Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios*. [S.l.], 2011. Citado na página 12.
- BT.709-6. *Parameter values for the HDTV standards for production and international programme exchange*. [S.l.], 2015. Citado na página 12.
- CASTRO, B. S. L. *Modelagem cross-layer da qualidade de experiência para transmissões de vídeo em sistemas sem fio ofdm*. Tese (Doutorado) — Universidade Federal Do Pará, Programa De Pós-graduação Em Engenharia Elétrica, 2014. Citado 2 vezes nas páginas 40 e 44.
- DUBOLIA, R. et al. Digital image watermarking by using discrete wavelet transform and discrete cosine transform and comparison based on PSNR. In: *Proc. Int Communication Systems and Network Technologies (CSNT) Conf.* [S.l.: s.n.], 2011. p. 593–596. Citado na página 17.

- EVALVID. *Versão 2.7*. 2016. <http://www2.tkn.tu-berlin.de/research/evalvid/EvalVid/docevalvid.html>. Citado na página 29.
- FFMPEG. *Version 2.8.10*. 2016. <https://ffmpeg.org/>. Citado na página 22.
- FOUNDATION, X. *Xiph.org Video Test Media*. 2016. Disponível em: <<https://media-xiph.org/video/derf/>>. Citado 2 vezes nas páginas 19 e 20.
- FRNDA, J.; VOZNAK, M.; SEVCIK, L. Impact of packet loss and delay variation on the quality of real-time video streaming. *Telecommunication Systems*, 2016. Springer, v. 62, n. 2, p. 265–275, 2016. Citado 2 vezes nas páginas 2 e 26.
- H.264. *SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS Infrastructure of audiovisual services – Coding of moving video*. [S.l.], 2016. Citado na página 15.
- IEEE 802.11AC'. *IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz*. dez. 2013. 1–425 p. Citado na página 10.
- JAIN, R. *Art of Computer Systems Performance Analysis Techniques For Experimental Design Measurements Simulation And Modeling*. [S.l.]: Wiley Computer, 1991. Citado na página 30.
- JUNIOR, J. J. H. F. *Predição de qualidade de experiência para internet do futuro em arquiteturas heterogêneas de redes sem fio móveis*. Tese (Doutorado) — Universidade Federal Do Pará, Programa De Pós-graduação Em Engenharia Elétrica, 2013. Citado 2 vezes nas páginas 5 e 7.
- KLAUE, J.; RATHKE, B.; WOLISZ, A. Evalvid—a framework for video transmission and quality evaluation. In: *Computer performance evaluation. Modelling techniques and tools*. [S.l.]: Springer, 2003. p. 255–272. Citado 3 vezes nas páginas 23, 24 e 26.
- MATLAB. *Versão 8.5.0.197613 (R2015a)*. 2015. The MathWorks Inc. Citado na página 30.
- MATOS, E. M. C. de et al. Modelo de extrapolação para perda de propagação para frequência de 10 ghz em ambiente indoor. In: *Simpósio Brasileiro De Telecomunicações - Sbrt2016*. [S.l.: s.n.], 2016. Citado na página 44.
- MP4BOX. *GPAC version 0.5.2*. 2016. <https://gpac.wp.imt.fr/mp4box/>. Citado na página 22.
- ONG, E. H. et al. IEEE 802.11ac: Enhancements for very high throughput wlans. In: *Proc. Indoor and Mobile Radio Communications 2011 IEEE 22nd Int. Symp. Personal*. [S.l.: s.n.], 2011. p. 849–853. ISSN 2166-9570. Citado na página 10.
- PRASAD, R. *OFDM for wireless communications systems*. [S.l.]: Artech House, 2004. Citado na página 9.
- RICHARDSON, I. E. *H.264 and MPEG-4 video compression: video coding for next-generation multimedia*. [S.l.]: John Wiley & Sons, 2004. Citado 3 vezes nas páginas 13, 15 e 19.

RICHARDSON, I. E. *The H.264 advanced video compression standard*. [S.l.]: John Wiley & Sons, 2011. Citado 2 vezes nas páginas 11 e 14.

ROZAS-RAMALLAL, I. et al. Improving performance of H.264/AVC transmissions over vehicular networks. In: *Proc. IFIP/IEEE Int. Symp. Integrated Network Management (IM 2013)*. [S.l.: s.n.], 2013. p. 1324–1327. ISSN 1573-0077. Citado na página 39.

SEIDEL, I. *Análise do impacto de pel decimation na codificação de vídeos de alta resolução*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Ciência da Computação, 2014. Citado 4 vezes nas páginas 11, 12, 13 e 15.

TANENBAUM, A. S. *Computer networks, 4-th edition*. [S.l.: s.n.], 2003. Citado 2 vezes nas páginas 6 e 7.

TCPDUMP. *Versio 4.7.4*. 2016. <http://www.tcpdump.org/>. Citado na página 23.

UHRINA, M. et al. The impact of gop pattern and packet loss on the video quality of h. 264/avc compression standard. In: BALICKI, J. (Ed.). *Advances in Neural Networks, Fuzzy Systems and Artificial Intelligence*. [S.l.: s.n.], 2014. Citado 2 vezes nas páginas 2 e 39.

UHRINA, M.; HLUBIK, J.; VACULIK, M. Correlation between objective and subjective methods used for video quality evaluation. In: *Proc. ELEKTRO*. [S.l.: s.n.], 2012. p. 103–108. Citado na página 16.

UHRINA, M.; HLUBIK, J.; VACULIK, M. Impact of compression on the video quality. *Advances in Electrical and Electronic Engineering*, 2012. Faculty of Electrical Engineering and Computer Science VSB-Technical University of Ostrava, v. 10, n. 4, p. 251, 2012. Citado na página 16.

WOOTTON, C. *A practical guide to video and audio compression: from sprockets and rasters to macroblocks*. [S.l.]: Taylor & Francis, 2005. Citado na página 12.

YIM, C.; BOVIK, A. C. Evaluation of temporal variation of video quality in packet loss networks. *Signal Processing: Image Communication*, 2011. Elsevier, v. 26, n. 1, p. 24–38, 2011. Citado na página 1.

YINTERNIX. *O que é 4K e qual a diferença pra Full HD?* 2016. <http://www.ynternix.com/o-que-e-4k-e-qual-diferenca-para-full-hd/>. Último acesso: 16-02-2017. Citado na página 14.

ZULPRATITA, U. S. GOP length effect analysis on H.264/AVC video streaming transmission quality over lte network. In: *3rd International Conference on Computer Science and Information Technology (ICCSIT'2013) January 4-5, 2013 Bali (Indonesia)*. [S.l.]: Planetary Scientific Research Centre, 2013. Citado 2 vezes nas páginas 2 e 39.