

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**UM SISTEMA DE GERENCIAMENTO DA QUALIDADE DE
EXPERIÊNCIA ORIENTADA À TRANSMISSÃO DE VÍDEOS PARA
DISPOSITIVOS MÓVEIS EM REDES SEM FIO**

ORLEWILSON BENTES MAIA

BELO HORIZONTE – MG
FEVEREIRO DE 2015

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ORLEWILSON BENTES MAIA

**UM SISTEMA DE GERENCIAMENTO DA QUALIDADE DE
EXPERIÊNCIA ORIENTADA À TRANSMISSÃO DE VÍDEOS PARA
DISPOSITIVOS MÓVEIS EM REDES SEM FIO**

Tese apresentada à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Dr. Luciano de Errico

Co-Orientador: Prof. Dr. Hani Camille Yehia

BELO HORIZONTE – MG
FEVEREIRO DE 2015

Em memória de meu avô Francisco (Chicute) e da minha sogra Lucielena (Tia Lulu).

*Dedico este trabalho aos meus pais Antônia e Jesus, à minha irmã Jeseanne,
aos meus sobrinhos Danilo e Lucas e à minha amada esposa Débora
Mendonça.*

AGRADECIMENTOS

Agradeço à Deus por sempre me fortalecer nos momentos felizes e difíceis durante esta jornada.

Aos meus pais Antônia e Jesus e a minha esposa Débora Mendonça pela compreensão, consolo e motivação.

Aos meus orientadores Luciano de Errico e Hani Camille pela paciência, conselhos e ajuda durante a minha jornada do doutorado.

Aos meus amigos Dona Norma, Pedro Dias e Álvaro Lemos pelas dicas e pelo apoio durante minhas estadias em Belo Horizonte.

Aos meus colegas de doutorado José Ricardo, Ricardo Erikson, Gilberto, Guilherme, Daniel Patrick e Ruan, e de trabalho Fábio, Alison, Mychelle, Cláudio Eduardo, Cláudio Júnior, Admilson, Lúcio, Gustavo, Vitor, Fágner, Sabrina, Fabrício, Philipe, Áurea, Ângela, Sancha, Ivana, Diego e Leandro pelo apoio e incentivo a minha pesquisa.

A professora Regina Mota pelas longas discursões acerca do meu trabalho.

Aos professores Pinheiro e Vicente pela oportunidade de ingresso no programa de doutorado da Universidade Federal de Minas Gerais (UFMG) através do Dinter entre o Instituto Federal de Educação, Ciência e Tecnologia do Amazonas (Ifam) e UFMG.

Ao Centro de Pesquisa e Desenvolvimento de Tecnologia Eletrônica e da Informação (Ceteli) localizado na Universidade Federal do Amazonas (Ufam) pelo espaço cedido para a realização das avaliações subjetivas, bem como às pessoas que responderam e avaliaram a qualidade dos vídeos.

Ao Centro de Ciência, Tecnologia e Inovação do Pólo Industrial de Manaus (CTPIM) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo suporte financeiro concebido pela bolsa de estudo.

RESUMO

O cenário atual aponta para o consumo crescente de serviços relacionados a conteúdos multimídia, especialmente de vídeo sob demanda e televisão IP, por dispositivos móveis com capacidade de reproduzir vídeos de alta definição e em qualquer lugar. No entanto, o sucesso desses serviços está fortemente relacionado com a qualidade do vídeo percebida pelo usuário, também conhecida como a qualidade da experiência (*quality of experience* – QoE). Neste trabalho é proposto um sistema de gerenciamento da qualidade de experiência orientada à transmissão de vídeos para dispositivos móveis. São caracterizados os sistemas de *streaming* de vídeo, assim como os elementos que compõem sua arquitetura. As formas de avaliação de qualidade de um vídeo são descritas, bem como as técnicas utilizadas para incrementar esta qualidade. Com base nesses estudos, é detalhado o desenvolvimento de um gerenciador de qualidade de vídeos para um serviço de *streaming* em dispositivos móveis, levando em consideração informações sobre a transmissão, as características do dispositivo móvel e a codificação do vídeo. Como a escolha da codificação do vídeo é realizada no servidor de *streaming*, esse gerenciamento não sobrecarrega o dispositivo móvel. Para seu funcionamento, o gerenciador atua com o auxílio de dois preditores: um para a escolha da codificação, quando o conteúdo é escolhido pela primeira vez; outro para ajuste automático da qualidade de vídeo de acordo com as informações coletadas durante a reprodução do conteúdo no dispositivo móvel. Para a criação desses preditores, vários experimentos foram realizados para a escolha das informações necessárias. Com uma proposta de tratamento das informações, os preditores obtiveram uma melhora de resultado de 22% em média, quando comparado com *Peak Signal-to-Noise Ratio* (PSNR).

Palavras-chave: serviço de *streaming* de vídeos, qualidade de experiência, gerenciamento automatizado.

ABSTRACT

The present scenario indicates an increased access to multimedia content services, especially video on demand and IP television, by mobile devices capable of playing high-definition videos anywhere. However, the success of these services is closely related to the video quality perceived by the user, also known as the quality of experience (QoE). This work proposes a quality management system for video transmission to mobile devices. The video streaming systems are characterized as well as the elements of its architecture. The different forms of video quality evaluation and the techniques used to improve this quality are described. Based on these studies, the development of a quality video manager for a streaming service on mobile devices is detailed, taking into account information about the transmission, the mobile device and the video encoding. As the video encoding choice is performed in the streaming server, this management does not overload the mobile device. For its operation, two predictors are required: one for the encoding choice, at the beginning, when the content is chosen; another for the automatic adjustment of the video quality according to the information collected during content playback on the mobile device. Several experiments were conducted for choosing the relevant features to be considered by these predictors. With a proposal of information treatment, an average improvement of 22% was achieved in the results, compared with Peak Signal-to-Noise Ratio (PSNR).

Keywords: *video streaming service, quality of experience, automated management.*

LISTA DE FIGURAS

Fig. 1	O que acontece na Internet em 60 segundos.	26
Fig. 2	Elementos que compõem uma arquitetura de serviço de vídeos em redes IP.	27
Fig. 3	Comparação entre MPEG-2, MPEG-4 e H.264.	29
Fig. 4	Exemplo da estrutura de um GoP e a relação entre as imagens.. . . .	30
Fig. 5	Formas de transmissão sobre uma rede IP.	34
Fig. 6	Exemplo de troca de mensagens entre servidor e cliente utilizando os protocolos RTSP, RTP e RTCP.	37
Fig. 7	Exemplo de transmissão de conteúdos através do <i>streaming</i> adaptativo.	38
Fig. 8	Exemplos de <i>smart TVs</i> , <i>smartphones</i> e <i>tablets</i>	38
Fig. 9	Exemplo de falha durante a transmissão de um vídeo.	46
Fig. 10	Modelo conceitual do processo de formação da opinião sobre qualidade.	47
Fig. 11	Uma plausível relação entre QoS e QoE.	48
Fig. 12	Fatores que podem influenciar na avaliação da QoE.	49
Fig. 13	Diagrama para avaliação da qualidade de vídeos.	49
Fig. 14	Exemplos de artefatos de erro durante a criação dos vídeos distorcidos.	52
Fig. 15	Exemplos da tabela de Snellen.	53
Fig. 16	Exemplos de placas utilizadas no teste de Ishihara.	54
Fig. 17	Diagrama utilizado pelas métricas classificadas como referência completa para avaliar a qualidade do vídeo distorcido.	59
Fig. 18	Diagrama utilizado pelas métricas classificadas como referência reduzida para avaliar a qualidade do vídeo distorcido.	61
Fig. 19	Diagrama utilizado pelas métricas classificadas como sem referência para avaliar a qualidade do vídeo distorcido.	63
Fig. 20	Arquitetura de transmissão de vídeos em que o sistema proposto atua.	70
Fig. 21	Etapas seguidas para a elaboração do sistema proposto.	72
Fig. 22	Informações espacial e temporal dos vídeos de referência selecionados para investigação.	72

Fig. 23	<i>Snapshot</i> dos vídeos selecionados para investigação.	73
Fig. 24	Informações espacial e temporal dos vídeos de referência selecionados.	75
Fig. 25	<i>Snapshot</i> dos vídeos selecionados para as avaliações subjetivas. . . .	76
Fig. 26	Representação do modelo de perda de pacotes utilizando Gilbert-Elliott.	78
Fig. 27	Fluxograma utilizado para avaliação subjetiva dos vídeos distorcidos. .	81
Fig. 28	Dispositivos utilizados nas avaliações subjetivas.	81
Fig. 29	Ordem de apresentação dos vídeos.	82
Fig. 30	Ambiente utilizado nas avaliações subjetivas.	83
Fig. 31	Sistema gerenciador de QoE elaborado neste trabalho.	88
Fig. 32	Fluxograma empregado para elaboração dos preditores utilizados no sistema proposto neste trabalho.	89
Fig. 33	Fluxograma da implementação dos experimentos.	96
Fig. 34	Topologia da rede utilizada nas simulações.	96
Fig. 35	Componentes utilizados no NS e EvalVid para representar a topologia mostrada na Fig. 34.	97
Fig. 36	Aplicativo criado para coleta das notas nas avaliações subjetivas. . . .	98
Fig. 37	Influência do <i>bitrate</i> no vídeo Akiyo de acordo com o GoP e MTU. . . .	99
Fig. 38	Influência do <i>bitrate</i> no vídeo Tempete de acordo com o GoP e MTU. .	99
Fig. 39	Influência do <i>bitrate</i> no vídeo Flower de acordo com o GoP e MTU. . .	100
Fig. 40	Influência da perda de pacotes no vídeo Akiyo com <i>bitrate</i> igual a 100 kbps de acordo com o GoP e MTU.	101
Fig. 41	Influência da perda de pacotes no vídeo Akiyo com <i>bitrate</i> igual a 500 kbps de acordo com o GoP e MTU.	101
Fig. 42	Influência da perda de pacotes no vídeo Akiyo com <i>bitrate</i> igual a 1000 kbps de acordo com o GoP e MTU.	102
Fig. 43	Influência da perda de pacotes no vídeo Akiyo com <i>bitrate</i> igual a 2000 kbps de acordo com o GoP e MTU.	102
Fig. 44	Influência da perda de pacotes no vídeo Tempete com <i>bitrate</i> igual a 100 kbps de acordo com o GoP e MTU.	103
Fig. 45	Influência da perda de pacotes no vídeo Tempete com <i>bitrate</i> igual a 500 kbps de acordo com o GoP e MTU.	103
Fig. 46	Influência da perda de pacotes no vídeo Tempete com <i>bitrate</i> igual a 1000 kbps de acordo com o GoP e MTU.	104
Fig. 47	Influência da perda de pacotes no vídeo Tempete com <i>bitrate</i> igual a 2000 kbps de acordo com o GoP e MTU.	104

Fig. 48	Influência da perda de pacotes no vídeo Flower com <i>bitrate</i> igual a 100 kbps de acordo com o GoP e MTU.	105
Fig. 49	Influência da perda de pacotes no vídeo Flower com <i>bitrate</i> igual a 500 kbps de acordo com o GoP e MTU.	105
Fig. 50	Influência da perda de pacotes no vídeo Flower com <i>bitrate</i> igual a 1000 kbps de acordo com o GoP e MTU.	106
Fig. 51	Influência da perda de pacotes no vídeo Flower com <i>bitrate</i> igual a 2000 kbps de acordo com o GoP e MTU.	106
Fig. 52	Distribuição dos entrevistados de acordo com a escolaridade.	110
Fig. 53	Distribuição de dispositivos utilizados para assistir vídeos provenientes da Internet.	111
Fig. 54	Distribuição de espera pela normalidade de exibição do vídeo após o surgimento de alguma falha por tipo de serviço.	111
Fig. 55	Histograma das avaliações subjetivas coletadas com vídeos distorcidos (sem introdução de erros).	112
Fig. 56	Distribuição das avaliações subjetivas em cada vídeo distorcido (sem introdução de erros).	113
Fig. 57	Distribuição das avaliações subjetivas dos vídeos distorcidos (sem introdução de erros) coletadas no <i>smartphone</i> com o seu intervalo de confiança.	114
Fig. 58	Distribuição das avaliações subjetivas dos vídeos distorcidos (sem introdução de erros) coletadas no <i>tablet</i> com o seu intervalo de confiança.	114
Fig. 59	Histograma das avaliações subjetivas coletadas com vídeos distorcidos (com introdução de erros).	115
Fig. 60	Distribuição das avaliações subjetivas em cada vídeo distorcido (com introdução de erros).	115
Fig. 61	Distribuição das avaliações subjetivas dos vídeos distorcido (com introdução de erros) coletadas no <i>smartphone</i> com o seu intervalo de confiança.	116
Fig. 62	Distribuição das avaliações subjetivas dos vídeos distorcido (com introdução de erros) coletadas no <i>tablet</i> com o seu intervalo de confiança.	116
Fig. 63	Resultados obtidos para regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos sem introdução de erros com o uso do método de amostragem <i>random subsampling</i>	120

Fig. 64	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos sem introdução de erros (1 de 3).	120
Fig. 65	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos sem introdução de erros (2 de 3).	121
Fig. 66	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos sem introdução de erros (3 de 3).	122
Fig. 67	Resultados obtidos para regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos sem introdução de erros com o uso do método de amostragem <i>cross validation</i>	124
Fig. 68	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>cross validation</i> para os vídeos distorcidos sem introdução de erros (1 de 3).	124
Fig. 69	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>cross validation</i> para os vídeos distorcidos sem introdução de erros (2 de 3).	125
Fig. 70	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>cross validation</i> para os vídeos distorcidos sem introdução de erros (3 de 3).	126
Fig. 71	Resultados obtidos para regressão logarítmico-linear mais tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos com introdução de erros com o uso do método de amostragem <i>random subsampling</i>	130
Fig. 72	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos com introdução de erros (1 de 3).	130

Fig. 73	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos com introdução de erros (2 de 3).	131
Fig. 74	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>random subsampling</i> para os vídeos distorcidos com introdução de erros (3 de 3).	132
Fig. 75	Resultados obtidos para regressão logarítmico-linear mais tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos com introdução de erros com o uso do método de amostragem <i>cross validation</i>	133
Fig. 76	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>cross validation</i> para os vídeos distorcidos com introdução de erros (1 de 2).	134
Fig. 77	Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem <i>cross validation</i> para os vídeos distorcidos com introdução de erros (2 de 2).	135
Fig. 78	Relação entre os componentes que compõem o sistema proposto. . . .	136
Fig. 79	Informações utilizadas pelo preditor de qualidade de vídeo sem erros introduzidos durante a transmissão.	137
Fig. 80	Informações utilizadas pelo preditor de qualidade de vídeo com erros introduzidos durante a transmissão.	137
Fig. 81	Gerenciador de QoE proposto atuando para o primeiro cenário. . . .	138
Fig. 82	Gerenciador de QoE proposto atuando para o segundo cenário. . . .	138
Fig. 83	Tabela de Snellen utilizada nas avaliações de acuidade visual. . . .	170
Fig. 84	Placas utilizadas no exame simplificado do teste de Ishihara (1 de 3). .	173
Fig. 85	Placas utilizadas no exame simplificado do teste de Ishihara (2 de 3). .	174
Fig. 86	Placas utilizadas no exame simplificado do teste de Ishihara (3 de 3). .	175

LISTA DE TABELAS

Tab. 1	Comparação entre os padrões MPEG-2 e H.264.	30
Tab. 2	Comparação entre as tecnologias utilizadas em redes sem fios.	33
Tab. 3	Comparação entre algumas <i>smart TVs</i> encontradas no mercado.	39
Tab. 4	Comparativo de alguns <i>smartphones</i> existentes no mercado (1 de 2).	40
Tab. 5	Comparativo de alguns <i>smartphones</i> existentes no mercado (2 de 2).	41
Tab. 6	Comparativo de alguns <i>tablets</i> existentes no mercado (1 de 2).	42
Tab. 7	Comparativo de alguns <i>tablets</i> existentes no mercado (2 de 2).	43
Tab. 8	Lista de bases de vídeos de referência disponível publicamente.	50
Tab. 9	Escala de qualidade.	55
Tab. 10	Escala de degradação.	55
Tab. 11	Escala de qualidade contínua.	56
Tab. 12	Escala de comparação utilizada no SCACJ.	58
Tab. 13	Análise comparativa dos métodos utilizados na abordagem subjetiva.	58
Tab. 14	Conversão do PSNR para o MOS.	61
Tab. 15	Análise comparativa dos métodos utilizados na abordagem objetiva.	64
Tab. 16	Análise comparativa dos métodos utilizados na abordagem híbrida.	67
Tab. 17	Informações sobre os vídeos escolhidos para a investigação.	74
Tab. 18	Informações sobre os vídeos escolhidos para as avaliações subjetivas.	75
Tab. 19	Parâmetros utilizados para a criação dos vídeos distorcidos utilizados nas avaliações subjetivas para o primeiro cenário.	77
Tab. 20	Parâmetros utilizados para a criação dos vídeos distorcidos utilizados nas avaliações subjetivas para o segundo cenário.	78
Tab. 21	Parâmetros utilizados para a criação dos vídeos distorcidos utilizados na investigação da variação do GoP e MTU para o primeiro cenário.	79
Tab. 22	Parâmetros utilizados para a criação dos vídeos distorcidos utilizados na investigação da variação do GoP e MTU para o segundo cenário.	79
Tab. 23	Características dos dispositivos utilizados nas avaliações subjetivas.	81

Tab. 24	Ordem de apresentação dos vídeos distorcidos (sem introdução de erros) nas avaliações subjetivas.	84
Tab. 25	Ordem de apresentação dos vídeos distorcidos (com introdução de erros) nas avaliações subjetivas.	85
Tab. 26	Algoritmos utilizados para o treinamento dos preditores.	92
Tab. 27	Parâmetros utilizados nas simulações.	97
Tab. 28	Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Akiyo de acordo com o <i>bitrate</i> , GoP e MTU.	107
Tab. 29	Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Tempete de acordo com o <i>bitrate</i> , GoP e MTU.	108
Tab. 30	Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Flower de acordo com o <i>bitrate</i> , GoP e MTU.	109
Tab. 31	Agrupamento dos atributos utilizados para criação do preditor de qualidade de vídeos distorcidos sem introdução de erros.	118
Tab. 32	Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem <i>random subsampling</i> para os vídeos distorcidos sem introdução de erros.	119
Tab. 33	Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem <i>cross validation</i> para os vídeos distorcidos sem introdução de erros.	123
Tab. 34	Agrupamento dos atributos utilizados para criação do preditor de qualidade de vídeos distorcidos com introdução de erros.	128
Tab. 35	Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem <i>random subsampling</i> para vídeos com introdução de erros.	129
Tab. 36	Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem <i>cross validation</i> para vídeos com introdução de erros.	133
Tab. 37	Valores lidos de acordo com a deficiência visual de cores.	172

LISTA DE ALGORITMOS

Alg. 1	Proposta de tratamento aplicando uma linha de tendência em cada atributo nas bases de treinamento e teste.	91
--------	--	----

LISTA DE ABREVIATURAS E SIGLAS

3GPP	<i>3rd Generation Partnership Project</i>
AAC	<i>Advanced Audio Codec</i>
AAC-LC	<i>Advanced Audio Codec Low Complexity</i>
AAC-SSR	<i>Advanced Audio Codec Scalable Sampling Rate</i>
ACR	<i>Absolute Category Rating</i>
ACR-HR	<i>Absolute Category Rating with Hidden Reference</i>
ANN	<i>Redes Neurais Artificiais (Artificial Neural Networks)</i>
ASU	<i>Arizona State University</i>
AVC	<i>Advanced Video Coding</i>
CDA	<i>Análise de Discriminante Canônico (Canonical Discriminant Analysis)</i>
CDMA	<i>Code Division Multiple Access</i>
CS	<i>Estímulo por Comparação (Comparison Stimulus)</i>
DASH	<i>Dynamic Adaptive Streaming over HTTP</i>
DMOS	<i>Differential Mean Opinion Score</i>
DoS	<i>Negação de Serviço (Denial of Service)</i>
DRM	<i>Segurança e Direitos Autorais (Digital Rights Management)</i>
DS	<i>Estímulo Duplo (Double Stimulus)</i>
DSCQS	<i>Double Stimulus Continuous Quality Scale</i>
DSIS	<i>Double Stimulus Impairment Scale</i>
ED	<i>Enhanced Definition</i>
EPFL	<i>École Polytechnique Fédérale de Lausanne</i>
FEC	<i>Forward Error Correction</i>
FR	<i>Referência Completa (Full-Reference)</i>
GoP	<i>Grupo de Imagens (Group of Picture)</i>
GoS	<i>Grau de Serviço (Grade of Service)</i>
HD	<i>High Definition</i>
HE-AAC	<i>High Efficiency Advanced Audio Codec</i>
HEVC	<i>High Efficiency Video Coding</i>
HSPA+	<i>Evolved High Speed Packet Access</i>

HVS	Sistema Visual Humano (<i>Human Visual System</i>)
IETF	<i>Internet Engineering Task Force</i>
IP	<i>Internet Protocol</i>
IPTV	Televisão por Protocolo de Internet (<i>Internet Protocol Television</i>)
IRCCyN	<i>Institut de Recherche en Communication et Cybernétique de Nantes</i>
ISP	Provedor de Acesso à Internet (<i>Internet Server Provider</i>)
ITU	<i>International Telecommunication Union</i>
LHS	Força Harmônica Local (<i>Local Harmonic Strength</i>)
LIVE	<i>Laboratory for Image & Video Engineering</i>
LTE	<i>Long Term Evolution</i>
MAE	Distância Absoluta Média (<i>Mean Absolute Error</i>)
MANN	Múltiplas Redes Neurais Artificiais (<i>Multiple Artificial Neural Network</i>)
MDC	Codificação Múltipla Descritiva (<i>Multiple Description Coding</i>)
MMSPG	<i>Multimedia Signal Processing Group</i>
MPLS	<i>Multiprotocol Label Switching</i>
MOS	Índice de Opinião Média (<i>Mean Opinion Score</i>)
MSE	Erro Quadrático Médio (<i>Mean Squared Error</i>)
MSSIM	<i>Multi Structural SIMilarity</i>
MTU	<i>Maximum Transmission Unit</i>
NQM	<i>Noise Quality Measure</i>
NQoS	QoS da Rede (<i>Network QoS</i>)
NR	Sem Referência (<i>No-Reference</i>)
NS	<i>Network Simulator</i>
PCA	Análise de Componentes Principais (<i>Principal Component Analysis</i>)
PCC	Coefficiente de Correlação de Pearson (<i>Pearson Correlation Coefficient</i>)
PNS	<i>Perceptual Noise Shaping</i>
Polimi	<i>Politecnico di Milano</i>
PQoS	Qualidade Percebida pelo Usuário (<i>Perceived QoS</i>)
PSNR	<i>Peak Signal-to-Noise Ratio</i>
PSQA	Avaliação de Qualidade Pseudo-Subjetiva (<i>Pseudo-Subjective Quality Assessment</i>)
Qualinet	<i>European Network on Quality of Experience in Multimedia Systems and Services</i>

GoE	Qualidade de Experiência (<i>Quality of Experience</i>)
GoR	Qualidade de Resiliência (<i>Quality of Resilience</i>)
GoS	Qualidade de Serviço (<i>Quality of Service</i>)
TRACE	<i>Video Trace Library</i>
RNN	Rede Neural Aleatória (<i>Random Neural Network</i>)
ROI	Regiões de Interesse (<i>Regions of Interest</i>)
RR	Referência Reduzida (<i>Reduced-Reference</i>)
RST	Teoria dos Conjuntos Aproximativos (<i>Rough Set Theory</i>)
SD	<i>Standard Definition</i>
SDN	<i>Software Defined Networks</i>
SDSCE	<i>Simultaneous Double Stimulus for Continuous Evaluation</i>
SI	Informação Espacial (<i>Spatial Information</i>)
SNR	<i>Signal-to-Noise Ratio</i>
SROCC	Coefficiente de Correlação de Postos de Spearman (<i>Spearman Rank Order Correlation Coefficient</i>)
SS	Estímulo Simples (<i>Single Stimulus</i>)
SSCQE	<i>Single Stimulus Continuous Quality Evaluation</i>
TCL	<i>Tool Command Language</i>
TI	Informação Temporal (<i>Temporal Information</i>)
TVD	TV Digital
TIFF	<i>Tagged Image File Format</i>
TNS	<i>Temporal Noise Shaping</i>
TUM	<i>Technische Universität München</i>
UMTS	<i>Universal Mobile Telecommunication System</i>
UQI	<i>Universal Quality Index</i>
URL	<i>Uniform Resource Locator</i>
VIFP	<i>Pixel Based VIF</i>
VoD	Vídeo sob Demanda (<i>Video on Demand</i>)
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
WiMAX	<i>Worldwide Interoperability for Microwave Access</i>
WLAN	Rede Local Sem Fio (<i>Wireless Local Area Network</i>)
WMAN	Redes Metropolitanas (<i>Wireless Metropolitan Area Network</i>)
WNSR	<i>Weighted Signal-to-Noise Ratio</i>

SUMÁRIO

Capítulo 1	Introdução	20
1.1	Definição do Problema	21
1.2	Objetivos	22
1.3	Organização do Trabalho	23
Capítulo 2	Aspectos Relevantes sobre Transmissão de Vídeos em Redes IP	25
2.1	Provedor de Conteúdos	28
2.2	Transmissão de Conteúdos	31
2.3	Recepção de Conteúdos	38
2.4	Conclusões	43
Capítulo 3	Avaliação da Qualidade de Vídeos	44
3.1	O que é Qualidade de Serviço (QoS)?	45
3.2	O que é Qualidade de Experiência (QoE)?	46
3.3	Metodologias Utilizadas na Avaliação da Qualidade de Vídeos	49
3.3.1	Abordagem Subjetiva	54
3.3.2	Abordagem Objetiva	58
3.3.3	Abordagem Híbrida	65
3.4	Conclusões	67
Capítulo 4	Gerenciamento da QoE na Transmissão de Vídeos em Redes Sem Fio	68
4.1	Delimitação do Problema	70
4.2	Metodologia	71
4.3	Conclusões	92
Capítulo 5	Experimentos e Resultados	94
5.1	Ferramentas Utilizadas	94

5.2 Implementação dos Experimentos	95
5.3 Resultados Obtidos	98
5.4 Conclusões	139
Capítulo 6 Conclusões	140
6.1 Resultados Obtidos	141
6.2 Propostas de Continuidade	143
6.3 Considerações Finais	146
Referências Bibliográficas	147
Apêndice A Publicações	163
Apêndice B Termo de Consentimento Livre e Informado	164
Apêndice C Formulário Sobre os Dados Pessoais do(a) Voluntário(a)	166
Apêndice D Teste de Acuidade Visual	169
Apêndice E Teste de Deficiência Visual de Cores	171

INTRODUÇÃO

“Oportunidades na vida vêm pela criação e não por acaso. Você mesmo, agora ou no passado, criou todas as oportunidades que passaram pelo seu caminho”

Autobiografia de um iogue

O rápido desenvolvimento de dispositivos móveis, de vídeo de alta definição e da infraestrutura de rede usada para *streaming* de vídeo exige uma permanente evolução das técnicas utilizadas para avaliar a qualidade de vídeo do conteúdo recebido pelo cliente em seu dispositivo. Na última década, o tráfego de voz interativa (*Voice over IP* – VoIP) foi adicionado ao tráfego de dados da rede tradicional (web, e-mail, transferência de arquivos). Hoje, VoIP é comum em redes IP e a tendência é a de rápido aumento do tráfego de vídeo, como vídeo sob demanda (VoD), WebTV e televisão IP (IPTV).

Em um serviço de IPTV, os conteúdos são organizados em uma grade de programação e distribuídos em redes gerenciáveis, baseadas no protocolo IP, além de proporcionar uma qualidade igual ou superior ao oferecido em uma transmissão de TV digital (TVD). Outra característica desse serviço é a possibilidade de oferecer milhares de canais, enquanto que em um sistema de televisão digital, devido a faixa limitada de espectro do sinal, podem ser oferecidos dezenas ou centenas de canais. No VoD, os conteúdos estão disponíveis em um servidor e o usuário os assiste quando ele quiser, após uma busca. A WebTV é semelhante ao IPTV com relação à organização dos conteúdos em uma grade de programação, mas os conteúdos são distribuídos em redes não gerenciáveis.

O sucesso dos serviços de distribuição de vídeos está associado à garantia de qualidade, ou seja, que os conteúdos sejam exibidos nos dispositivos dos usuários com um mínimo de falhas e atrasos. Para alcançar isso, o administrador da rede monitora informações da infraestrutura, tais como largura de banda, atraso, vazão, dentre outros, para proporcionar uma qualidade adequada para cada cliente. No entanto, devido à complexidade da rede, torna-se difícil gerenciá-la de modo que a qualidade seja garantida.

Quando se avalia a qualidade desse tipo de serviço em dispositivos móveis, além das condições da infraestrutura e da conexão, devem ser observadas as características de cada dispositivo, como tamanho da tela, quantidade de memória interna e capacidade do processador. Mais do que isso, é importante saber como está sendo exibido o conteúdo no dispositivo do usuário, pois de nada vale garantir a qualidade da conexão em um nível aceitável se a qualidade de vídeo do conteúdo exibido não está sendo satisfatória para o usuário.

Fundado nas concepções descritas, neste trabalho apresenta-se uma proposta de sistema que auxilia no gerenciamento da qualidade de um serviço de *streaming* de vídeo para dispositivos móveis. Esse sistema leva em consideração informações coletadas sobre os formatos de vídeo do conteúdo disponíveis no servidor de *streaming*, sobre a qualidade da comunicação e sobre o dispositivo móvel em que está sendo reproduzido o conteúdo.

1.1 Definição do Problema

Uma das dificuldades do gerenciamento de um serviço de *streaming* de vídeos é manter a qualidade de vídeo dos conteúdos recebidos nos dispositivos dos clientes. Hoje, com a popularização dos dispositivos móveis, essa dificuldade é maior ainda, pois existem diversas variações de características de dispositivos móveis.

Além disso, com a evolução das técnicas de codificação de vídeo, possibilita-se a escolha de várias resoluções de vídeo pelos clientes. No entanto, em muitas vezes, essa escolha pode interferir na qualidade percebida pelo usuário, pois podem ser transmitidos dados que não agregam na qualidade percebida pelo usuário.

Diante disso, neste trabalho foi proposto a elaboração de um gerenciador automatizado da qualidade de vídeos exibidos em dispositivos móveis. Foram levados em consideração somente vídeos sem áudio. Esse gerenciador

atua na escolha dos parâmetros oridundos da transmissão do conteúdo, dos formatos disponíveis do conteúdo no servidor de *streaming* e do dispositivo móvel, com o intuito de maximizar a qualidade de vídeo do conteúdo exibido no dispositivo móvel do cliente.

Outra característica desse gerenciador é a não atuação na rede, ou seja, ele não tem o papel de escolher as melhores rotas para a transmissão dos pacotes, origidados a partir do vídeo escolhido anteriormente, entre o servidor de *streaming* e o dispositivo móvel. Ele observa em tempo real os parâmetros de transmissão, de vídeo e do dispositivo móvel para selecionar o formato de vídeo de acordo com a situação atual da comunicação e proporcionar, assim, uma qualidade de vídeo percebida aceitável pelo usuário.

1.2 Objetivos

O objetivo geral foi avaliar a qualidade de vídeo de forma automatizada e em tempo real, a partir do monitoramento de informações acerca da transmissão do conteúdo, dos formatos de vídeo disponibilizados para o conteúdo escolhido e do dispositivo móvel, com o intuito de oferecer a codificação de vídeo do conteúdo que melhor se compatibiliza com as condições atuais da comunicação entre o servidor de *streaming* e o dispositivo móvel. Para isso, procurou-se responder os seguintes questionamentos:

- De acordo com os participantes das avaliações subjetivas, há diferença da qualidade de vídeo percebida pela pessoa enquanto assiste ao conteúdo em um *tablet* e *smartphone*?
- Quando o vídeo é exibido em tela cheia no *tablet* ou *smartphone*, será que a resolução de vídeo influencia na percepção da qualidade de vídeo?
- A respeito da perda de pacotes durante a transmissão do conteúdo, será que existe uma relação entre resolução de vídeo e taxa de perda de pacotes?
- Diante das informações disponibilizadas sobre os formatos, a transmissão e a recepção dos conteúdos em um serviço de *streaming* de vídeos, quais informações podem ser empregadas para elaborar um avaliador de qualidade de vídeo sem levar em consideração o vídeo original do conteúdo?

- Tendo em vista que a análise da qualidade de vídeo será realizada de tempos em tempos, qual o período a ser adotado para executar essa análise de forma automática?

A partir disso, foram observados os seguintes objetivos específicos:

- Investigar o impacto da percepção da qualidade de vídeo em dispositivos móveis por meio de avaliações subjetivas;
- Analisar quais fatores pertinentes à codificação do vídeo e à perda de pacotes influenciam na percepção da qualidade de vídeo pelas pessoas;
- Averiguar quais informações observadas na codificação, na transmissão e na recepção dos conteúdos são relevantes para conceber o gerenciador de qualidade de vídeo;
- Elaborar os componentes que integram o gerenciador de qualidade de vídeo utilizando técnicas de inteligência computacional.

1.3 Organização do Trabalho

Para facilitar a exposição deste trabalho, o texto foi organizado em seis capítulos, dos quais este foi o primeiro. No Capítulo 2, o processo de criação, distribuição e recepção de conteúdos em um serviço de *streaming* de vídeos são descritos, assim como os tipos de conteúdos, a codificação de áudio/vídeo e as técnicas utilizadas para transmissão e recepção dos conteúdos.

Um levantamento de trabalhos relacionados à avaliação da qualidade de um serviço *streaming* de vídeos é descrito no Capítulo 3. Nessa análise, são levadas em consideração duas formas de avaliação: a da qualidade da comunicação oferecida pela rede, conhecida como Qualidade de Serviço (*Quality of Service* – QoS), e a da qualidade do conteúdo entregue ao usuário, baseada na percepção humana, conhecida como Qualidade de Experiência (*Quality of Experience* – QoE). Também são mostradas algumas abordagens e técnicas utilizadas para melhorar a qualidade nesse tipo de serviço.

No Capítulo 4 é relatada a metodologia para o desenvolvimento do gerenciador de QoE para dispositivos móveis, bem como a delimitação do problema e a arquitetura em que esse gerenciador está inserido. Nele também são explicados as características dos vídeos utilizados nas avaliações subjetivas e os componentes que integram o gerenciador de QoE.

A configuração do ambiente e as ferramentas utilizadas para o desenvolvimento do gerenciador de QoE são explanadas no Capítulo 5. Também são detalhados os experimentos realizados durante este trabalho, bem como as análises dos resultados obtidos, principalmente com relação às avaliações subjetivas de vídeos sem e com introdução de erros gerados durante a transmissão em uma rede sem fio. Já as perspectivas de continuidade deste trabalho e as considerações finais são expostas no Capítulo 6.

ASPECTOS RELEVANTES SOBRE TRANSMISSÃO DE VÍDEOS EM REDES IP

*“Perder tempo em aprender coisas que não interessam,
priva-nos de descobrir coisas interessantes”*

Carlos Drummond de Andrade

Nos últimos anos, a Internet vem passando por uma grande transformação devido ao surgimento e crescimento de serviços relacionados a conteúdos multimídia (áudio, vídeo, imagens e textos). Isso se tornou possível graças ao avanço científico nas áreas de telecomunicação, processamento de dados, redes de computadores e armazenamento de informações.

Outro fato que contribuiu para esse cenário foi a popularização de dispositivos móveis (*notebooks*, *tablets* ou *smartphones*) e a disseminação de redes sem fio, seja em redes locais (WLANs) ou redes de telefonia celular (3G e 4G). Na Fig. 1 é ilustrada a diversidade de serviços disponíveis na Internet em decorrência dessa transformação, mostrando a quantidade de informações geradas e consumidas em 60 segundos durante o ano de 2014.

Essa nova tendência trouxe novos perfis de serviços, principalmente relacionados à transmissão de vídeos. Embora as formas mais comuns de transmissão de conteúdo para um receptor de TV sejam via radiodifusão, satélite ou cabo, hoje, graças à evolução das técnicas de compressão de conteúdos multimídia e à melhoria significativa na infraestrutura de telecomunicações, surgiram vários serviços de distribuição de conteúdos multimídia na Internet,

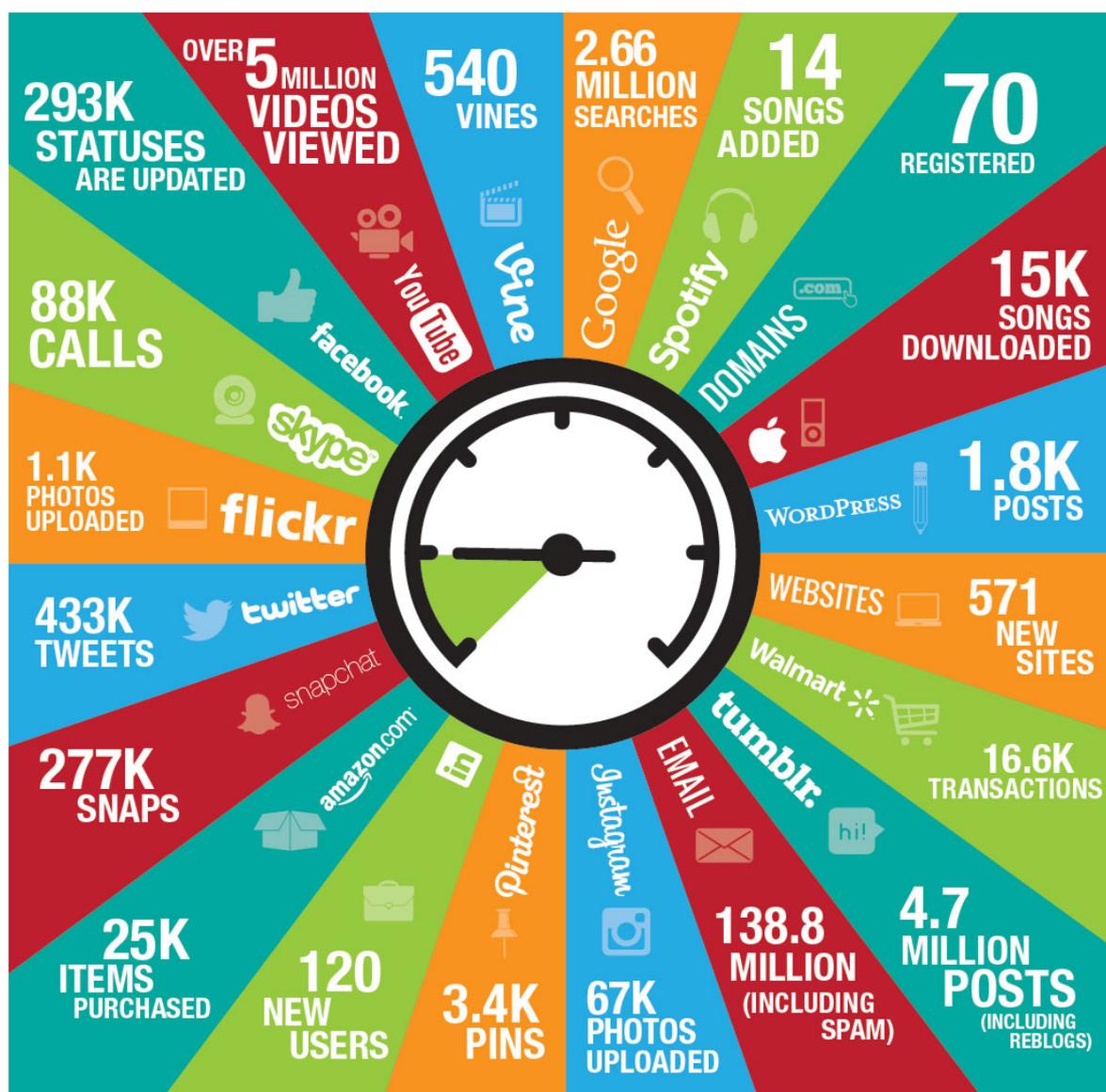


Figura 1: O que acontece na Internet em 60 segundos.

Fonte: Adaptado de [Qmee \(2015\)](#).

tais como YouTube, Vimeo, Livestream e, recentemente, NetFlix, Amazon Fire TV, Apple TV e Google TV. Acredita-se que, em poucos anos, 90% dos conteúdos transmitidos na Internet serão relacionados a vídeos assistidos por mais de um bilhão de pessoas ([BEGEN et al., 2011](#); [CISCO, 2013](#)).

Esses serviços podem ser classificados como distribuição de vídeo sob demanda (*Video on Demand – VoD*), canais de TV distribuídos em redes não gerenciáveis (*WebTV*) ou pacotes de canais de TV com qualidade de TV paga e distribuídos em redes gerenciáveis (*Internet Protocol TeleVision – IPTV*) ([SIMPSON, 2008](#)).

Enquanto que no VoD os conteúdos estão disponíveis em um servidor e o usuário os escolhe e assiste quando desejar, sendo possíveis controles de exibição como pausar, avançar e retroceder, no WebTV e IPTV são enviados canais com uma programação contínua, definida pelo produtor do conteúdo e sem esses controles. Em comum, esses serviços são transmitidos sobre redes IP (seja na Internet ou em redes privadas de provedores de serviço) utilizando a técnica de transmissão contínua conhecida como *streaming*. Em *streaming*, o conteúdo é exibido para o usuário à medida que é enviado pela fonte, sem necessidade de armazenamento do arquivo completo para posterior exibição. Em geral, um *buffer* é utilizado para armazenar uma certa quantidade de conteúdo, em média 10s de conteúdo, antes de sua exibição, de forma a minimizar flutuações e falhas esporádicas na transmissão pela rede.

Uma arquitetura necessária para disponibilizar um serviço de vídeos em rede IP é composta por diversos elementos (ver Fig. 2). São eles:

- **Provedor de Conteúdos:** os conteúdos são criados, editados, codificados e armazenados em uma base multimídia, a qual é disponibilizada por intermédio de um servidor de *streaming*;
- **Transmissão de Conteúdos:** os conteúdos são encapsulados em pacotes IP (*Internet Protocol*) e são transmitidos na Internet ou em uma rede de Provedor de Acesso à Internet (*Internet Server Provider – ISP*);
- **Recepção de Conteúdos:** os conteúdos são visualizados nos receptores dos clientes, como televisão, computador pessoal ou dispositivos móveis.

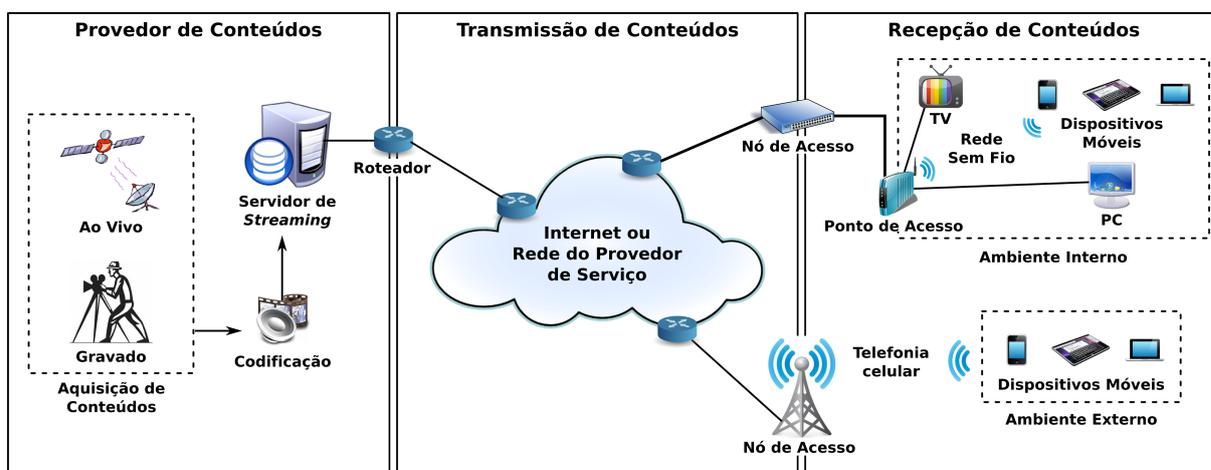


Figura 2: Elementos que compõem uma arquitetura de serviço de vídeos em redes IP.

2.1 Provedor de Conteúdos

No provedor de conteúdo podem ser disponibilizados vídeos gravados ou ao vivo, além do que, eles podem ser oferecidos em formato 2D, onde tem-se a percepção de largura e altura; 3D, que possibilita o incremento do 2D com percepção de profundidade dos objetos da cena; ou 4D, que utiliza a percepção de todos os sentidos do corpo humano (visão, audição, tato, paladar e olfato).

Antes que esses conteúdos sejam disponibilizados em servidores de *streaming*, técnicas de codificação são aplicadas no áudio e vídeo, que avaliam quais informações são relevantes a partir da percepção auditiva e visual humana, de modo que, o que não for perceptível será descartado durante o processo de compressão. Essas técnicas são definidas pelo *Motion Pictures Experts Group* (MPEG) e pela *International Telecommunication Union* (ITU). Segundo Paul (2011), os principais padrões de codificação utilizados em um serviço de transmissão de vídeos são o MPEG-2 e MPEG-4.

O MPEG-2 foi homologado em 1996 pelo MPEG e, desde então, tem sido utilizado para vários propósitos, como produção e transmissão de conteúdos pelas emissoras de televisão e visualização de vídeos pela Internet. O MPEG-2 é composto por perfis e níveis para codificar e decodificar o vídeo. Os perfis são classificados em Simples, Principal, *Signal-to-Noise Ratio* (SNR), Espacial e Alto, enquanto que os níveis são caracterizados pelas resoluções e taxas de transmissão dos vídeos (O'DRISCOLL, 2008).

O padrão MPEG-4, sucessor natural do MPEG-2, foi homologado em 2000 pelo MPEG. Ele é composto por 22 especificações, que oferecem suporte a uma grande variedade de formatos multimídia, como conteúdos de áudio e vídeo, *framework* para animação e modelos para implementação em *hardware* e *software* (O'DRISCOLL, 2008). O MPEG-4 Parte 2 e o MPEG-4 Parte 10 são as especificações utilizadas para a codificação de vídeo.

A ideia principal dos algoritmos utilizados pelo MPEG-4 é identificar no vídeo elementos que são estáticos (fundo do vídeo) e dinâmicos (movimentação de pessoas) (KUMAR, 2010). Assim, os elementos estáticos e dinâmicos são codificados separadamente e, para cada um deles, são empregados algoritmos específicos para cada tipo de conteúdo. Em alguns casos, pode-se reduzir a taxa de transmissão de conteúdos com alta resolução quase pela metade, quando comparado ao MPEG-2. Outra característica do MPEG-4 é a possibilidade de adicionar outros elementos ao conteúdo, como objetos interativos, descrição da localização dos objetos, animações etc. (HWANG, 2010).

O MPEG-4 Parte 10, conhecido como H.264 ou *advanced video coding* (AVC), foi introduzido no mercado em 2002 e o seu grande diferencial, quando comparado ao MPEG-4 Parte 2, é a compressão ainda maior de vídeos de alta resolução, a uma taxa de transmissão de dados baixa. Isso se dá ao fato do uso de várias técnicas inovadoras, tais como predição de imagem intraespacial, compensação de movimento, otimização da taxa de distorção e codificação de entropia, que possibilitam alto grau de adaptabilidade, sofisticação e diversificação em comparação com outros codificadores existentes para vídeos (HWANG, 2010).

Na Fig. 3 é exemplificado o uso dos padrões MPEG-2, MPEG-4 (Parte 2) e H.264 em um vídeo com duração de 120 minutos, resolução espacial de 720x480 e taxa de transmissão a 768 kbps. Foram utilizados como parâmetros: largura de banda (Mbps), armazenamento (GBytes) e tempo de *download* (minutos).

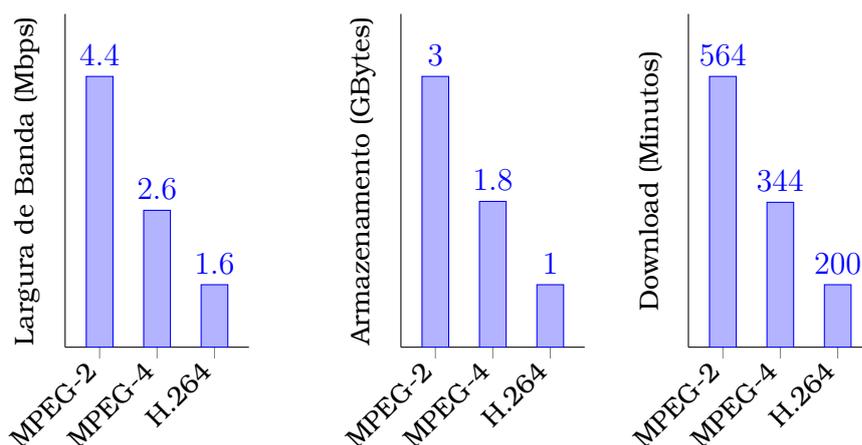


Figura 3: Comparação entre MPEG-2, MPEG-4 e H.264.

Fonte: adaptado de Kumar (2010).

Uma peculiaridade entre os padrões MPEG é a estrutura do vídeo. Um vídeo é representado por um ou vários grupos de imagens (*Group of Pictures* – GoP). Em cada GoP pode incluir até três tipos de imagens: *Intra-coded* (I), *Predictive-coded* (P) e *Bi-directionally predictive-coded* (B). Uma imagem I é codificada independente das outras imagens e a compressão é fundamentada somente nas informações contida nela. As imagens P utilizam informações da imagem anterior a ela (que pode ser do tipo I ou P). Por fim, as imagens B empregam informações das imagens I ou P precedentes e/ou posteriores (SIMPSON, 2008; BOVIK, 2009). O GoP possui as seguintes propriedades: deve

possuir pelo menos uma imagem do tipo I e seu tamanho é definido a partir da distância entre duas imagens do tipo I. Também pode ser determinada a quantidade de imagens B entre as imagens I e P. Usualmente, utiliza-se entre 12 e 15 para tamanho do GoP e entre 1 e 3 para a quantidade de imagens B (WU E RAO, 2006). Na Fig. 4 é ilustrada a sequência inicial de um vídeo com GoP de tamanho igual a 12.

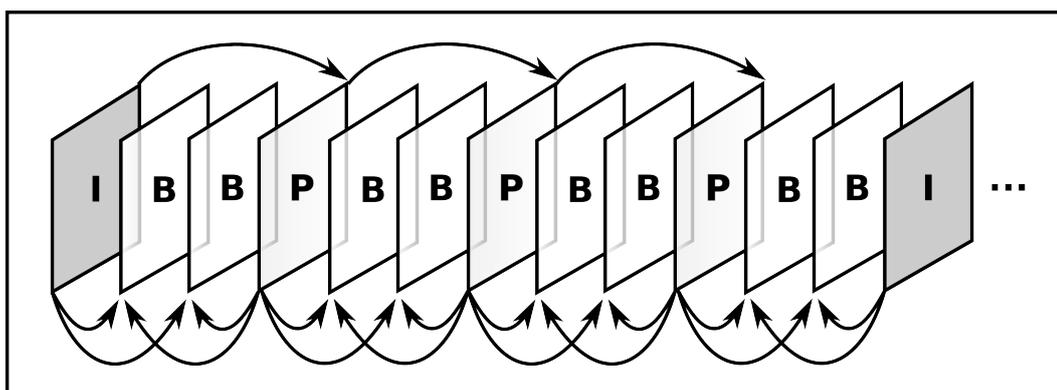


Figura 4: Exemplo da estrutura de um GoP e a relação entre as imagens..

Outros fatores, que podem contribuir durante o processo de compressão do vídeo, são a taxa de transmissão de bits por segundo (*bitrate*), que pode ser constante ou variável, e a resolução espacial. As principais resoluções utilizadas são: *Standard Definition* (SD), com resoluções de 480i e 576i; *Enhanced Definition* (ED), com resoluções de 480p e 576p; e *High Definition* (HD), com resoluções de 720p, 1080i e 1080p. Na Tabela 1 é mostrada uma comparação entre os padrões MPEG-2 e H.264, levando em conta a taxa de bits necessária para a transmissão de cada tipo de resolução.

Tabela 1: Comparação entre os padrões MPEG-2 e H.264.

Codificação	Resolução	Taxa de Bits
MPEG-2	SD/ED	3 a 8 Mbps
MPEG-2	HD	15 a 18 Mbps
H.264	SD/ED	1,5 a 8 Mbps
H.264	HD	8 a 12 Mbps

Fonte: (GROUP, 2015).

Com relação à compressão de áudio, os principais métodos utilizados são o *advanced audio codec* (AAC) e *high efficiency advanced audio codec* (HE-AAC) (PAUL, 2011).

O AAC foi criado pelo MPEG e desenvolvido a partir do MP3 (MPEG-1 Parte 3). Ele é uma das especificações do MPEG-2 (Parte 7) e pode ser classificado em três perfis: baixa complexidade (*low complexity* – AAC-LC); principal (AAC); e taxa de amostragem escalável (*scalable sampling rate* – AAC-SSR) (KUMAR, 2010). O método AAC tem suporte de até 48 canais e emprega diversas ferramentas que ajudam na compressão do áudio: *temporal noise shaping* (TNS), que minimiza o efeito de propagação temporal; e *perceptual noise shaping* (PNS), que escolhe os bits mais significativos na quantização do áudio original, tornando a compressão mais eficaz (SALOMON E MOTTA, 2010).

Já o HE-AAC pertence ao padrão MPEG-4. Ele é semelhante ao AAC, diferenciando-se pelo fato de exigir uma licença para a sua utilização. A sua atual versão, o HE-AAC v2, possui um método denominado parâmetro estéreo, que seleciona os sinais estéreos do áudio e os comprime com mais eficácia (PAUL, 2011).

Ao final do processo de codificação, os vídeos são armazenado no servidor de *streaming*. Esses vídeos são, então, disponibilizados para transmissão em redes IP. Para isso, existem diversos *softwares* no mercado com essa funcionalidade. Dentre eles, estão: *Darwin QuickTime Streaming Server* (DARWIN, 2015), *FFserver* (FFSERVER, 2015), *VLC* (VLC, 2015), *LIVE555 Media Server* (LIVE555, 2015), *Adobe Media Server* (ADOBE, 2015) e o *Windows Streaming Media Services* (WINDOWS, 2015).

2.2 Transmissão de Conteúdos

Para que o usuário tenha acesso ao serviço de vídeos sobre IP, é necessário que ele se comunique com um nó de acesso, cujo papel é estabelecer uma comunicação entre o consumidor e a rede do provedor. Neste caso, podem ser usadas redes locais sem fio ou redes de telefonia celular, permitindo a distribuição de serviços de vídeo para dispositivos móveis. As redes sem fio podem ser classificadas em dois grupos: curto alcance (até 300 m) e longo alcance (alguns quilômetros). Dentre as tecnologias de curto alcance estão: Bluetooth, WiBree, WirelessHD e WiFi.

A implementação mais popular do Bluetooth tem alcance de sinal pequeno (até 10 m), taxa de transmissão de dados variando até 1 Mbps e suporte a qualidade de serviço (FERRO E POTORTI, 2005). Um possível sucessor do Bluetooth é o WiBree, que diferencia-se no consumo de até 10 vezes menor de energia (PEI *et al.*, 2008).

O WirelessHD é um novo padrão de comunicação sem fio, o qual permite o *streaming* de conteúdos multimídia de alta definição entre dispositivos domésticos, com alcance de até 10 m. Ele suporta medição de qualidade e a sua taxa de transmissão varia de 10 Mbps até 28 Gbps (WIRELESSHD, 2015).

O WiFi (IEEE 802.11) foi criado em 1999 por um grupo de empresas com o objetivo de compartilhar conteúdos através de um padrão de comunicação sem fio (WiFi, 2015). Ele possui várias taxas de transmissão de dados: 11 Mbps (802.11b), 54 Mbps (802.11a e 802.11g), 600 Mbps (802.11n), 500 Mbps (802.11ac) e 1 Gbps (802.11ad). Os padrões a/b/g/n utilizam a técnica do melhor esforço para o envio de dados, sendo que o alcance do sinal pode variar de 50 m (ambientes fechados) até 100 m (ambientes abertos).

As tecnologias de redes sem fio de longo alcance mais comuns são utilizadas em comunicação de dados via sistemas de telefonia celular, compreendendo as redes de terceira geração (3G) e quarta geração (4G). Os padrões reconhecidos como 3G são o *Code Division Multiple Access* (CDMA) e o *Universal Mobile Telecommunication System* (UMTS). Já a 4G é representada pelos padrões *Worldwide Interoperability for Microwave Access* (WiMAX), *Long Term Evolution* (LTE) e *Evolved High Speed Packet Access* (HSPA+).

O surgimento do CDMA e UMTS tornou possível a transmissão de conteúdos multimídia para telefones celulares, devido ao aumento da taxa de transmissão de dados para 300 kbps e 2 Mbps, respectivamente (SAUTER, 2009). Seu ponto negativo é a ausência de medição de qualidade.

O WiMAX é um padrão de comunicação para redes metropolitanas (*Wireless Metropolitan Area Network* – WMAN) (WiMAX, 2015). Ele possui alcance de até 30 quilômetros, possibilita a recepção de conteúdos em dispositivo móvel em movimento por até 120 quilômetros por hora, possui métodos para melhorar a qualidade do serviço e oferece taxa de transmissão de até 3,74 Mbps por MHz (KUMAR, 2008).

O HSPA+ é uma evolução do UMTS. Ele trouxe como benefício um melhor gerenciamento do consumo de energia nos aparelhos celulares e o aumento da taxa de transmissão de dados, que pode chegar até 42 Mbps (SAUTER, 2009). Já o LTE, concorrente direto do HSPA+, tem como grande aliado a sua taxa de transmissão de dados, que pode alcançar até 100 Mbps. Dentre as características em comum do HSPA+ e LTE estão o uso do IP para identificar os aparelhos celulares e suporte a garantia de medição de qualidade.

Na Tabela 2 é apresentada uma comparação entre as tecnologias utilizadas em redes sem fio, levando em consideração alcance, taxa de transmis-

são, frequência utilizada e suporte a medição de qualidade.

Tabela 2: Comparação entre as tecnologias utilizadas em redes sem fios.

Tecnologia	Alcance	Taxa de Transmissão	Frequência Utilizada	Suporte a Medição de Qualidade
Bluetooth (IEEE 802.15)	10 m	3 Mbps	2,4 GHz	Sim
Wibree (IEEE 802.15.2)	10 m	1 Mbps	2,4 GHz	Sim
WirelessHD	10 m	de 10Mbps até 28 Gbps	60 GHz	Sim
WiFi (IEEE 802.11a)	15 m	54 Mbps	5 GHz	Não
WiFi (IEEE 802.11b)	30 m	11 Mbps	2,4 GHz	Não
WiFi (IEEE 802.11g)	30 m	54 Mbps	2,4 GHz	Não
WiFi (IEEE 802.11n)	50 m	600 Mbps	2,4 GHz e/ou 5 GHz	Sim
WiFi (IEEE 802.11ac)	200 m	500 Mbps	5 GHz	Sim
WiFi (IEEE 802.11ad)	20 m	1 Gbps	60 GHz	Sim
CDMA	5 km	307 kbps	800 MHz	Não
UMTS	5 km	2 Mbps	850 – 2100 MHz	Não
WiMAX (IEEE 806.16)	30 km	9,4 Mbps	2 – 66 GHz	Sim
HSPA+	5 km	42 Mbps	850 – 2100 MHz	Sim
LTE	5 km	100 Mbps	2,5 MHz	Sim

Fonte: (KURAN E TUGCU, 2007; HWANG, 2010; STANKIEWICZ E JAJSZCZYK, 2011; RAYCHAUDHURI E MANDAYAM, 2012).

A maioria dessas redes utiliza o IP para estabelecer uma comunicação entre servidores e clientes. Esse protocolo é baseado na comutação de pacotes através de datagramas. Cada pacote possui o endereço do remetente (servidor) e do destinatário (cliente). Suas principais características são a adoção de serviços sem confirmação, a ausência de mecanismos de controle de erro e de fluxo e o mecanismo do melhor esforço possível para entregar corretamente os datagramas ao destinatário (*best effort delivery*) (CARISSIMI *et al.*, 2009).

Atualmente, o IP (versão 4 – IPv4) tem passado por algumas dificuldades, como falta de endereços IP e problemas no tamanho das tabelas de roteamento. Devido a isso, um novo protocolo tem sido desenvolvido pelo IETF desde a década de 1990. Recentemente, este novo protocolo, denominado IPv6 (versão 6 do IP), tem sido disponibilizado e suas principais características são:

suporte a bilhões de endereços (ao invés de milhões do IPv4), redução do tamanho das tabelas de roteamento, suporte nativo a medição de qualidade, autenticação de usuários e privacidade de dados (CARISSIMI *et al.*, 2009).

A transmissão de dados, sobre uma rede IP, é baseada na quantidade de destinos em uma comunicação entre cliente e servidor. Ela pode ser via *unicast*, *broadcast*, *multicast* ou *peer-to-peer* (P2P) (ver Fig. 5).

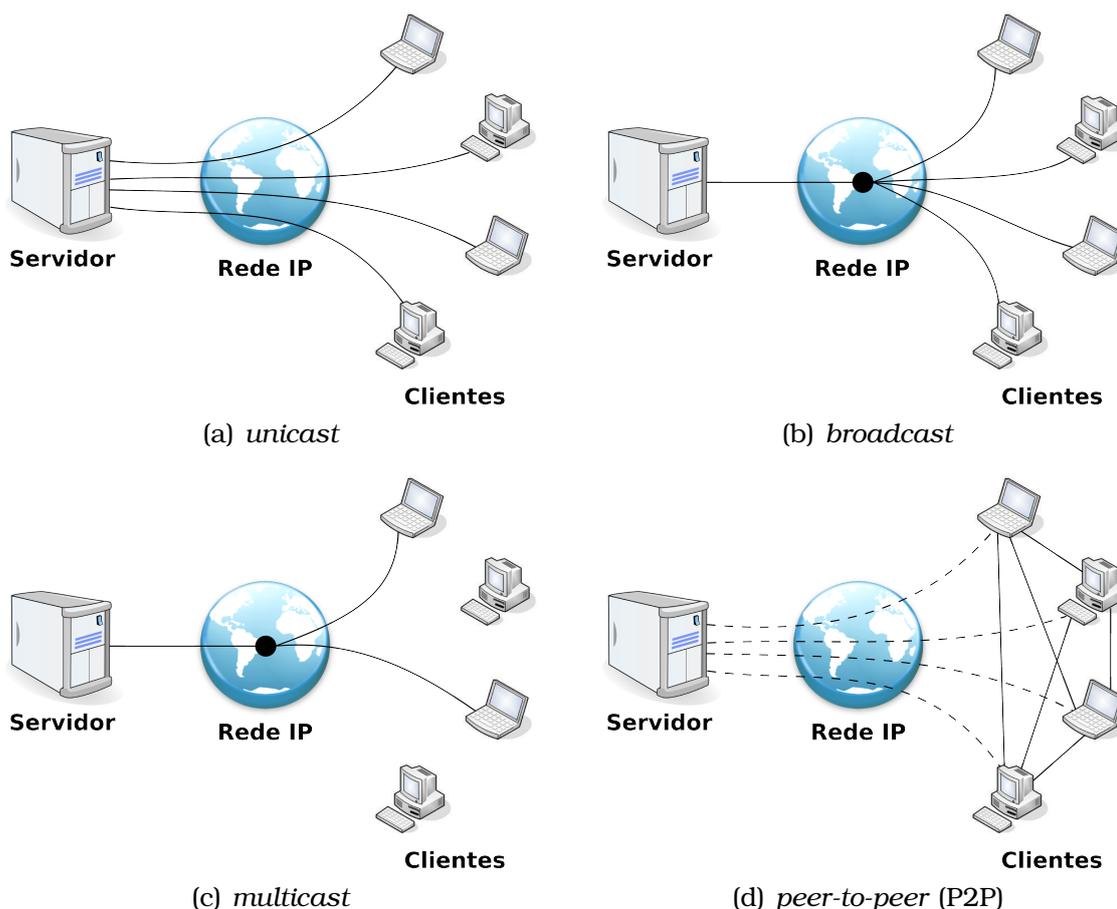


Figura 5: Formas de transmissão sobre uma rede IP.

No *unicast* (Fig. 5 (a)), o conteúdo é disponibilizado a um único destino, ou seja, o conteúdo é enviado ao cliente através de um servidor por meio de uma solicitação (PETERSON E DAVIE, 2004). Caso o mesmo conteúdo seja requisitado por outro cliente, outra solicitação será criada pelo servidor para enviar o conteúdo. Sua principal desvantagem é a sobrecarga na rede com várias cópias de um mesmo conteúdo.

Ao contrário do *unicast*, no *broadcast* (Fig. 5 (b)), o servidor envia o conteúdo para todos os clientes disponíveis na rede ou sub-rede. Nesse caso, o cliente não necessita solicitar o conteúdo. Esse tipo de transporte é utilizado

para disponibilizar um conteúdo para vários clientes de uma vez, ao invés de esperar a solicitação de um cliente para enviar o conteúdo.

Todavia, poderão haver clientes na rede que não gostem de receber o conteúdo. Uma forma de melhorar isso é a utilização do *multicast* (Fig. 5 (c)). O *multicast* permite que o servidor envie o conteúdo para um grupo de clientes através de uma única operação de envio (KUROSE E ROSS, 2003). A vantagem dessa metodologia é a diminuição do tráfego de dados na rede. Por outro lado, atualmente, a Internet não possui muitas facilidades para lidar com redes IP multicast (COMER, 2014).

Outra forma de distribuir conteúdos é através do P2P (Fig. 5 (d)). A ideia por trás do P2P é em vez de ter uma única cópia de um determinado arquivo em um único servidor, o arquivo é dividido em vários pedaços que ficam disponibilizados em vários servidores. Outra característica dessa abordagem é que os próprios clientes são os servidores, não necessitando de máquinas robustas. Dessa forma, o cliente pode acessar o arquivo a partir do servidor mais próximo dele (COMER, 2014).

De acordo com Androutsellis-theotokis e Spinellis (2004), P2P é uma classificação usada tanto para sistemas quanto para aplicações, independente de suas operações internas, baseada na percepção externa de fornecimento de interação direta entre computadores. Segundo Schollmeier (2001), a arquitetura de um sistema distribuído pode ser denominada de P2P se o participante, denominado de *peer*, for capaz de: a) auto-organizar dentro da topologia da rede com o propósito de compartilhar recursos (conteúdo ou poder de processamento); b) adaptar a falhas; c) acomodar transientes de *peers* enquanto mantêm conectividade e desempenho aceitável; e d) ser acessível diretamente por outros *peers* sem passar por entidades intermediárias.

Uma das vantagens do P2P é a diminuição da latência de chegada dos pedaços do conteúdo desejado, pois os pedaços estão localizados em vários *peers* e não em um único local, como em um servidor. Por outro lado, quando um pedaço só existir em um *peer* e esse estiver indisponível, o cliente não obterá por completo o conteúdo desejado.

Para viabilizar a transmissão de conteúdos multimídia sobre redes IP, utilizam-se os protocolos *Real-Time Transport Protocol* (RTP), *Real-Time Control Protocol* (RTCP) e *Real-Time Streaming Protocol* (RTSP).

O RTP foi projetado para trabalhar junto com a camada de aplicação e o protocolo de transporte UDP em aplicações de conteúdo em tempo real, tais como voz e vídeo (SIMPSON, 2008). Ele possui suporte a vários formatos de

áudio e vídeo, suporte ao *multicasting* e numeração de pacotes para ordenação na aplicação. Já o RTCP complementa o RTP, com informações e estatísticas sobre o tráfego de pacotes entre servidor e cliente. Segundo [Peterson e Davie \(2004\)](#), o RTCP possui as seguintes funcionalidades:

- Sincronização entre os pacotes;
- Estatísticas sobre a quantidade de perda de pacotes, o tempo de chegada entre os pacotes, a variação do tempo de chegada dos pacotes, taxa de transmissão dos pacotes, dentre outros;
- Identificação dos pacotes para que possam ser sincronizados no receptor;
- Identificação dos participantes durante uma sessão RTP.

O RTSP trabalha na camada de aplicação e tem como objetivo criar uma sessão de comunicação entre servidor e cliente. Um benefício ao usuário com o seu uso é o controle do conteúdo, como iniciar, pausar, parar, avançar e retroceder. Suas principais características são a utilização do modelo cliente-servidor, o suporte à transmissão dos conteúdos via *unicast* ou *multicast* através dos protocolos TCP ou UDP, e o trabalho em parceria com o RTP e RTCP ([O'DRISCOLL, 2008](#)).

Enquanto o RTP e RTCP garantem a configuração e a conexão entre servidor e cliente, o RTSP disponibiliza uma identificação para o acesso ao conteúdo localizado no servidor e um controle do conteúdo no dispositivo do cliente. A identificação dos conteúdos é dada através de uma *Uniform Resource Locator* (URL), como “rtsp://192.168.0.103:1234/video.mpg”, por exemplo. Uma vez estabelecida a comunicação através dessa URL, o RTP e RTCP entram em ação no gerenciamento da transmissão do conteúdo. Na [Fig. 6](#) é demonstrada uma solicitação de conteúdo ao servidor até o recebimento do mesmo pelo cliente através dos protocolos RTSP, RTP e RTCP.

Uma desvantagem do RTSP está relacionada à largura de banda necessária para transmitir o conteúdo multimídia. Se ela não for suficiente para a recepção do conteúdo após o usuário escolher a qualidade desejada, ocorrerão muitas pausas no decorrer da exibição do conteúdo. Caso o usuário modifique a qualidade do conteúdo, uma nova solicitação ao servidor será feita após saber o instante em que está sendo exibido o conteúdo, para que o usuário não assista desde o começo.

Para contornar esse tipo de problema, uma nova abordagem, denominada *streaming* adaptativo, tem sido utilizada. Nessa abordagem, o servidor

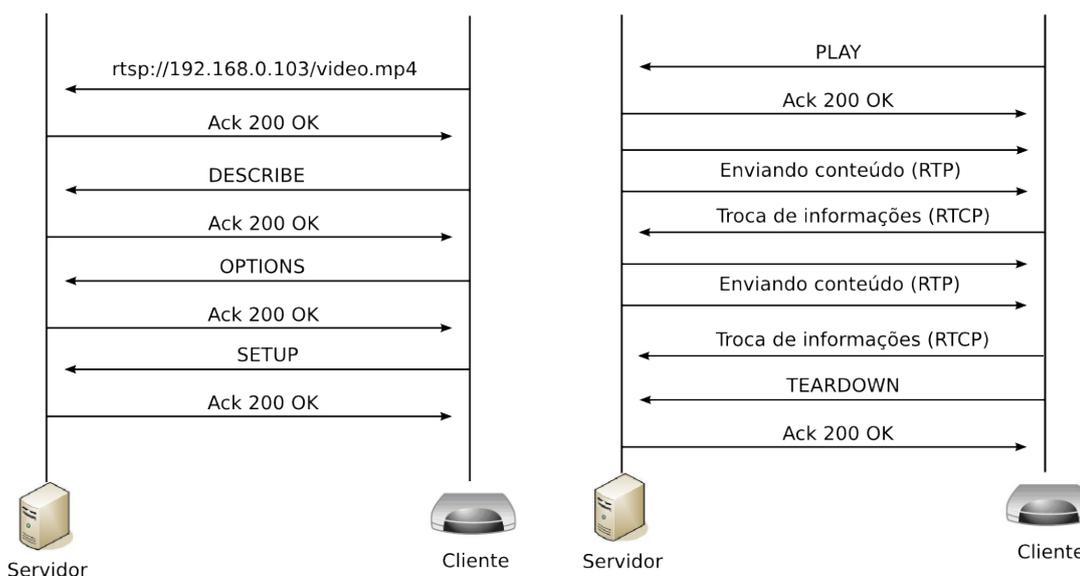


Figura 6: Exemplo de troca de mensagens entre servidor e cliente utilizando os protocolos RTSP, RTP e RTCP.

Fonte: adaptado de O'Driscoll (2008).

de conteúdos contém diferentes níveis de qualidade (resolução e taxa de transmissão) de um mesmo vídeo, divididos em vários fragmentos (em tempos de segundos) (AKHSHABI *et al.*, 2011). Desse modo, o conteúdo exibido no cliente pode ser adaptado automaticamente às condições da comunicação. Por exemplo, se a rede estiver sobrecarregada, a qualidade do vídeo é ajustada para uma qualidade inferior durante a exibição do conteúdo. Caso contrário, a qualidade do vídeo será ajustada automaticamente para uma qualidade superior. Na Fig. 7 é demonstrado o uso do *streaming* adaptativo durante a transmissão de um vídeo entre servidor e cliente.

Algumas soluções proprietárias utilizando o conceito de *streaming* adaptativo têm sido desenvolvidas, como *Smooth Streaming* pela Microsoft, *HTTP Live Streaming* pela Apple, e *Dynamic HTTP Streaming* pela Adobe. Além disso, o MPEG em parceria com o *3rd Generation Partnership Project* (3GPP) propuseram uma especificação aberta e global denominada *Dynamic Adaptive Streaming over HTTP* (DASH) (YITONG *et al.*, 2013).

A principal vantagem do uso do *streaming* adaptativo está relacionada à automatização da mudança de qualidade do conteúdo, para que o usuário não tenha problemas de várias pausas durante a exibição. Apesar disso, no servidor de conteúdos haverá diversos níveis de qualidades para cada fragmento de um determinado vídeo, exigindo uma grande quantidade de espaço para armazenamento dos conteúdos nos servidores.

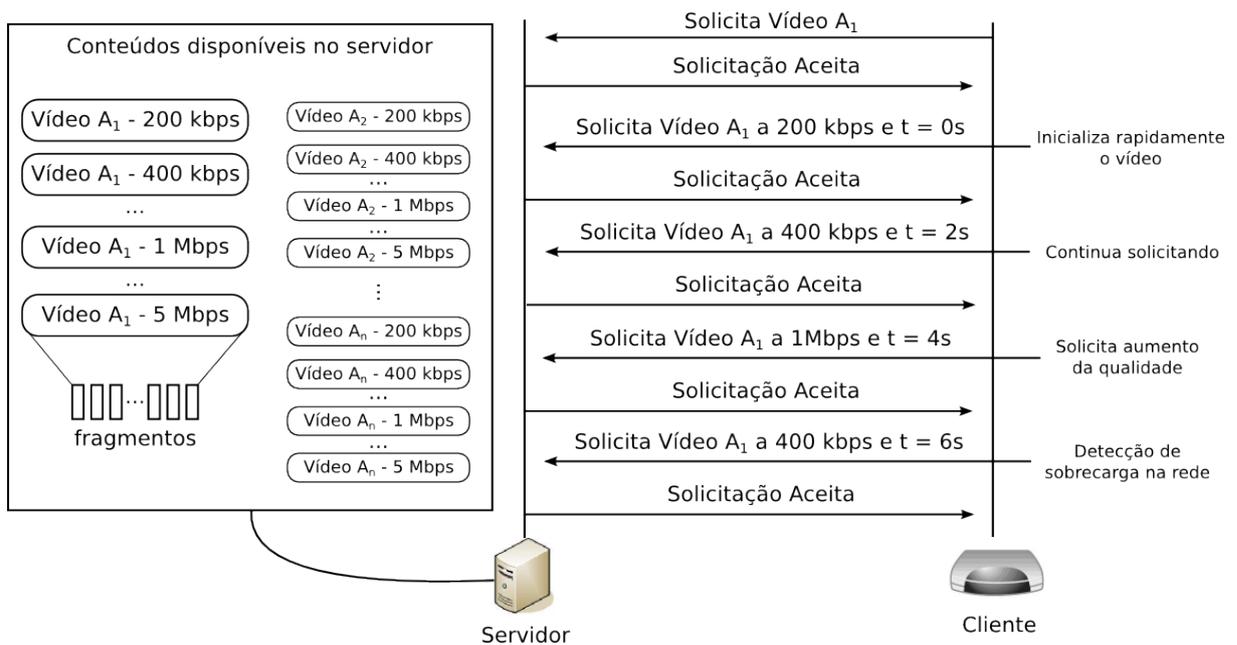


Figura 7: Exemplo de transmissão de conteúdos através do *streaming* adaptativo.

Fonte: adaptado de [Begen et al. \(2011\)](#).

2.3 Recepção de Conteúdos

Com a evolução da tecnologia, principalmente no campo de microprocessamento e armazenamento de dados, muitos dispositivos têm sido criados para ter acesso a conteúdos multimídia disponíveis na Internet, no conforto de uma residência ou fora dela. Dentre esses dispositivos, estão as *smart TVs* e os dispositivos móveis, que são representados em grande maioria por *smartphones* e, recentemente, *tablets* (ver Fig. 8).



Figura 8: Exemplos de *smart TVs*, *smartphones* e *tablets*.

Uma *smart TV* é uma TV que possui acesso à Internet. Ela possui suporte à visualização de conteúdos em formato 3D e podem ser instalados aplicativos, denominados *widgets*, que podem ter acesso a vários serviços dis-

poníveis na Internet, como redes sociais, e-mail, buscas, dentre outros. O tamanho da tela pode variar de 32 polegadas até 85 polegadas e a resolução de 1920x1080 (*full HD*) até 7680x4320 (*ultra HD*). Nesse ramo, as principais empresas são Samsung, Sony, LG, Panasonic, Philips e Toshiba. Na Tabela 3 é mostrada uma comparação entre algumas *smart TVs* encontradas no mercado, levando em consideração tecnologia da tela, tamanho da tela, resolução da tela, entradas disponíveis e conexões.

Tabela 3: Comparação entre algumas *smart TVs* encontradas no mercado.

Modelo (Empresa)	Parâmetros			
	Tecnologia da Tela	Tamanho da Tela	Resolução da Tela	Conexões
42LS5700 (LG)	LED	42"	1920x1080	4 HDMI, 3 USB, Ethernet, WiFi
KDL-48W605 (Sony)	LED	48"	1920x1080	4 HDMI, 2 USB, Ethernet, WiFi
H6400 (Samsung)	LED	65"	1920x1080	4 HDMI, 3 USB, Ethernet, WiFi
VIERA TC-L50E6B (Panasonic)	LED	50"	1920x1080	3 HDMI, 1 USB, Ethernet, WiFi
65PFG7459/78 (Philips)	LED	65"	1920x1080	4 HDMI, 2 USB, Ethernet, WiFi
LE4057i (Semp Toshiba)	LED	40"	1920x1080	3 HDMI, 2 USB, WiFi

Fonte: (LG, 2015; SONY, 2015; SAMSUNG, 2015; PANASONIC, 2015; PHILIPS, 2015; SEMPTOSHIBA, 2015).

Já no ramo dos dispositivos móveis, as principais empresas que estão atuando nesta área são Apple, Samsung e Microsoft. O tamanho das telas pode variar de 4" até 10,1". Diferentemente de suas origens, quando os telefones celulares foram criados para permitir comunicação via voz e sem fio entre duas pessoas, hoje em dia isso é apenas um mero detalhe. As pessoas utilizam os dispositivos móveis para trocar mensagens, acessar a Internet, tirar fotos e compartilhar em redes sociais, etc. No quarto trimestre de 2014, o sistema operacional Android era utilizado em 76,6% dos dispositivos móveis no mundo, seguido do iOS com 19,7% e Windows Phone com 2,8% (IDC, 2015).

Um comparativo de alguns *smartphones* e *tablets* encontrados no mercado é apresentado nas Tabelas 4 – 7. Foram levados em consideração: tamanho da tela; sistema operacional; processador; quantidade de memória RAM, interna e externa; resolução de reprodução de vídeos; duração da bateria; e tipos de conexão disponíveis.

Tabela 4: Comparativo de alguns *smartphones* existentes no mercado (1 de 2).

Características	Smartphones				
	iPhone 5s	iPhone 5c	Galaxy S V	Galaxy Ace 3	Moto G
Fabricante	Apple	Apple	Samsung	Samsung	Motorola
Tela (polegada/pixels)	4" 1136x640	4" 1136x640	4" 800x480	5,1" 1920x1080	4,5" 1280x720
Sistema Operacional	iOS 7	iOS 7	Android 4.2.2	Android 4.2	Android 4.2.2
Processador	dual core, 1,3 GHz	dual core, 1,3 GHz	quad core, 2,5 GHz	dual core, 1 GHz	quad core, 1,2 GHz
Memória RAM	1 GB	1 GB	2 GB	1 GB	1 GB
Memória Interna	16/32/64 GB	8/16/32 GB	16/32 GB	4/8 GB	8/16 GB
Memória Externa	microSD até 128 GB	microSD até 32 GB	microSD até 32 GB
Execução de Vídeo	1080p	1080p	2160p	720p	720p
Conexão WiFi	WLAN a/b/g/n	WLAN a/b/g/n	WLAN a/b/g/n/ac	WLAN b/g/n	WLAN b/g/n
Rede Móvel	3G/4G	3G/4G	3G/4G	3G/4G	3G/4G
Duração da Bateria	10 h	10 h	21 h	8 h	...

Fonte: (GSMARENA, 2015).

Tabela 5: Comparativo de alguns *smartphones* existentes no mercado (2 de 2).

Características	Smartphones			
	Moto E	Xperia T3	Lumia 930	HTC One
Fabricante	Motorola	Sony Ericsson	Nokia	HTC
Tela (polegada/pixels)	4,3" 960x540	5,3" 1280x720	5" 1920x1080	5" 1920x1080
Sistema Operacional	Android 4.2.2	Android 4.2.2	Windows Phone 8.1	Android 4.2.2
Processador	dual core, 1,2 GHz	quad core, 1,4 GHz	quad core, 2,2 GHz	quad core, 2,5 GHz
Memória RAM	1 GB	1 GB	2 GB	2 GB
Memória Interna	4 GB	8 GB	32 GB	16 GB
Memória Externa	microSD até 32 GB	microSD até 32 GB	..	microSD até 128 GB
Execução de Vídeo	854x480	1080p	1080p	1080p
Conexão WiFi	WLAN b/g/n	WLAN a/b/g/n	WLAN a/b/g/n/ac	WLAN a/b/g/n
Rede Móvel	3G	3G/4G	3G/4G	3G/4G
Duração da Bateria	...	13 h	12 h	26 h

Fonte: (GSMARENA, 2015).

Tabela 6: Comparativo de alguns *tablets* existentes no mercado (1 de 2).

Características	Tablets				
	iPad Air	iPad mini 2	Galaxy Tab 3 10,1	Galaxy Tab 3 7,0	Xoom 2
Fabricante	Apple	Apple	Samsung	Samsung	Motorola
Tela (polegada/pixels)	9,7" 2048x1536	7,9" 2048x1536	10,1" 1280x800	7" 1024x600	10,1" 1280x800
Sistema Operacional	iOS 7	iOS 7	Android 4.2.2	Android 4.1.2	Android 4.0.4
Processador	dual core, 1,3 GHz	dual core, 1,3 GHz	dual core, 1,6 GHz	dual core, 1,2 GHz	dual core, 1,2 GHz
Memória RAM	1 GB	1 GB	1 GB	1 GB	1 GB
Memória Interna	16/32/ 64/128 GB	16/32/ 64/128 GB	16/32 GB	8/16 GB	16/32 GB
Memória Externa	microSD até 64 GB	microSD até 64 GB	microSD até 32 GB
Execução de Vídeo	1080p	1080p	720p	720p	720p
Conexão WiFi	WLAN a/b/g/n	WLAN a/b/g/n	WLAN a/b/g/n	WLAN a/b/g/n	WLAN a/b/g/n
Rede Móvel	3G/4G	3G/4G	3G	3G	3G
Duração da Bateria	10 h	10 h	10 h

Fonte: (GSMARENA, 2015).

Tabela 7: Comparativo de alguns *tablets* existentes no mercado (2 de 2).

Características	Tables			
	Xperia Z2	G Pad 10,1	Surface 2	Memo Pad 10
Fabricante	Sony	LG	Microsoft	Asus
Tela (polegada/pixels)	10,1" 1920x1200	10,1" 1920x1200	10,6" 1920x1080	10,1" 1280x800
Sistema Operacional	Android 4.2.2	Android 4.2.2	Microsoft Windows 8.1	Android 4.2
Processador	quad core, 2,3 GHz	quad core	quad core, 1,7 GHz	quad core, 1,6 GHz
Memória RAM	3 GB	...	2 GB	1 GB
Memória Interna	16 GB	...	32/64 GB	8/16 GB
Memória Externa	microSD até 64 GB	microSD até 64 GB	microSD até 64 GB	microSD até 64 GB
Execução de Vídeo	1080p	...	1080p	720p
Conexão WiFi	WLAN a/b/g/n/ac	WLAN a/b/g/n	WLAN a/b/g/n	WLAN b/g/n
Rede Móvel	3G/4G	3G
Duração da Bateria	13 h	9 h

Fonte: (GSMARENA, 2015).

2.4 Conclusões

O respectivo capítulo teve como objetivo introduzir os principais conceitos e elementos que constituem um serviço de *streaming* de vídeos em redes IP. Foram apresentadas as tecnologias utilizadas, desde a aquisição dos conteúdos até a recepção dos mesmos pelos dispositivos dos usuários e ainda foram discutidas algumas técnicas usadas para transmissão dos conteúdos.

Uma questão relevante para o sucesso de um serviço de transmissão de vídeos é a garantia de qualidade percebida pelo cliente. Essa garantia, seja ela avaliada durante a transmissão ou a recepção do conteúdo pelo dispositivo do cliente, é o tema principal do próximo capítulo. Nele, são abordadas as formas de avaliação de qualidade, assim como as técnicas que são utilizadas para incrementar a qualidade em serviço de transmissão de vídeos em redes IP.

AVALIAÇÃO DA QUALIDADE DE VÍDEOS

“Uma grande qualidade ou talento desculpa muitos pequenos defeitos”

Marquês de Maricá

Diversos serviços podem ser encontrados na Internet, tais como enviar e-mails, conversar com alguém através de mensagens de texto ou áudio, assistir vídeos, dentre outros. Para cada um desses serviços, seja ele gratuito ou pago, uma avaliação da qualidade é necessária para garantir o recebimento correto do conteúdo pelo usuário de modo que ele sinta satisfação em usufruí-lo.

O sucesso dos serviços de distribuição de vídeos está fortemente ligado à garantia de qualidade, de modo que os conteúdos sejam exibidos nos dispositivos dos usuários com um mínimo de falhas e atrasos. Para alcançar isso, o administrador da rede monitora informações da infraestrutura (largura de banda, atraso ou vazão) para proporcionar uma qualidade adequada para cada cliente. Todavia, devido à complexidade da rede, torna-se difícil gerenciá-la de modo que a qualidade seja garantida.

Quando dispositivos móveis são inseridos nessa infraestrutura, as dificuldades são ainda maiores, pois surgem novos problemas como, por exemplo, cobertura do sinal sem fio, alta taxa de perda de pacotes e instabilidade do canal sem fio. Além disso, quando se avakua o serviço de vídeo para dispositivos móveis, deve-se observar, além das condições da infraestrutura e conexão, as características de cada dispositivo (tamanho da tela, quantidade de memória interna, capacidade do processador, etc.), para oferecer a qualidade adequada

a cada um deles. Mais do que isso, é importante saber como está sendo exibido o conteúdo no dispositivo do usuário, pois de nada valerá garantir a qualidade da conexão em um nível aceitável, se a qualidade do conteúdo exibido não está sendo satisfatória para o usuário.

Dada às condições necessárias para que um serviço de *streaming* de vídeos sobre a rede IP possa alcançar seus objetivos, algumas informações coletadas na rede IP podem ser consideradas durante a transmissão de um fluxo de vídeo para o usuário. Essas informações são usadas para definir o conceito de qualidade de serviço (*quality of service* – QoS).

Outras informações podem ser mensuradas (resolução exibida no dispositivo ou codificação do conteúdo) para avaliar a qualidade de vídeo. Todos esses fatores podem influenciar na qualidade percebida pelo usuário, determinando, assim, o nível de qualidade de experiência (*quality of experience* – QoE).

3.1 O que é Qualidade de Serviço (QoS)?

A QoS é definida pela ITU como um conjunto de características de um serviço de telecomunicação cujo cerne seja a satisfação do usuário (ITU-T, 2008). A IETF sumariza QoS como uma coleção de requisitos a ser cumprida durante o transporte de informações de um serviço em particular (CRAWLEY, 1998). Os parâmetros comumente utilizados para mensurar a QoS são largura de banda, tempo de atraso entre pacotes, variação do atraso entre os pacotes, índice de perda de pacote e probabilidade de erro entre os pacotes.

Além da QoS, os serviços podem ser avaliados de acordo com o grau de serviço (*grade of service* – GoS) e a qualidade de resiliência (*quality of resilience* – QoR). O GoS está relacionado aos acontecimentos que ocorrem durante a comunicação entre servidor e cliente, como configuração, liberação e manutenção (STANKIEWICZ E JAJSZCZYK, 2011). Nele, são utilizados os seguintes parâmetros: tempo de configuração, probabilidade de que a comunicação esteja bloqueada, tempo para autenticação do cliente e probabilidade de cair a conexão.

Na QoR, o grau de sobrevivência de um conteúdo na rede é verificado, ou seja, quanto tempo é necessário para o serviço se recupere após uma quebra da conexão ou como está a disponibilidade do serviço junto ao servidor (TAPOLCAI *et al.*, 2005). Quando é reestabelecida a conexão após uma falha na comunicação, a QoR é responsável em verificar se o nível de QoS e GoS são

os mesmos antes da quebra da conexão, se o percurso reestabelecido para entregar o conteúdo está congestionado e se houve muitos pacotes perdidos durante uma paralisação de serviço.

Um exemplo de falha durante a transmissão de um vídeo para o dispositivo do usuário é ilustrado na Fig. 9. Em algum instante da exibição do conteúdo (Fig. 9 (b)), algumas regiões da imagem não estão de acordo com o que foi transmitido pelo servidor (Fig. 9 (a)). Isso pode ter acontecido por vários fatores como, por exemplo, atraso dos pacotes devido ao congestionamento da rede ou perda de pacotes devido a cobertura do sinal sem fio.

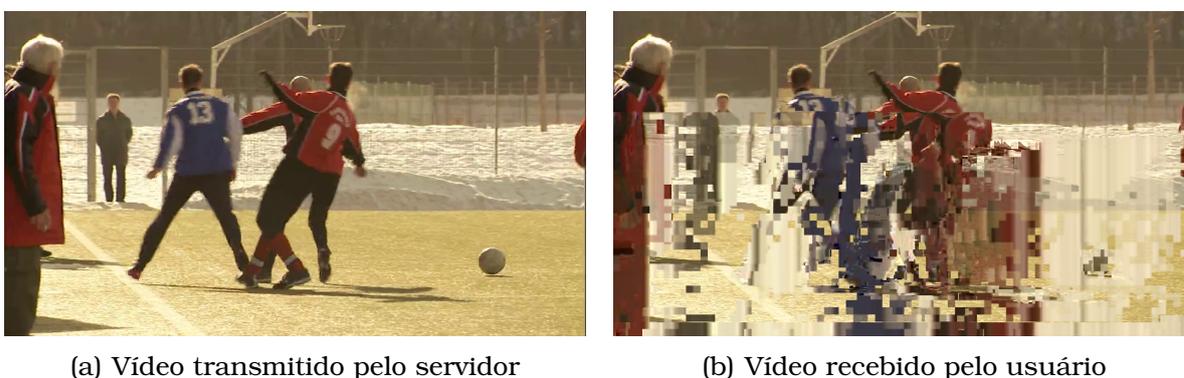


Figura 9: Exemplo de falha durante a transmissão de um vídeo.

Em um serviço de *streaming* de vídeos, a mensuração da QoS ocorre no instante em que os pacotes gerados pelo servidor para cada vídeo são transmitidos para o dispositivo do usuário. Esse tipo de avaliação é chamada de QoS da rede (*network QoS* – NQoS). Por outro lado, pode-se investigar o relacionamento entre os parâmetros da QoS com a qualidade percebida pelo usuário (*perceived QoS* – PQoS). Com o passar dos anos, esse termo evoluiu para qualidade de experiência (*quality of experience* – QoE), onde a experiência do usuário é mais valiosa do que a qualidade provida pelo serviço.

3.2 O que é Qualidade de Experiência (QoE)?

A QoE é uma avaliação da satisfação do usuário com relação ao conteúdo exibido em seu dispositivo (ZEPERNICK E ENGELKE, 2011). Para Raake e Egger (2014), a QoE é o grau de prazer ou aborrecimento de uma pessoa sobre uma aplicação, um serviço ou um sistema. Tudo isso está relacionado à percepção da pessoa com relação ao conteúdo mostrado no dispositivo.

Segundo [Sternberg \(2010\)](#), a percepção compreende processos em que é possível reconhecer, organizar e entender as sensações provenientes de estímulos ambientais. O processo de percepção da QoE é fundamentado no sistema audiovisual humano, o qual relaciona a percepção auditiva (volume de áudio, momentos de silêncio ou distorções de áudio) ou visual (mudança de cor, variação da intensidade de luz ou falha em algum pixel) do usuário sobre um conteúdo exibido a ele em algum dispositivo. O presente trabalho salientará somente no componente visual da QoE. Na Fig. 10 é mostrado um modelo conceitual do processo de formação da opinião sobre qualidade.

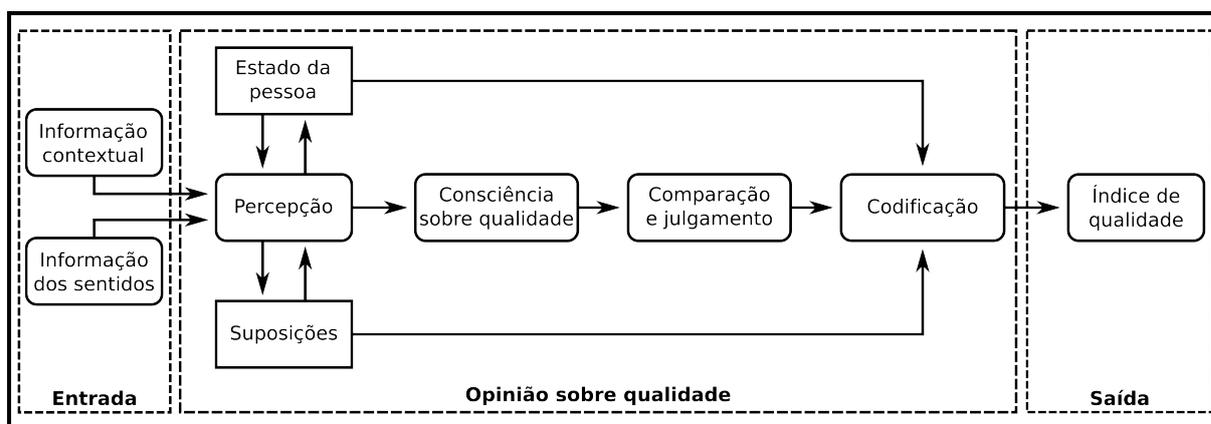


Figura 10: Modelo conceitual do processo de formação da opinião sobre qualidade.

Fonte: adaptado de [Raake e Egger \(2014\)](#).

O processo de percepção inicia nos estímulos originados do ambiente onde a pessoa está inserida (informações contextuais) ou dos seus sentidos (o que é visualizado no dispositivo ou o que é ouvido ao seu redor). Esses estímulos são convertidos em representações neurais, que formam a percepção em conjunto com o estado da pessoa (estado mental e físico) e com as suposições (atitudes e conceitos). Consecutivamente, a consciência sobre a qualidade entra em ação, onde a pessoa fica concentrada em algum tipo de avaliação de qualidade. O resultado disso é a comparação com casos anteriores para, enfim, julgar de forma positiva ou negativa a qualidade. Finalmente, com base nesse julgamento mais as condições de seu estado e suas suposições, a pessoa opina sobre a qualidade ([RAAKE E EGGER, 2014](#)).

Durante a avaliação da QoE, os parâmetros utilizados são subjetivos, ou seja, a percepção do usuário com relação ao conteúdo exibido no dispositivo é mensurada por palavras, tais como excelente, bom, aceitável, ruim ou péssimo. Além disso, outros parâmetros podem ser levados em consideração, tais como custo, disponibilidade, usabilidade e fidelidade ([KIM E CHOI, 2010](#)).

Quando comparada com a QoS, a QoE não possui métricas bem definidas, pois a sua avaliação depende da percepção de cada usuário. Por outro lado, existe uma relação entre QoS e QoE. Na Fig. 11 é ilustrada uma plausível relação entre QoS e QoE por meio de três regiões.

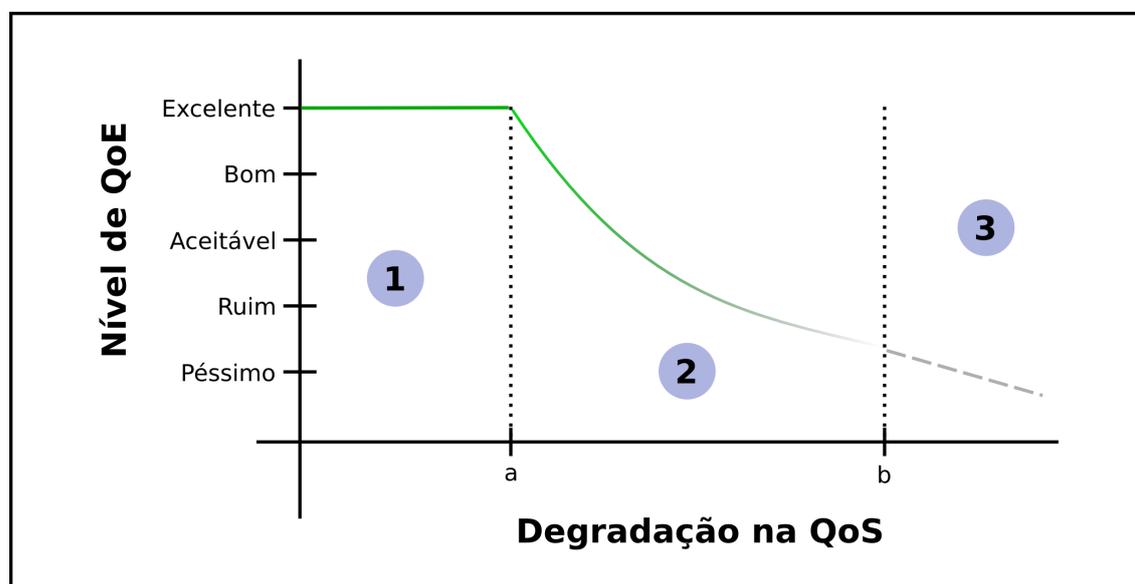


Figura 11: Uma plausível relação entre QoS e QoE.

Fonte: adaptado de [Fiedler et al. \(2010\)](#).

Na Região 1, não existe algum distúrbio perceptível durante a transmissão do vídeo a partir do servidor para o receptor, proporcionando, assim, uma avaliação excelente da QoE. Na Região 2, surgem algumas falhas durante a transmissão (perda de pacote ou atraso) e a satisfação do usuário com relação ao conteúdo diminui. Finalmente, na Região 3, o usuário desiste de assistir ao conteúdo porque está ocorrendo muita degradação na QoS, proporcionando, assim, uma QoE inaceitável.

Existem outros fatores que podem influenciar na avaliação da QoE, como apresentado na Fig. 12. [Baraković et al. \(2010\)](#) classificaram esses fatores em cinco categorias: (1) desempenho tecnológico; (2) usabilidade; (3) avaliação subjetiva; (4) expectativas; e (5) contexto. No trabalho de [Xue e Chen \(2012\)](#) foram levados em consideração alguns fatores (tamanho da tela, distância entre usuário e dispositivo, iluminação e movimentação do usuário) que podem influenciar na percepção dos usuários quando assistiam vídeos em dispositivos móveis.

O *European Network on Quality of Experience in Multimedia Systems and Services* (Qualinet) definiu três principais grupos de influência sobre a

QoE: (1) fatores humanos, tais como demografia, informações sócio-econômica, estrutura mental e psicológica e estado emocional; (2) fatores do sistema como, por exemplo, captura, transmissão e exibição no dispositivo; e (3) fatores de contexto, tais como o ambiente físico onde está localizado o usuário, fatores social, temporal e econômica (CALLET, 2013).

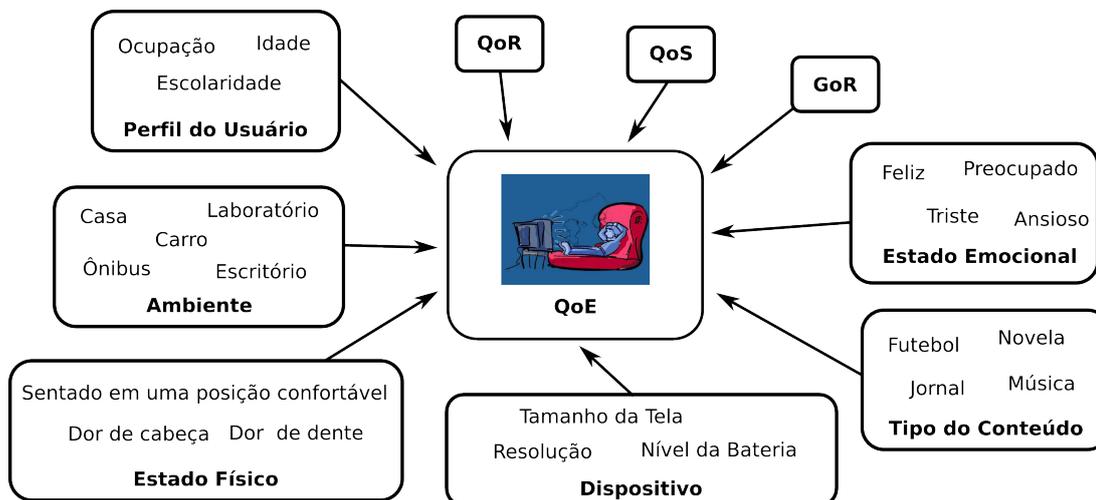


Figura 12: Fatores que podem influenciar na avaliação da QoE.

Fonte: adaptado de De Moor *et al.* (2010) e Stankiewicz e Jajszczyk (2011).

3.3 Metodologias Utilizadas na Avaliação da Qualidade de Vídeos

As metodologias usadas para avaliação da qualidade de vídeos são normalmente divididas em quatro etapas, conforme ilustrado na Fig. 13. São elas: (1) escolha dos vídeos de referência (sem introdução de erro); (2) criação dos vídeos distorcidos (com introdução de erro); avaliação dos vídeos distorcidos por pessoas; e (4) análise estatística das avaliações dos vídeos distorcidos.



Figura 13: Diagrama para avaliação da qualidade de vídeos.

Os vídeos de referência, também conhecidos como dados não comprimidos, não estão codificados com algum algoritmo de compressão e não

possuem falhas. As resoluções espaciais comumente utilizadas vão desde CIF (352x288) até alta definição (1280x720 e 1920x1080). Atualmente, também estão disponibilizadas bases de vídeos com resolução 4K (3840x2160) e formato em 3D. Diferentes grupos de pesquisa dessa área criam suas próprias bases de vídeos de referência e os disponibilizam em seus websites (ver Tabela 8).

Tabela 8: Lista de bases de vídeos de referência disponível publicamente.

Nome da Base de Vídeos	Instituição	Resolução Espacial (pixels)	Frame rate (fps)	Quantidade de Vídeos de Referência
EPFL-PoliMI Video Quality Assessment Database (DE SIMONE <i>et al.</i> , 2009, 2010; EPFL-POLI MI, 2015)	Politecnico di Milano (PoliMI) and École polytechnique fédérale de Lausanne (EPFL)	352x288, 704x576	25, 30	12
IRCCyN/IVC 1080i Database (PÉCHARD <i>et al.</i> , 2008; IRCCyN/IVC, 2015A)	Institut de recherche en communication et cybernétique de Nantes (IRCCyN)	1920x1080	50	24
IRCCyN/IVC H264 AVC vs SVC VGA Video database (PITREY <i>et al.</i> , 2010; IRCCyN/IVC, 2015B)		320x240, 640x480	30	8
IRCCyN/IVC Influence Content Video VGA database (PITREY <i>et al.</i> , 2012; IRCCyN/IVC, 2015C)		640x480	30	60
IT-IST Lisbon H.264/MPEG-2 Video Quality Database (BRANDÃO E QUELUZ, 2010A; BRANDÃO <i>et al.</i> , 2010B; BRANDÃO E QUELUZ, 2010C; IT-IST, 2015)	IT-IST Lisbon	352x288	25, 30	12
Video Trace Library (SEELING E REISSLEIN, 2012; PULIPAKA <i>et al.</i> , 2013; TRACE, 2015)	Arizona State University (ASU)	176x144, 352x288, 1920x1080	25, 30	26
TUM 1080p25 and 1080p50 Data Set (KEIMEL <i>et al.</i> , 2010; TUM, 2015A)	Technische Universität München (TUM)	1920x1080	25, 50	9
TUM Multi Format Test Set (KEIMEL <i>et al.</i> , 2012; TUM, 2015B)		720x576, 1280x720, 1920x1080	24, 25, 50, 60	48

Continua na próxima página

Continuação da Tabela 8

Nome da Base de Vídeos	Instituição	Resolução Espacial (pixels)	Frame rate (fps)	Quantidade de Vídeos de Referência
LIVE Video Quality Database (SESHADRINATHAN <i>et al.</i> , 2010A,B; LIVE, 2015A)	Laboratory for Image & Video Engineering (LIVE) – University of Texas	768x432	25, 50	10
LIVE Mobile Video Quality Database (MOORTHY <i>et al.</i> , 2012A,B; LIVE, 2015B)		1280x720	30	10
4K-HEVC Video Subjective Quality Database (SONG <i>et al.</i> , 2013; 4KHEVC, 2015)	Shanghai Jiao Tong University	3840x2160	30	15
Ultra High Definition HEVC DASH Data Set (LE FEUVRE <i>et al.</i> , 2014; TSI, 2015)	Telecom ParisTech	1280x720, 1920x1080, 3840x2160	30, 60	20
3D Video Quality Assessment Database (GOLDMANN <i>et al.</i> , 2010; EPFL, 2015)	Multimedia Signal Processing Group (MMSPG) – École polytechnique fédérale de Lausanne (EPFL)	1920x1080	25	6

Um dos critérios utilizados para selecionar os vídeos de referência é a análise das informações espacial e temporal. Na informação espacial (*spatial information* – SI) são avaliados os detalhes espaciais do n -ésimo quadro F do vídeo por intermédio da aplicação do filtro de Sobel ($Sobel()$). Essa informação é obtida a partir da equação 1. Quando o valor resultante é alto, diz-se que o vídeo possui alta complexidade para codificação (ITU-T, 2009).

$$SI = \max_{time} \{std_{space}[Sobel(F_n)]\} \quad (1)$$

A informação temporal (*temporal information* – TI) é calculada a partir da diferença entre os pixels na posição i (linha) e j (coluna) do quadro atual (F_n) com os pixels na mesma posição do seu quadro antecessor (F_{n-1}). Essa informação é representada pela equação 2. O vídeo que possui muitas cenas em movimento terá como resultado um valor alto (ITU-T, 2009).

$$TI = \max_{time} \{std_{space}[F_n(i, j) - F_{n-1}(i, j)]\} \quad (2)$$

Na próxima etapa, os vídeos distorcidos são criados. Esses vídeos contêm artefatos de erros ou variação na qualidade durante a codificação do conteúdo. Os erros comumente introduzidos são o *blockiness*, *blurring*, *flickering*,

perda de pacote e *jitter* (FARIAS, 2010). A taxa de bits por segundo (*bitrate*) e de frames por segundo (*frame rate*) podem ser ajustadas durante o processo de codificação. Na Fig. 14 são expostos alguns *snapshots* de artefatos de erro introduzidos durante a criação dos vídeos distorcidos.



Figura 14: Exemplos de artefatos de erro durante a criação dos vídeos distorcidos.

Posteriormente, a qualidade dos vídeos distorcidos são avaliados por um grupo de pessoas. O número de participantes pode variar de 4 até 40 (ITU-T, 2009), mas pelo menos 15 participantes são necessários para obter uma avaliação estatística significativa sobre a qualidade dos vídeos distorcidos (ITU-T, 2002). Estudos apresentados em Staelens *et al.* (2010) mostraram que a duração máxima de cada sessão para avaliação da qualidade dos vídeos não exceda 30 minutos, para que o avaliador não tenha fadiga. Além disso, antes do início da sessão de avaliação, recomenda-se que cada participante seja examinado quanto à acuidade visual (teste de Snellen) e à deficiência de cores (teste de Ishihara) (ITU-T, 2007).

O teste de acuidade visual tem por objetivo avaliar a limitação da visão em identificar detalhes espaciais (forma e contorno) de objetos (MINISTÉRIO DA SAÚDE, 2008). Para a realização desse teste, adota-se o uso da Tabela de Snellen (ver Fig. 15). Essa tabela é constituída por letras, organizadas de

maneira padronizada, de tamanhos progressivamente menores, chamados de optotipos. Em cada linha, existe um número decimal que corresponde à medida de acuidade visual da pessoa. A pessoa apresenta visão normal quando, ao ser colocada em frente à tabela de Snellen, consegue ler as menores letras que nela se encontram (acuidade visual 20/20). Quando a pessoa não enxerga uma ou mais letras (menor que acuidade visual 20/20), ela apresenta alguma limitação da visão.

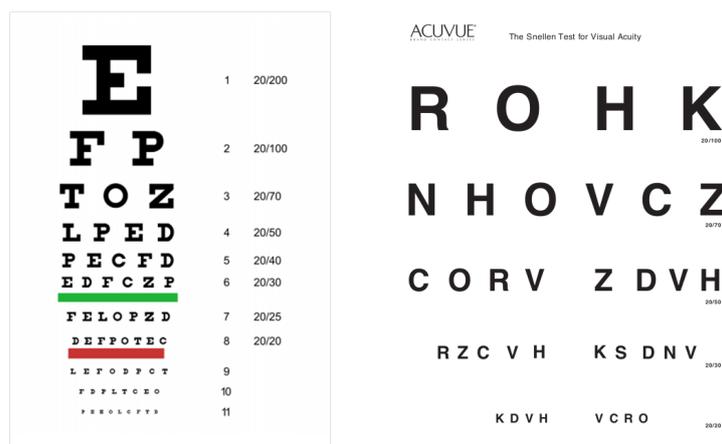


Figura 15: Exemplos da tabela de Snellen.

Fonte: [Acuvue \(2015\)](#) e [Provisu \(2015\)](#).

O teste de cores visa averiguar se uma pessoa possui alguma deficiência visual de origem congênita para cores, também conhecida como daltonismo. Para isso, utiliza-se o teste de Ishihara ([ISHIHARA, 1972](#)). Ele é constituído de placas onde estão expostos vários círculos com cores ligeiramente diferentes. No centro dessas placas, estão presentes alguns números ou algumas linhas (ver Fig. 16). Em um exame simplificado, escolhem-se seis placas aleatoriamente. A visão da pessoa será considerada normal para cores se ela conseguir identificar os números e as linhas mostrados nas seis placas. Caso contrário, ela possui alguma deficiência visual de cor.

Ao final desse processo, elabora-se um modelo de predição para avaliar a qualidade dos vídeos distorcidos, o qual pode ser baseado no vídeo original ou na opinião dos usuários com relação a qualidade dos vídeos distorcidos. Além disso, outros parâmetros podem ser coletados a partir da camada de aplicação (tamanho e resolução do dispositivo), camada de rede (taxa de perda de pacotes, atraso e *jitter*), e camada de conteúdo (*frame rate*, *bitrate*, resolução espacial e informação temporal).

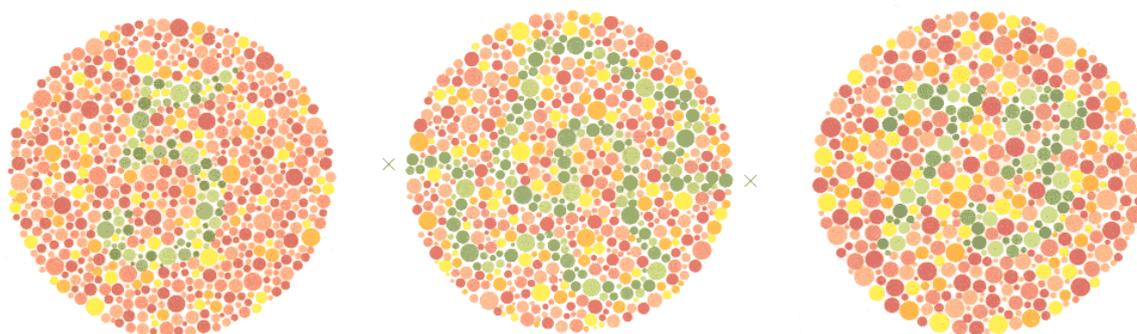


Figura 16: Exemplos de placas utilizadas no teste de Ishihara.

Fonte: Ishihara (1972).

Para melhorar o desempenho do preditor, podem-se utilizar técnicas para reduzir a quantidade de atributos de entrada, tais como análise de componentes principais (*principal component analysis* – PCA) (YANG *et al.*, 2011), análise de discriminante canônico (*canonical discriminant analysis* – CDA) (SHMUELI *et al.*, 2008) ou teoria dos conjuntos aproximativos (*rough set theory* – RST) (LAGHARI *et al.*, 2012). Os métodos comumente empregados para a criação do preditor de qualidade de vídeos são baseados em regressão, redes neurais artificiais (*artificial neural networks* – ANN), e lógica *fuzzy* (MSAKNI E YOUSSEF, 2013).

Para avaliar a acurácia do preditor, são utilizadas as seguintes métricas: erro quadrático médio (*mean squared error* – MSE), distância absoluta média (*mean absolute error* – MAE), coeficiente de correlação de Pearson (*Pearson correlation coefficient* – PCC), e coeficiente de correlação de postos de Spearman (*Spearman rank order correlation coefficient* – SROCC) (KORHONEN *et al.*, 2012). Complementando essa avaliação, podem-se ser empregados métodos de amostragem, tais como *holdout*, amostragem aleatória, validação cruzada e *bootstrap* (TAN *et al.*, 2009; FACELI *et al.*, 2012).

Considerando o processo de avaliação da qualidade de vídeos mostrado na Fig. 13, pode-se classificar em três principais abordagens: subjetiva, objetiva e híbrida.

3.3.1 Abordagem Subjetiva

A abordagem subjetiva produz uma avaliação da qualidade do vídeo baseada na percepção humana dos avaliadores. Usualmente, os experimentos são realizados em laboratórios, onde os vídeos são exibidos em uma tele-

visão. Entretanto, outros lugares podem ser utilizados, como em ambientes naturais (casa ou trabalho), bem como outros meios para exibição dos vídeos (dispositivos móveis) (KORHONEN *et al.*, 2012; CATELLIER *et al.*, 2012).

O índice de opinião média (*mean opinion score* – MOS) é uma métrica comumente utilizada nas avaliações subjetivas. Aqui, o participante avalia a qualidade de cada vídeo por meio de um valor contido na escala de qualidade, a qual varia de 1 até 5 (ver Tabela 9). Outra opção é a escala de degradação (ver Tabela 10), onde o participante relata o nível de degradação do vídeo distorcido, comparando-o com o vídeo de referência exibido anteriormente.

A forma como os vídeos são exibidos aos participantes pode ser dividido em estímulo simples (*single stimulus* – SS), estímulo duplo (*double stimulus* – DS), e estímulo por comparação (*comparison stimulus* – CS) (ITU-T, 1999).

Tabela 9: Escala de qualidade.

Esca	Qualidade
5	Excelente
4	Bom
3	Aceitável
2	Ruim
1	Péssimo

Fonte: (BROOKS E HESTNES, 2010).

Tabela 10: Escala de degradação.

Esca	Prejuízo
5	Imperceptível
4	Perceptível, mas não irritante
3	Ligeiramente irritante
2	Irritante
1	Muito irritante

Fonte: (ITU-T, 1999).

Estímulo simples (single stimulus – SS)

No estímulo simples, os vídeos distorcidos são mostrados um a um e avaliados individualmente por cada participante. Às vezes, os vídeos de referência são exibidos, não informando ao avaliador. Os principais métodos dessa categoria são: *absolute category rating* (ACR), *absolute category rating with hidden reference* (ACR-HR), e *single stimulus continuous quality evaluation* (SSCQE) (ITU-T, 1999).

No ACR, somente vídeos distorcidos são exibidos para o participante e eles são avaliados individualmente por meio da escala ilustrada na Tabela 9. Esses vídeos podem ser visualizados mais de uma vez, desde que as condições definidas previamente sejam as mesmas. Cada vídeo tem, em média, duração de dez segundos.

O ACR-HR é similar ao ACR, com a diferença que os vídeos de referência são utilizados. No entanto, o participante não sabe qual é o tipo do vídeo que ele está avaliando. A escala de qualidade é usada para a avaliação dos vídeos (ver Tabela 9), mas o valor final da qualidade de cada vídeo distorcido será calculado a partir da diferença do valor da qualidade (*differential mean opinion score – DMOS*) entre o vídeo distorcido e seu respectivo vídeo de referência. Esse processo é conhecido como “referência escondida” (*hidden reference*).

O SSCQE possibilita avaliar os vídeos distorcidos de forma contínua, realizada enquanto o participante assiste aos vídeos. Cada vídeo é mostrado uma única vez, com duração mínima de cinco minutos. A escala utilizada é similar a escala de qualidade, com a diferença que ela é contínua e os valores são normalizados para um intervalo de 0 até 100 (ver Tabela 11).

Tabela 11: Escala de qualidade contínua.

Intervalo	Qualidade
80 a 100	Excelente
60 a 80	Bom
40 a 60	Aceitável
20 a 40	Ruim
0 a 20	Péssimo

Fonte: (ITU-T, 1999).

Estímulo duplo (double stimulus – DS)

No estímulo duplo, os vídeos tanto de referência quanto os distorcidos são mostrados ao avaliador. Para avaliar a qualidade dos vídeos distorcidos, é solicitado ao participante que leve em consideração a diferença de qualidade com o seu respectivo vídeo de referência. Os principais métodos classificados como estímulo duplo são: *double stimulus impairment scale* (DSIS), *double stimulus continuous quality scale* (DSCQS), e *simultaneous double stimulus for continuous evaluation* (SDSCE) (ITU-T, 1999).

No DSIS, vários pares de vídeos são apresentados em sequência para o participante, onde o primeiro é o vídeo de referência enquanto que o segundo é o vídeo distorcido. Cada sequência pode ser reproduzida até duas vezes. Após isso, o participante avalia a qualidade do vídeo distorcido por intermédio da escala de degradação (ver Tabela 10).

Similar ao método anterior, no DSCQS cada participante avalia a qualidade de ambos os vídeos, mas o participante não sabe qual é o tipo do vídeo (referência ou distorcido) que está sendo avaliado. A escala utilizada nesse método é o mesmo empregado no SSCQE (ver Tabela 11).

No SDSCE, um par de vídeos é mostrado simultaneamente e o vídeo de referência conhecido pelo participante. A avaliação da qualidade do vídeo distorcido é realizada de forma contínua, através de uma escala padronizada (ver Tabela 11), após observar a diferença de qualidade com o vídeo de referência. A duração de cada vídeo é de, aproximadamente, cinco minutos (ITU-T, 2002).

Estímulo por comparação (comparison stimulus – CS)

O estímulo por comparação é similar ao estímulo duplo, com a diferença que somente os vídeos distorcidos são exibidos aos participantes. A avaliação da qualidade pode ser feita de duas maneiras. Na primeira, chamada de *pair comparison* (PC), é perguntado ao participante que indique quais dos vídeos possui a melhor qualidade, sem a necessidade de informar o valor da qualidade. Na segunda, conhecida como *stimulus comparison adjectival categorical judgment* (SCACJ), o participante indica, através de uma escala de valores mostrada na Tabela 12, a qualidade do segundo vídeo após compará-lo com o primeiro vídeo.

Tabela 12: Escala de comparação utilizada no SCACJ.

Escola	Comparação
-3	Muito pior
-2	Pior
-1	Um pouco pior
0	O mesmo
1	Um pouco melhor
2	Melhor
3	Muito melhor

Fonte: (ITU-T, 1999).

Na Tabela 13 é provida uma análise comparativa das metodologias subjetivas para avaliação da qualidade de vídeos discutidas nessa sessão. Os métodos são comparados um a um, usando os seguintes parâmetros: tipo de estímulo, uso explícito ou escondido dos vídeos de referência, duração dos vídeos, repetição dos vídeos, escala de avaliação da qualidade utilizada, e avaliação contínua.

Tabela 13: Análise comparativa dos métodos utilizados na abordagem subjetiva.

Parâmetros	Métodos							
	ACR	ACR-HR	SSCQE	DSIS	DSCQS	SDSCE	PC	SCACJ
Tipo de estímulo	SS	SS	SS	DS	DS	DS	CS	CS
Vídeo de referência explícito	Não	Não	Não	Sim	Não	Sim	Não	Não
Vídeo de referência escondido	Não	Sim	Não	Não	Sim	Não	Não	Não
Duração dos vídeos	10 seg	10 seg	5 min	10 seg	10 seg	5 min
Repetição da exibição do vídeo	Sim	Sim	Não	Sim	Sim	Não	Sim	Sim
Escala de qualidade utilizada	Tab. 9	Tab. 9	Tab. 11	Tab. 10	Tab. 11	Tab. 11	..	Tab. 12
Avaliação contínua	Não	Não	Sim	Sim	Sim	Sim	Não	Não

3.3.2 Abordagem Objetiva

A principal dificuldade na abordagem subjetiva é encontrar pessoas que estejam dispostas a avaliar a qualidade dos vídeos. Além disso, a per-

cepção de cada pessoa pode ser influenciada por fatores externos, tais como ambiente (em casa, no trabalho ou na rua), estado emocional e físico, tipo do conteúdo exibido (música, jornal ou esporte), perfil da pessoa (escolaridade, idade ou ocupação) e assim por diante (STANKIEWICZ E JAJSZCZYK, 2011).

Na abordagem objetiva, criam-se modelos matemáticos (algoritmos) baseados no sistema visual humano (*human visual system* – HVS), a fim de avaliar a qualidade dos vídeos sem interferências externas. Esses modelos são classificados em referência completa (*full-reference* – FR), referência reduzida (*reduced-reference* – RR), e sem referência (*no-reference* – NR) (HEMAMI E REIBMAN, 2010; SESHADRINATHAN E BOVIK, 2011).

Referência completa (*full-reference* – FR)

Na referência completa, o vídeo distorcido é comparado quadro-por-quadro com seu respectivo vídeo de referência para obter uma medida de qualidade (ver Fig. 17). Diferentes aspectos são considerados, tais como processamento da cor, informação espacial e temporal, percepção a intensidade de luz, e sensibilidade ao contraste (YOU *et al.*, 2010). As métricas mais representativas desse modelo são o erro quadrático médio (*mean squared error* – MSE), *structural similarity* (SSIM), *visual signal-to-noise ratio* (VSNR), *information fidelity criterion* (IFC), *visual information fidelity* (VIF), e *peak SNR* (PSNR).

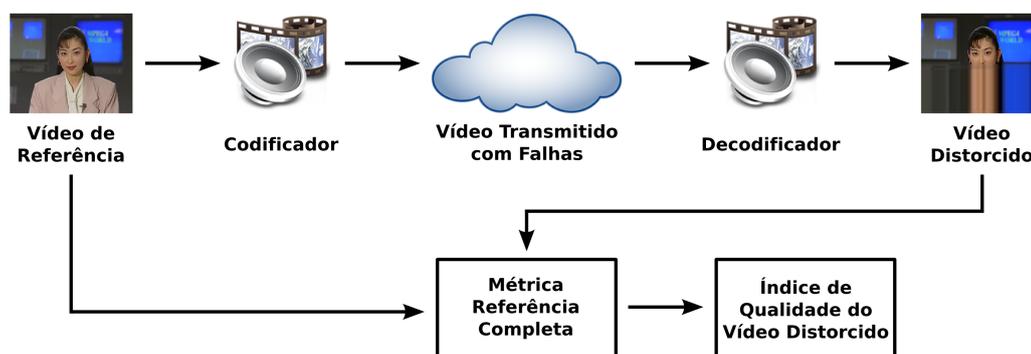


Figura 17: Diagrama utilizado pelas métricas classificadas como referência completa para avaliar a qualidade do vídeo distorcido.

O MSE é um estimador estatístico usado para mensurar a diferença aritmética entre o conteúdo de referência e o conteúdo distorcido (SHEIKH E BOVIK, 2006). Ele avalia o valor de distorção de todos os pixels e não leva em consideração os aspectos visuais da percepção humana. A qualidade do

vídeo distorcido é considerada boa quando o MSE retorna um valor baixo, após comparação com a variância do sinal analisado.

O SSIM avalia a variação dos padrões de intensidade de um pixel local (WANG *et al.*, 2004) e está relacionado ao contraste do pixel ou as mudanças de luminância. A qualidade do vídeo distorcido está associada com três aspectos de perda de informação: distorção da correlação, distorção do contraste e distorção da luminância. Quando o valor da qualidade mensurada estiver próximo de um, a qualidade do conteúdo distorcido será considerada excelente.

O VSNR é um método inspirado no modelo psicoacústico do ser humano e sua abordagem é dividida em dois estágios. No primeiro estágio, detectam-se as distorções usando limiares de contraste. Se as distorções estiverem abaixo da percepção humana, o conteúdo distorcido está adequado (VNSR = infinito) e não é mais necessário continuar com a análise. Entretanto, se as distorções são visíveis (segundo estágio), um valor é obtido a partir do contraste percebido (CHANDLER E HEMAMI, 2007). As limitações desse método estão relacionadas ao uso específico do componente de luminância para obter o índice e o não uso de informações espaciais do conteúdo distorcido.

O IFC utiliza modelos de ambientes naturais em conjunto com modelos de distorção para quantificar estatisticamente informações compartilhadas entre os conteúdos de referência e distorcidos, produzindo um valor de qualidade que varia de zero (sem fidelidade) a infinito (fidelidade perfeita) (SHEIKH *et al.*, 2005). Ele é baseado em propriedades estatísticas de ambientes naturais relacionadas à evolução do sistema visual humano (SHEIKH E BOVIK, 2006).

No VIF, o conteúdo de referência é modelado como sendo a saída de uma fonte natural estocástica que passa por um “canal HSV” e que, posteriormente, será processada pelo cérebro (SHEIKH E BOVIK, 2006). Ele calcula a informação mútua entre a entrada e a saída do “canal HSV” e então insere um canal de distorção entre elas. O valor resultante é mensurado a partir da informação que o cérebro poderia extrair a partir do conteúdo distorcido.

Finalmente, o PSNR é uma das métricas mais utilizadas na referência completa. Ele realiza uma comparação pixel-a-pixel entre o vídeo de referência e o vídeo distorcido, não levando em consideração o que eles realmente representam (WINKLER E MOHANDAS, 2008). O PSNR pode ser aproximadamente mapeado para o MOS utilizando a Tabela 14.

Tabela 14: Conversão do PSNR para o MOS.

PSNR (dB)	MOS
> 37	5 (Excelente)
31 – 37	4 (Bom)
25 – 31	3 (Aceitável)
20 – 25	2 (Ruim)
< 20	1 (Péssimo)

Fonte: (GROSS *et al.*, 2004).

Referência reduzida (reduced-reference – RF)

Na referência reduzida, algumas propriedades ou características são extraídas a partir do vídeo de referência (detalhes espacial ou informação temporal) e elas são utilizadas na avaliação da qualidade do vídeo distorcido (ver Fig. 18).

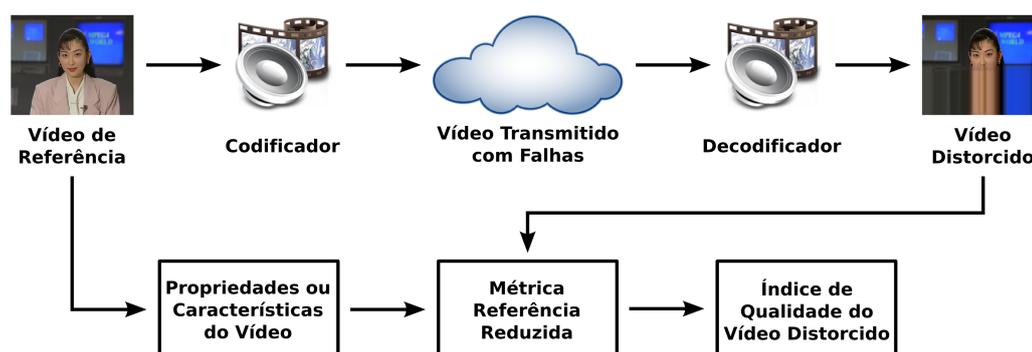


Figura 18: Diagrama utilizado pelas métricas classificadas como referência reduzida para avaliar a qualidade do vídeo distorcido.

Um dos primeiros trabalhos relacionados à referência reduzida foi proposto por Webster *et al.* (1993). No vídeo tanto de referência quanto no distorcido são extraídas uma grande quantidade de características dos componentes de cor e luminância do vídeo, tais como filtro de Sobel, filtro de Laplace, transformada rápida de Fourier, derivada de primeira ordem e medições de distorção das cores. Além disso, o algoritmo também analisa como as pessoas percebem as informações temporal e espacial.

Gunawan e Ghanbari (2008) propuseram um avaliador de qualidade do vídeo baseado na força harmônica local (*local harmonic strength – LHS*), ferramenta utilizada para quantificar múltiplas distorções de todos os quadros de

uma sequência sem a necessidade de usar um modelo HVS. Eles investigaram o impacto dos artefatos *blockiness* e *blurriness* sobre o vídeo distorcido após aplicar o modelo LHS.

Um modelo que leva em consideração a percepção de movimento do ser humano em vídeos é proposto por [Rohani et al. \(2010\)](#). A mensuração da qualidade do vídeo é feita após a aplicação de um mapa de pesos baseado na diferença de velocidade dos objetos em cada quadro do vídeo.

Um avaliador de qualidade espaço-temporal é descrito por [Amirshahi e Larabi \(2011\)](#). Eles observaram que o HVS é sensível a mudanças de formas em um vídeo por meio da análise de regiões correspondentes no domínio do tempo. A qualidade do vídeo distorcido é calculada com base na densidade de atividade do movimento.

[Ma et al. \(2012\)](#) exploraram as informações temporal e espacial de ambos os vídeos (de referência e distorcido) extraídos de cada quadro. Além disso, eles analisaram a mudança de energia a partir do domínio do espaço e a estatística da distribuição do histograma dos interquadros a partir do domínio do tempo. Essas informações foram obtidas a partir do descritor de variação de energia e da função de densidade generalizada Gaussiana, respectivamente.

Sem referência (no-reference – NR)

Na metodologia sem referência (ver Fig. 19) não utiliza alguma informação do vídeo de referência para obter a qualidade do vídeo distorcido e ela é normalmente utilizada quando é necessário obter o índice de qualidade em tempo real. Alguns trabalhos relacionados a essa metodologia são [Fu-zheng et al. \(2003\)](#), [Saad e Bovik \(2009\)](#), [Yao et al. \(2009\)](#), [Kim et al. \(2010\)](#), [Kawano et al. \(2010\)](#) e [Lin et al. \(2012\)](#).

[Fu-zheng et al. \(2003\)](#) usaram uma técnica de marca d'água para indicar a qualidade de vídeo distorcido. Esse método adiciona uma marca d'água no vídeo antes de ser transmitido e a extrai quando o vídeo é recebido pelo usuário. A avaliação da qualidade é obtida a partir da comparação entre a marca d'água extraída e a marca d'água original. Eles simularam vários cenários de transmissão de vídeos em redes sem fio e IP. Obteve-se uma correlação de 0,992 com o PSNR a partir de resultados experimentais.

[Saad e Bovik \(2009\)](#) descreveram um avaliador de qualidade de vídeos híbrido utilizando a metodologia sem referência baseado em modelos de pixel e taxa de bits. O modelo criado correlaciona as deficiências ocorridas durante

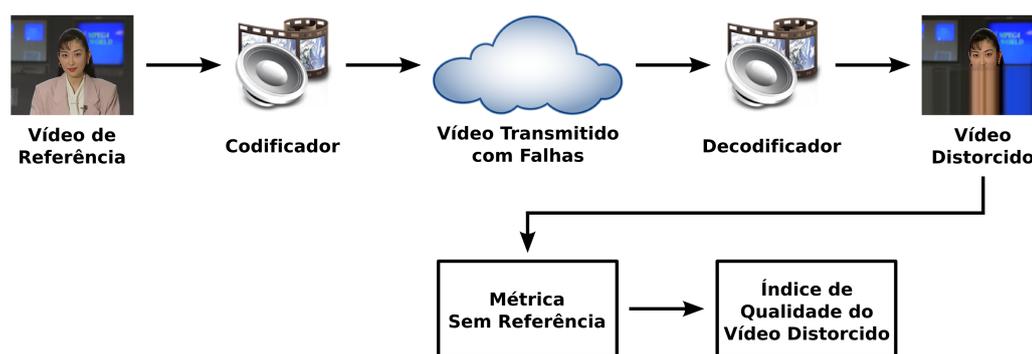


Figura 19: Diagrama utilizado pelas métricas classificadas como sem referência para avaliar a qualidade do vídeo distorcido.

a transmissão (*blockiness* ou efeito fantasma) com o vídeo distorcido.

Um estudo sobre a estatística da movimentação em vídeos de referência e distorcido gerado pela perda de pacotes em uma rede IP foi apresentado por Yao *et al.* (2009). O modelo tem como entrada a informação temporal da intensidade dos pixels de movimento em uma sequência de quadros.

Kim *et al.* (2010) criaram um novo parâmetro para melhorar a acurácia de um avaliador de qualidade de vídeo que utiliza a metodologia sem referência. Eles observaram que uma câmera, após o processo de captura de uma sequência de imagens, controla dinamicamente o intervalo do sinal digital devido ao processo de quantização. Se esse intervalo não for corretamente criado, valores próximos às bordas serão mostrados repetidamente. Um modelo de qualidade de vídeo é então criado, o qual possui como entrada os seguintes parâmetros: brilho, análise de movimento, contraste, pixels das bordas e a cromaticidade. Esse modelo foi validado com vinte estudantes e obteve uma correlação de Pearson igual a 0,847.

Um estimador de QoE para serviços de *streaming* de vídeos foi proposto por Kawano *et al.* (2010). Esse estimador é baseado nos artefatos de *blockiness* e *motion-blur*, que são extraídos a partir do vídeo distorcido. Testes subjetivos permitiram ajustar esse estimador e a sua acurácia foi comparada com o PSNR.

Todos os trabalhos mostrados anteriormente levaram em consideração todos os pixels contidos em cada quadro. Lin *et al.* (2012) introduziram um método baseado em regiões de interesse (*regions of interest* – ROI). Como os olhos do ser humano se concentram em regiões específicas (ROIs), essas regiões podem ser utilizadas para melhorar a acurácia do preditor da qualidade de vídeos. Nesse trabalho, para cada quadro, eles analisaram a influência dos

artefatos de *blockiness* e *blurriness* nas regiões de interesse encontradas para, então, predizer a qualidade do vídeo distorcido.

Na Tabela 15 é sumarizada uma análise comparativa das metodologias objetivas utilizadas para avaliar a qualidade de vídeos, relacionando os seguintes parâmetros: metodologia de referência, tipo de métrica e o que foi analisado em cada proposta.

Tabela 15: Análise comparativa dos métodos utilizados na abordagem objetiva.

Propostas	Parâmetros		
	Metodologia	Métrica	O que é analisado?
MSE (SHEIKH E BOVIK, 2006)	Referência Completa	Dados	Todos os pixels
SSIM (WANG <i>et al.</i> , 2004)	Referência Completa	Imagem	Variação dos padrões locais da intensidade dos pixels
VSNR (CHANDLER E HEMAMI, 2007)	Referência Completa	Imagem	Luminância dos pixels
IFC (SHEIKH <i>et al.</i> , 2005)	Referência Completa	Imagem	Informação estatística dos pixels
VIF (SHEIKH E BOVIK, 2006)	Referência Completa	Imagem	Relação entre a entrada e a saída o canal HVS
PSNR (WINKLER E MOHANDAS, 2008)	Referência Completa	Dados	Todos os pixels
Webster <i>et al.</i> (1993)	Referência Reduzida	Imagem	Componente de luminância do vídeo e informações temporal e espacial
Gunawan e Ghanbari (2008)	Referência Reduzida	Imagem	Características do LHS, e artefatos de <i>blockiness</i> e <i>blurriness</i>
Rohani <i>et al.</i> (2010)	Referência Reduzida	Imagem	Percepção humana com relação ao movimento em vídeos
Amirshahi e Larabi (2011)	Referência Reduzida	Imagem	Sensibilidade do modelo HVS com relação às mudanças de formas em vídeos
Ma <i>et al.</i> (2012)	Referência Reduzida	Imagem	Informações temporal e espacial
Fu-zheng <i>et al.</i> (2003)	Sem Referência	Imagem	Informações sobre a técnica de marca d'água
Saad e Bovik (2009)	Sem Referência	Pacote	Falhas ocorridas durante a transmissão
Yao <i>et al.</i> (2009)	Sem Referência	Pacote	Influência da perda de pacotes
Kim <i>et al.</i> (2010)	Sem Referência	Imagem	Brilho, análise do movimento, contraste e crominância
Kawano <i>et al.</i> (2010)	Sem Referência	Imagem	Influência dos artefatos <i>blockiness</i> e <i>motion-blur</i>
Lin <i>et al.</i> (2012)	Sem Referência	Imagem	Influência dos artefatos <i>blockiness</i> e <i>motion-blur</i> nas regiões de interesse

3.3.3 Abordagem Híbrida

Na abordagem híbrida, combinam-se vantagens das abordagens subjetiva e objetiva. A qualidade do vídeo distorcido é avaliada por um grupo de pessoas e os resultados obtidos dessas avaliações são utilizados para formular o modelo de predição de qualidade. Características não relacionadas ao vídeo podem ser empregadas pelo modelo, tais como características do dispositivo (tamanho da tela ou resolução de reprodução do vídeo) e informações sobre a transmissão (perda de pacotes, atraso ou *jitter*).

Mohamed e Rubino (2002) descreveram uma metodologia chamada avaliação de qualidade pseudo-subjetiva (*pseudo-subjective quality assessment* – PSQA) para automatizar a avaliação da qualidade de vídeos no dispositivo do usuário. Ela analisa parâmetros de comportamento do vídeo e da rede (*bitrate*, tipo do frame, *frame rate* e taxa de perda de pacotes em rajada). Após isso, o modelo de predição da qualidade do vídeo distorcido é criado a partir do treinamento de uma rede neural aleatória (*random neural network* – RNN).

Rubino *et al.* (2006) usaram o PSQA para gerenciar a qualidade percebida pelo usuário em tempo real. A qualidade do vídeo exibido no receptor é controlada por um esquema de QoS, o qual gerencia o controle de admissão, alteração nos parâmetros do codificador, filtra o tráfego para evitar congestionamentos, e utiliza o esquema DiffServ para priorizar o acesso. Os resultados obtidos em experimentos foram analisados com classificadores ANN e Bayesiano.

Piamrat *et al.* (2008) propuseram um mecanismo de controle de admissão inspirado na QoE em redes sem fio. O recurso da transmissão em redes sem fio é dinamicamente controlado na rede de acesso (*access point*), enquanto a QoE é obtida em tempo real. A taxa de perda de pacotes e o tamanho médio das rajadas de perda são considerados durante o treinamento da RNN para, então, predizer a qualidade do vídeo percebida pelo usuário.

Khan *et al.* (2008) apresentaram um sistema de inferência adaptativo neuro-fuzzy. Esse modelo de aprendizagem estima a qualidade do vídeo percebida pelo usuário para um serviço de transmissão contínua de vídeos em redes sem fio por meio da combinação de parâmetros da rede e da aplicação (*frame rate*, taxa de bits por segundo, largura da banda e perda de pacotes). O resultado é expresso em termos do MOS.

Yamagishi *et al.* (2009) propuseram um modelo híbrido utilizando a metodologia sem referência para avaliar a qualidade de vídeos. Esse modelo

analisa o cabeçalho dos pacotes recebidos baseado nas características do sinal do vídeo. O valor estimado da qualidade é obtido a partir da observação dos efeitos durante o processo de compressão do vídeo e não leva em consideração os efeitos da perda dos pacotes.

Ghareeb e Viho (2010) associaram o PSQA no contexto de codificação múltipla descritiva (*multiple description coding* – MDC) para melhorar a qualidade de um serviço de *streaming* de vídeos em vários caminhos de sobreposição. O MDC é uma técnica que provê resiliência de erros, a qual fragmenta o fluxo de mídia em vários sub-fluxos, designados como descritores. Esse trabalho analisa a influência da perda dos quadros tipo I, P e B, o tamanho do grupo de imagens (*Group of Picture* – GoP), e o uso ou não do MDC na avaliação da QoE percebida pelo usuário durante os experimentos.

Os artefatos *blockiness* e *blurriness* gerados a partir da perda de pacotes foram estudados por Farias *et al.* (2011). Alguns vídeos de referência são codificados com taxa de bits variando de 50 kbps até 400 kbps e, para cada vídeo, a perda de pacotes é simulada a partir da taxa de erros de 0,1% a 10%. A qualidade dos vídeos durante as simulações é calculada a partir do PSNR e mapeada para o MOS. Finalmente, cria-se a métrica de qualidade do vídeo a partir das medições de perda de pacotes, *blockiness* e *blurriness*.

Um monitor em tempo real de QoE para redes sem fio, baseada em RNN e múltiplas redes neurais artificiais (*multiple artificial neural network* – MANN), foi discutido por Cerqueira *et al.* (2012a,b) e Rosário *et al.* (2013). Cada técnica de rede neural correlaciona a qualidade obtida a partir de avaliações subjetivas com parâmetros da rede e do vídeo (porcentagem de perda de quadros do tipo I, P e B, total de perdas, tamanho do GoP e informações temporal e espacial). Usando esse monitor, a qualidade do vídeo pode ser observada em tempo real pelo usuário, enquanto que o serviço de transmissão contínua de vídeo é otimizado de forma a melhorar a QoE percebida pelo usuário.

Na Tabela 16 é mostrada uma análise comparativa das metodologias híbridas utilizadas para avaliar a qualidade de vídeos, baseada no tipo de rede, no método utilizado e nas características usadas.

Tabela 16: Análise comparativa dos métodos utilizados na abordagem híbrida.

Propostas	Parâmetros		
	Tipo da rede	Método	Informações utilizadas
Mohamed e Rubino (2002)	Com fio	RNN	Taxa de bits por segundo, tipo de quadro, <i>frame rate</i> , rajada de perda.
Rubino <i>et al.</i> (2006)	Sem fio	Classificadores ANN e Bayesiano	Taxa de perda, atraso e <i>jitter</i> .
Piamrat <i>et al.</i> (2008)	Sem fio	RNN	MOS.
Khan <i>et al.</i> (2008)	Sem fio	ANN	<i>Frame rate</i> , Taxa de bits por segundo, largura da banda, e taxa de perda de pacotes.
Yamagishi <i>et al.</i> (2009)	Com fio	Modelo de regressão	Taxa de bits por segundo, taxa de perda de pacotes, <i>frame rate</i> , GoP, formato e tipo de codec do vídeo.
Ghareeb e Viho (2010)	Com fio	RNN	Taxa de perda dos quadros do tipo I, P, e B, e tamanho do GOP.
Farias <i>et al.</i> (2011)	Com fio	Modelo de regressão	<i>Blurriness</i> , <i>blockiness</i> , e taxa de perda de pacotes.
Cerqueira <i>et al.</i> (2012a,b) e Rosário <i>et al.</i> (2013)	Sem fio	RNN e MANN	Porcentagem de perda dos quadros do tipo I, P e B, tamanho do GoP, e características do conteúdo dos vídeos.

3.4 Conclusões

Nesse capítulo foram discutidas algumas abordagens para avaliação da qualidade dos vídeos. Essas abordagens observam informações provenientes da comunicação (QoS) ou o *feedback* do usuário com relação ao conteúdo exibido no dispositivo (QoE). No decorrer do texto, foram expostas as etapas que constituem um processo de avaliação da qualidade de vídeos e uma análise comparativa entre os métodos utilizados em cada abordagem.

No próximo capítulo é descrita uma proposta de gerenciamento orientado a QoE em serviços de *streaming* de vídeos para dispositivos móveis. São detalhados o processo de coleta de dados e a elaboração do preditor de qualidade, assim como os componentes que auxiliam na manutenção da QoE percebida pelo usuário durante a recepção do conteúdo no dispositivo móvel.

GERENCIAMENTO DA QoE NA TRANSMISSÃO DE VÍDEOS EM REDES SEM FIO

“A ciência serve para nos dar uma ideia de quão extensa é a nossa ignorância”

Félicité Robert de Lamennais

O rápido desenvolvimento de dispositivos móveis e vídeos de alta definição, bem como a infraestrutura utilizada para transmissão de vídeos em redes IP, exigem uma permanente evolução das técnicas empregadas para avaliar a experiência do usuário com relação ao vídeo exibido em seu dispositivo (QoE). Complementando isso, outras situações podem ser examinadas durante o processo de elaboração ou aprimoramento da avaliação da qualidade do serviço. Dentre elas, estão:

- *Escolha do conteúdo:* diferente de um sistema de televisão por radio-difusão, onde os canais são transmitidos ao mesmo tempo para todos os consumidores e o tempo de troca entre canais é “instantânea”, em um serviço de IPTV ou VoD, por exemplo, os conteúdos são enviados somente para aqueles que escolheram o conteúdo desejado. Quando o usuário escolhe outro conteúdo, pode haver uma demora de poucos segundos para exibição do novo conteúdo. Esse tempo de espera pode ser decorrência de vários fatores, tais como congestionamento do tráfego de dados na rede, distância entre o provedor de conteúdos, características do dispositivo do usuário, dentre outros.

- *Segurança e direitos autorais (Digital Rights Management – DRM)*: quando o serviço de *streaming* de vídeos é pago, utilizam-se técnicas para embalar o conteúdo (criptografar) e garantir que ele seja exibido somente nos dispositivos dos assinantes. Também podem existir regras para gravação do conteúdo, com o objetivo de evitar distribuição não autorizada a outras pessoas. Outras situações podem ocorrer como, por exemplo, o provedor pode sofrer ataques de negação de serviço (*Denial of Service – DoS*), o conteúdo pode ser alterado ou as informações dos assinantes podem ser rastreadas indevidamente.
- *Padronização e interoperabilidade*: ainda não há um modelo que ofereça suporte a diferentes plataformas, dispositivos e codificações. O resultado foi a disseminação de diferentes serviços em diferentes formatos, com a popularização do serviço de transmissão contínua de vídeos e com a redução do custo para os consumidores. Por outro lado, uma padronização pode retardar futuras atualizações no sistema.
- *Cobertura do sinal*: com a popularização da transmissão de vídeos em redes IP, principalmente em dispositivos móveis, esse serviço deve se adaptar à heterogeneidade das redes sem fio, pois cada uma delas possui características distintas, tais como alcance do sinal, taxa de transmissão de dados, suporte a medição de qualidade, dentre outros. Além disso, podem acontecer muitas falhas de comunicação em decorrência do deslocamento do dispositivo móvel, por exemplo.
- *Codificação dos conteúdos*: diante da variedade de dispositivos, a codificação dos conteúdos precisa ser bem planejada, com o objetivo de atender as características de cada dispositivo e das condições atuais da comunicação. Para isso, devem-se levar em consideração a taxa de quadros por segundo, a resolução do vídeo, a taxa de bits por segundo etc.

Diante do que foi explicado, o presente trabalho contribui com a proposição, o desenvolvimento e a validação de um sistema que gerencia a qualidade do conteúdo, em tempo real, de um serviço de *streaming* de vídeos em redes sem fio. A modelagem e as principais características desse sistema são detalhadas nas próximas seções.

4.1 Delimitação do Problema

O sistema proposto opera na arquitetura ilustrada na Fig. 20. Nessa arquitetura, pressupõe-se que os conteúdos foram gravados e editados. Por esse motivo, utilizaram-se vídeos de referência disponibilizados em bases de vídeos, conforme relatado na Tabela 8. No presente trabalho, não foram utilizados vídeos com áudio. Os vídeos escolhidos foram codificados em vários formatos e armazenados em um servidor de *streaming*.

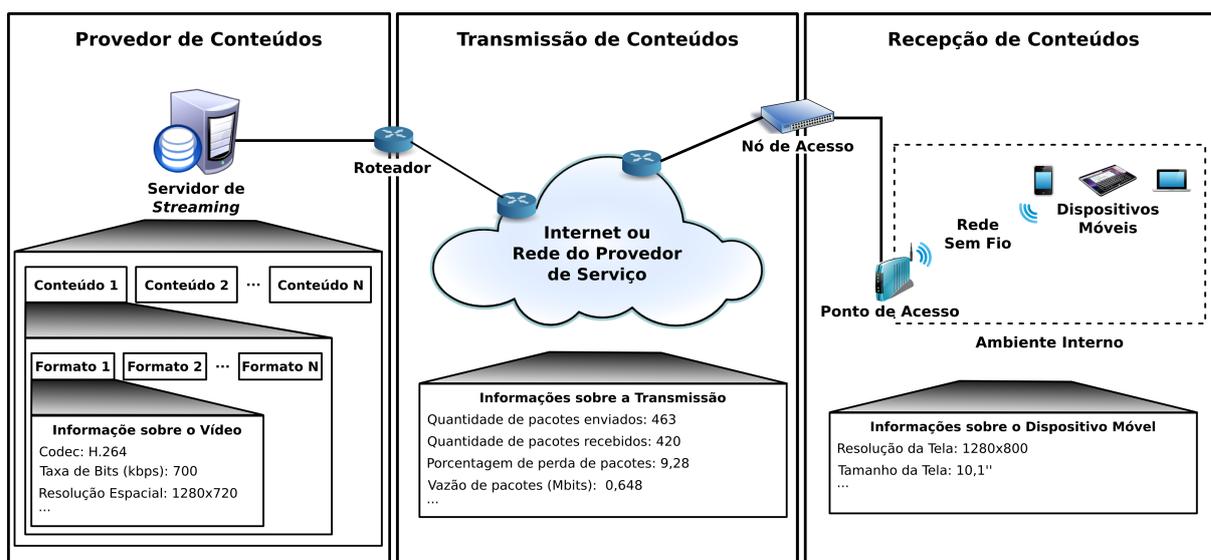


Figura 20: Arquitetura de transmissão de vídeos em que o sistema proposto atua.

Os conteúdos são transmitidos na própria rede do provedor de serviço através dos protocolos TCP, onde são trafegadas as informações sobre o dispositivo e as condições da transmissão para o servidor, e UDP, que é empregado para a transmissão dos conteúdos do servidor para o dispositivo do cliente. No transcorrer da transmissão, são coletados dados para o gerenciador de QoE proposto, tais como quantidade de pacotes enviados, recebidos e perdidos, por exemplo.

O sistema proposto não atua na rede, ou seja, ele não controla o roteamento e a admissão de novas aplicações na rede. O sistema proposto monitora o envio dos conteúdos pelo servidor e a chegada dos mesmos no dispositivo do cliente e, a partir disso, intervém na escolha do formato do conteúdo de acordo com a qualidade do conteúdo observada no dispositivo ou quando o cliente avalia a qualidade de forma negativa.

Os dispositivos utilizados pelos usuários são do tipo móvel (*smartphones* ou *tablets*), conectados com o servidor de *streaming* por meio de uma rede

sem fio (WiFi) em uma residência. Com relação à segurança, o sistema proposto não tem um plano de contingência caso ocorram ataques de negação de serviço (DoS), alteração do conteúdo ou rastreamento indevido das informações dos usuários.

Diante do que foi exposto, o sistema proposto tem como premissas:

- *Adaptação ao estado atual da transmissão e às características do dispositivo móvel:* o sistema coleta informações sobre os pacotes transmitidos (porcentagem de perda de pacotes, largura da banda, vazão da conexão, etc.) e sobre as características do dispositivo móvel (tamanho e resolução da tela) do cliente. Essas informações são úteis para que o sistema gereencie a qualidade do conteúdo exibido no dispositivo móvel do usuário.
- *Gerenciamento da qualidade do conteúdo:* para cada conteúdo oferecido pelo serviço, são disponibilizados vários formatos de vídeo. A escolha da qualidade do conteúdo (formato do vídeo) se dá após uma averiguação das informações oriundas da transmissão, do dispositivo móvel e do conteúdo em dois momentos: (1) quando o conteúdo é escolhido pela primeira vez pelo usuário; e (2) durante a transmissão do conteúdo para o dispositivo móvel do usuário.
- *Avaliação automatizada da qualidade do conteúdo:* para que o usuário não seja incomodado periodicamente com perguntas sobre a qualidade do conteúdo, o sistema analisa, em tempo real, informações sobre a transmissão e o dispositivo móvel. Quando observada alguma anormalia (perda de pacotes, por exemplo), o sistema escolhe um novo formato de vídeo adequado para a condição atual do serviço. A avaliação do usuário, caso ele queira, é levada em consideração nessa análise.

4.2 Metodologia

Para a elaboração do gerenciador da qualidade de experiência orientada a transmissão de vídeos em redes sem fio, definiu-se a metodologia apresentada na Fig. 21. Cada etapa mencionada nessa figura é detalhada a seguir.



Figura 21: Etapas seguidas para a elaboração do sistema proposto.

1 Seleção dos Vídeos de Referência

Diante da grande variedade de bases de vídeo de referência apresentada na Tabela 8, optou-se pela escolha de vídeos cujas informações espacial e temporal fossem distintas com o propósito de avaliar o impacto dessas informações nas avaliações subjetivas. Também foi realizada uma investigação sobre o impacto da variação do GoP e do tamanho máximo do pacote (*maximum transmission unit* – MTU) na qualidade dos vídeos recebidos pelo cliente.

Para a investigação, foram selecionados 12 vídeos disponibilizados em TRACE (2015) (ver Fig. 23), em formato YUV com resolução espacial 352x288 pixels, a uma taxa de 25 quadros por segundo. Na Fig. 22 e na Tabela 17 são mostradas as informações temporal e espacial e as principais características dos vídeos selecionados para investigação, respectivamente.

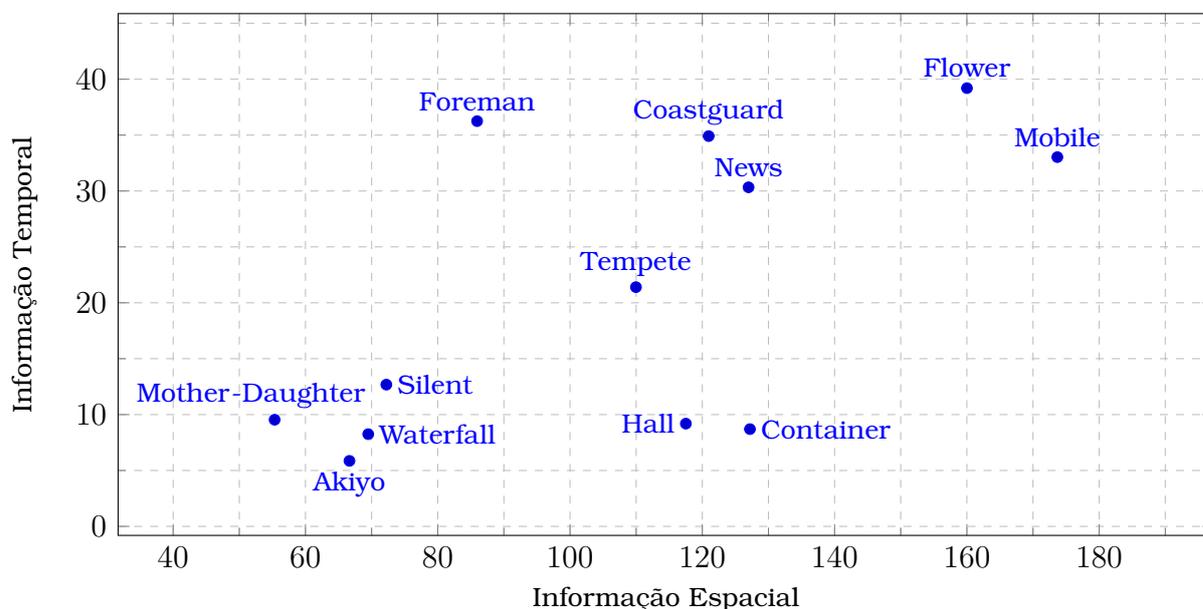


Figura 22: Informações espacial e temporal dos vídeos de referência selecionados para investigação.



(a) Akiyo



(b) Coastguard



(c) Container



(d) Flower



(e) Foreman



(f) Hall



(g) Mobile



(h) Mother-Daughter



(i) News



(j) Silent



(k) Tempete



(l) Waterfall

Figura 23: Snapshot dos vídeos selecionados para investigação.

Tabela 17: Informações sobre os vídeos escolhidos para a investigação.

Nome do Vídeo de Referência	Número de Quadros	Descrição
Akiyo	300	Transmissão de um telejornal com uma apresentadora. Câmera estática.
Coastguard	300	Filmagem de um barco em movimento. Câmera em movimento.
Container	300	Filmagem de um navio de container em movimento. Câmera estática.
Flower	250	Filmagem de um campo florido. Câmera em movimento.
Foreman	300	Gravação de um mestre de obras falando sobre seu trabalho. Câmera em movimento.
Hall	300	Filmagem de pessoas andando pelo corredor de um escritório. Câmera estática.
Mobile	300	Gravação de vários brinquedos em movimento. Câmera em movimento.
Mother-Daughter	300	Filmagem de uma mãe com sua filha. Câmera estática.
News	300	Transmissão de um telejornal com dois apresentadores e um vídeo de fundo. Câmera estática.
Silent	300	Filmagem de um tradutor-intérprete em libras. Câmera estática.
Tempete	260	Gravação de uma flor com folhas caindo ao seu redor. Câmera em movimento.
Waterfall	260	Vista aérea de uma cachoeira. Câmera em movimento.

A resolução espacial foi outra condição utilizada para a seleção dos vídeos de referência para a avaliação subjetiva. Eles deveriam ter pelo menos 1080x720 para que não houvesse perda de qualidade quando os vídeos fossem codificados em resoluções inferiores. O motivo dessa decisão foi pelo fato de averiguar se haveria alguma diferença de percepção da qualidade dos vídeos com relação à resolução dos vídeos e ao tamanho das telas dos dispositivos móveis. Assim, selecionaram-se três vídeos a partir das bases localizadas em [TUM \(2015a,b\)](#): Oktoberfest 2, Old Town Cross e Ski 2 (ver Fig. 25).

Esses foram capturados por uma câmera digital a uma resolução de 1920x1080 pixels e taxa de 25 quadros por segundo. Para cada quadro, foi criada uma imagem em formato *Tagged Image File Format* (TIFF). Os vídeos de referências (sem adição de erros) foram gerados no FFmpeg, um software de código aberto e livre utilizado para manipulação de dados multimídia ([FFMPEG, 2015](#)). O resultado final foi a criação de sequências de vídeo em formato

YUV, mantendo as características originais. Na Fig. 24 e na Tabela 18 são mostradas as informações temporal e espacial e as principais características dos vídeos selecionados para as avaliações subjetivas, respectivamente.

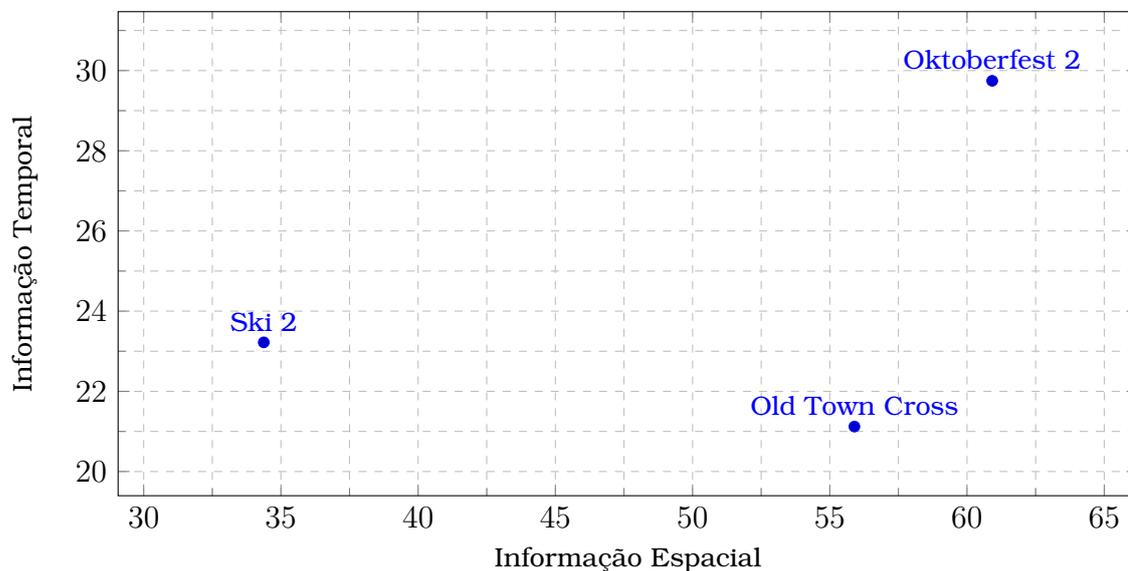


Figura 24: Informações espaciais e temporais dos vídeos de referência selecionados.

Tabela 18: Informações sobre os vídeos escolhidos para as avaliações subjetivas.

Nome do Vídeo de Referência	Número de Quadros	Descrição
Oktoberfest 2	300	Filmagem de cavalos decorados no Oktoberfest. Câmera estática.
Old Town Cross	250	Vista aérea de uma cidade antiga da Alemanha. Câmera em movimento.
Ski 2	300	Descida de um esquiador em uma montanha. Câmera em movimento.



(a) Oktoberfest 2



(b) Old Town Cross



(c) Ski 2

Figura 25: Snapshot dos vídeos selecionados para as avaliações subjetivas.

2 Criação dos Vídeos Distorcidos

O gerenciador de qualidade proposto neste trabalho auxilia na escolha do melhor formato do vídeo em face das informações analisadas acerca da transmissão dos vídeos, das características do dispositivo móvel e dos formatos do conteúdo disponibilizados no servidor. No presente trabalho, foram explorados dois cenários. No primeiro cenário, quando não há perda de pacotes durante a transmissão, o gerenciador escolhe o melhor formato a partir das informações sobre o dispositivo móvel do usuário e o conteúdo. No segundo e último cenário, o gerenciador escolhe o melhor formato do conteúdo conforme as informações observadas durante a transmissão dos pacotes agregadas com as informações do dispositivo móvel e dos formatos disponibilizados para o conteúdo. Esses dois cenários foram utilizados nas avaliações subjetivas.

Para simular o primeiro cenário, os vídeos de referência escolhidos para a avaliação subjetiva foram codificados conforme as características apresentadas na Tabela 19. A escolha dessas informações foi estabelecida com base nos possíveis valores da escala do MOS e na duração máxima para cada sessão de avaliação subjetiva, a qual não poderia ultrapassar 30 minutos. Assim, com o auxílio do FFmpeg e suporte do x264 (VLC, 2015), os vídeos de referência (.yuv) foram codificados em H.264 (.mp4). Ao total, foram criados 84 vídeos distorcidos.

Tabela 19: Parâmetros utilizados para a criação dos vídeos distorcidos utilizados nas avaliações subjetivas para o primeiro cenário.

Resolução espacial	426x240, 640x360, 854x480 e 1280x720
Bitrate (kbps)	50, 150, 300, 500, 700, 1000 e 2000
Frame rate	25
GoP	12
MTU	1024
Ordem de transmissão dos frames	IPBBPBBPBBPB
Ordem de visualização dos frames	IBBPBBPBBPBB

No segundo cenário, foi analisada a qualidade do vídeo percebida pelo usuário quando há perda de pacotes durante a transmissão. De acordo com Wang e Moayeri (1995) e Hasslinger e Hohlfeld (2008), a distribuição Gilbert-Elliott apresenta o melhor comportamento de perda de pacotes em redes sem

fio quando comparada com a distribuição Uniforme. O modelo Gilbert-Elliott é representado por uma cadeia de Markov de dois estados, conforme ilustrado na Fig. 26. No que diz respeito à taxa de perda de pacotes, recomenda-se que ela não ultrapasse 10% (ITU-T, 2007).

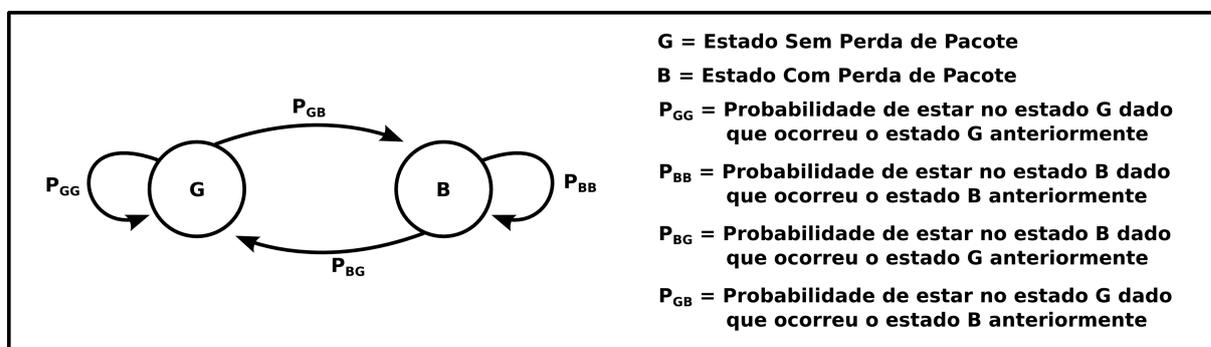


Figura 26: Representação do modelo de perda de pacotes utilizando Gilbert-Elliott.

Diante disso, foram criados 72 vídeos distorcidos para a avaliação subjetiva para o segundo cenário a partir das informações apresentadas na Tabela 20. Para o modelo de perda de pacotes, foram utilizados os seguintes valores: $P_{GG} = 0,96$; $P_{BB} = 0,94$; e $P_G = 0,001$ (HASSLINGER E HOHLFELD, 2008). Com relação à taxa de perda de pacotes (P_B), foram empregados os valores 1%, 3%, 5%, 7%, 9% e 12%.

Tabela 20: Parâmetros utilizados para a criação dos vídeos distorcidos utilizados nas avaliações subjetivas para o segundo cenário.

Resolução espacial (bitrate)	426x240 (700 kbps), 640x360 (700 kbps), 854x480 (1000 kbps) e 1280x720 (2000 kbps)
Frame rate	25
GoP	12
MTU	1024
Ordem de transmissão dos frames	IPBBPBBPBBPB
Ordem de visualização dos frames	IBBPBBPBBPBB

Para a investigação da variação do GoP e MTU na qualidade dos vídeos recebidos pelo cliente, também foram analisados dois cenários: (1) quando não há perda de pacotes durante a transmissão dos pacotes entre o servidor e o cliente; e (2) quando há perda de pacotes durante a transmissão dos conteúdos. A qualidade dos vídeos foi avaliada através do PSNR. Após isso, utilizou-se o mapeamento do PSNR para o MOS, conforme mostrado na Tabela 14.

Para simular o primeiro cenário, os vídeos de referência escolhidos para a investigação foram codificados de acordo com as características descritas na Tabela 21. A escolha dos parâmetros foi estabelecida com base nos possíveis valores do mapeamento do PSNR para o MOS. Dessa forma, com o auxílio do FFmpeg e suporte do x264, os vídeos de referência (.yuv) foram codificados em H.264 (.mp4). Ao total, foram criados 1296 vídeos distorcidos.

Tabela 21: Parâmetros utilizados para a criação dos vídeos distorcidos utilizados na investigação da variação do GoP e MTU para o primeiro cenário.

Resolução espacial	352x288
Bitrate (kbps)	10, 15, 20, 50, 100, 150, 300, 500, 700, 1000, 1500 e 2000
Frame rate	25
GoP	6, 12 e 18
MTU	216, 512 e 1024
Ordem de transmissão dos frames	IPBBPBBPBBPB
Ordem de visualização dos frames	IBBPBBPBBPBB

No segundo cenário, também foi utilizado o modelo Gilbert-Elliott para gerar vídeos distorcidos com introdução de erro. Assim, foram criados 7776 vídeos distorcidos para investigação da variação do GoP e MTU a partir das informações apresentadas na Tabela 22.

Tabela 22: Parâmetros utilizados para a criação dos vídeos distorcidos utilizados na investigação da variação do GoP e MTU para o segundo cenário.

Resolução espacial	352x288
Bitrate (kbps)	10, 15, 20, 50, 100, 150, 300, 500, 700, 1000, 1500 e 2000
Frame rate	25
GoP	6, 12 e 18
MTU	216, 512 e 1024
Taxa de Perda de Pacotes (%)	1, 3, 5, 7, 9 e 12
Ordem de transmissão dos frames	IPBBPBBPBBPB
Ordem de visualização dos frames	IBBPBBPBBPBB

3 Avaliação Subjetiva dos Vídeos Distorcidos

Nessa etapa, as avaliações da qualidade dos vídeos distorcidos (sem e com introdução de erros) criados para as avaliações subjetivas foram coletadas. Definiu-se que pelo menos 20 pessoas eram necessárias para obter um valor estatístico significativo sobre a qualidade desses vídeos. Antes do início das avaliações subjetivas, foram elaborados dois documentos: (1) Termo de Consentimento Livre e Informado e (2) Formulário sobre os dados pessoais do(a) voluntário(a).

No Termo de Consentimento Livre e Informado (ver Apêndice B) foi explicado qual era a finalidade da pesquisa e suas possíveis utilizações. Já no Formulário sobre os dados pessoais do(a) voluntário(a) (ver Apêndice C) foram perguntadas algumas informações pessoais (nome, e-mail para contato e escolaridade, por exemplo) e a respeito do tempo médio de visualização de vídeos provenientes da Internet e em que dispositivos (televisão, computador/laptop, *smartphone* ou *tablet*). Esses dois documentos foram enviados por e-mail para alunos de graduação e pós-graduação de algumas instituições de ensino localizadas em Manaus/AM e Minas Gerais/MG.

No decorrer da chegada das respostas, contactou-se com cada pessoa para saber a disponibilidade de um dia e horário para avaliar a qualidade dos vídeos distorcidos. Após isso, seguiu-se o fluxograma mostrado na Fig. 27. Antes da avaliação da qualidade dos vídeos distorcidos, cada voluntário(a) realizou um teste de acuidade visual (ver Apêndice D) e teste de deficiência visual de cores (ver Apêndice E). Caso o(a) voluntário(a) passasse nesses dois testes, ele(a) estaria apto(a) para avaliar a qualidade dos vídeos.

Os voluntários foram divididos em dois grupos: os que assistiriam os vídeos em um *tablet* e os que assistiriam os vídeos em um *smartphone*. O objetivo disso era averiguar se haveria alguma diferença de qualidade percebida quando a pessoa assistia o vídeo em uma tela grande (*tablet*) e em uma tela pequena (*smartphone*). Dessa forma, coletou-se pelo menos 20 amostras de avaliações de qualidade para cada dispositivo. Na Fig. 28 e na Tabela 23 são apresentados os dispositivos utilizados e suas principais características, respectivamente.

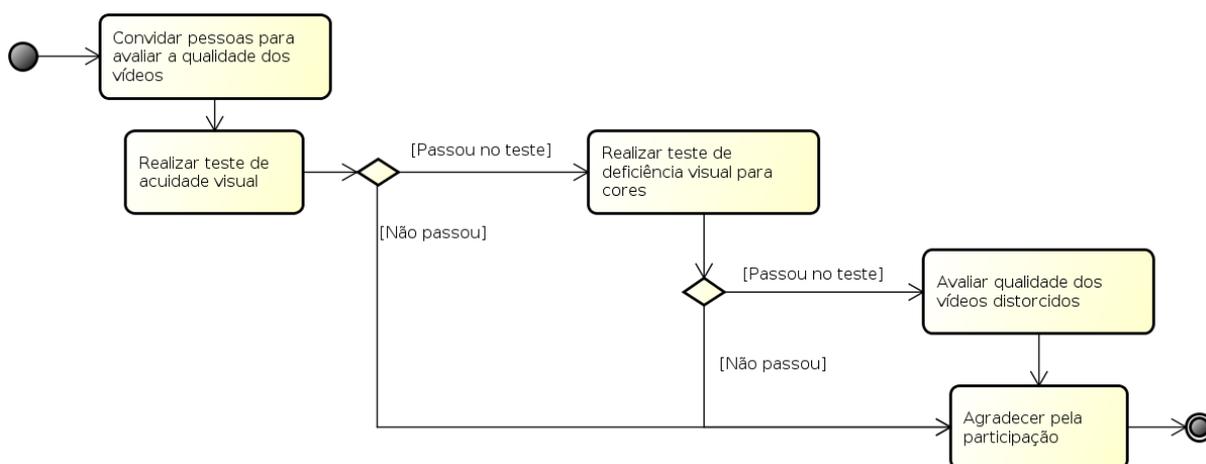


Figura 27: Fluxograma utilizado para avaliação subjetiva dos vídeos distorcidos.



Figura 28: Dispositivos utilizados nas avaliações subjetivas.

Tabela 23: Características dos dispositivos utilizados nas avaliações subjetivas.

Características	Samsung Galaxy Tab 3 10.1	Samsung Galaxy S4 Mini
Tela (polegada/pixels)	10,1" 1280x800	4,3" 960x540
Sistema Operacional	Android 4.2.2	Android 4.2.2
Processador	dual core, 1,6 GHz	dual core, 1,7 GHz
Memória RAM	3 GB	1,5 GB
Memória Interna	16 GB	8 GB
Reprodução de Vídeo	720p	1080p
Conexão WiFi	WLAN a/b/g/n	WLAN a/b/g/n
Rede Móvel	3G/4G	3G/4G

Tanto para a primeira etapa (vídeos sem introdução de erros) quanto para a segunda etapa (vídeos com introdução de erros), utilizou-se a ordem de exibição dos vídeos mostrada na Fig. 29. Ela foi definida dessa maneira para

evitar os efeitos contextuais e de memória durante o julgamento da qualidade do vídeo distorcido. Para a aquisição do MOS foi adotado o método ACR de estímulo simples devido as características dos vídeos e pelo tempo disponível (no máximo 30 minutos para cada sessão).

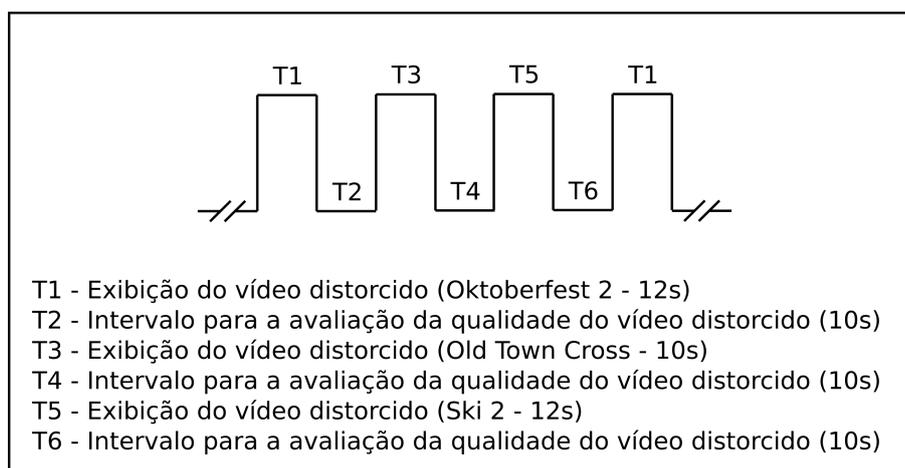


Figura 29: Ordem de apresentação dos vídeos.

Antes da avaliação da qualidade dos vídeos distorcidos, a pessoa assistia alguns vídeos de treinamento para se familiarizar com o ambiente. Eram reproduzidos 6 vídeos e eles não tinham alguma relação com os vídeos distorcidos. Após isso, era iniciada a sessão para reprodução e avaliação da qualidade dos vídeos distorcidos.

Cada vídeo era reproduzido uma única vez em tela cheia no dispositivo móvel. A distância entre o(a) voluntário(a) e o dispositivo era no máximo o comprimento de seu braço. Após a sua reprodução, uma escala de qualidade era exibida com os seguintes valores: Péssimo, Ruim, Aceitável, Bom e Excelente. Ao final da reprodução dos vídeos, era apresentada uma mensagem de agradecimento pela participação na pesquisa. Na Fig. 30 é demonstrado o ambiente utilizado nas avaliações subjetivas.

A ordem de apresentação dos vídeos distorcidos (sem e com introdução de erros) foi seguida conforme descritas nas Tabelas 24 e 25. Essas sequências foram geradas aleatoriamente de modo que dois vídeos consecutivos não poderiam ter o mesmo conteúdo do vídeo de referência.

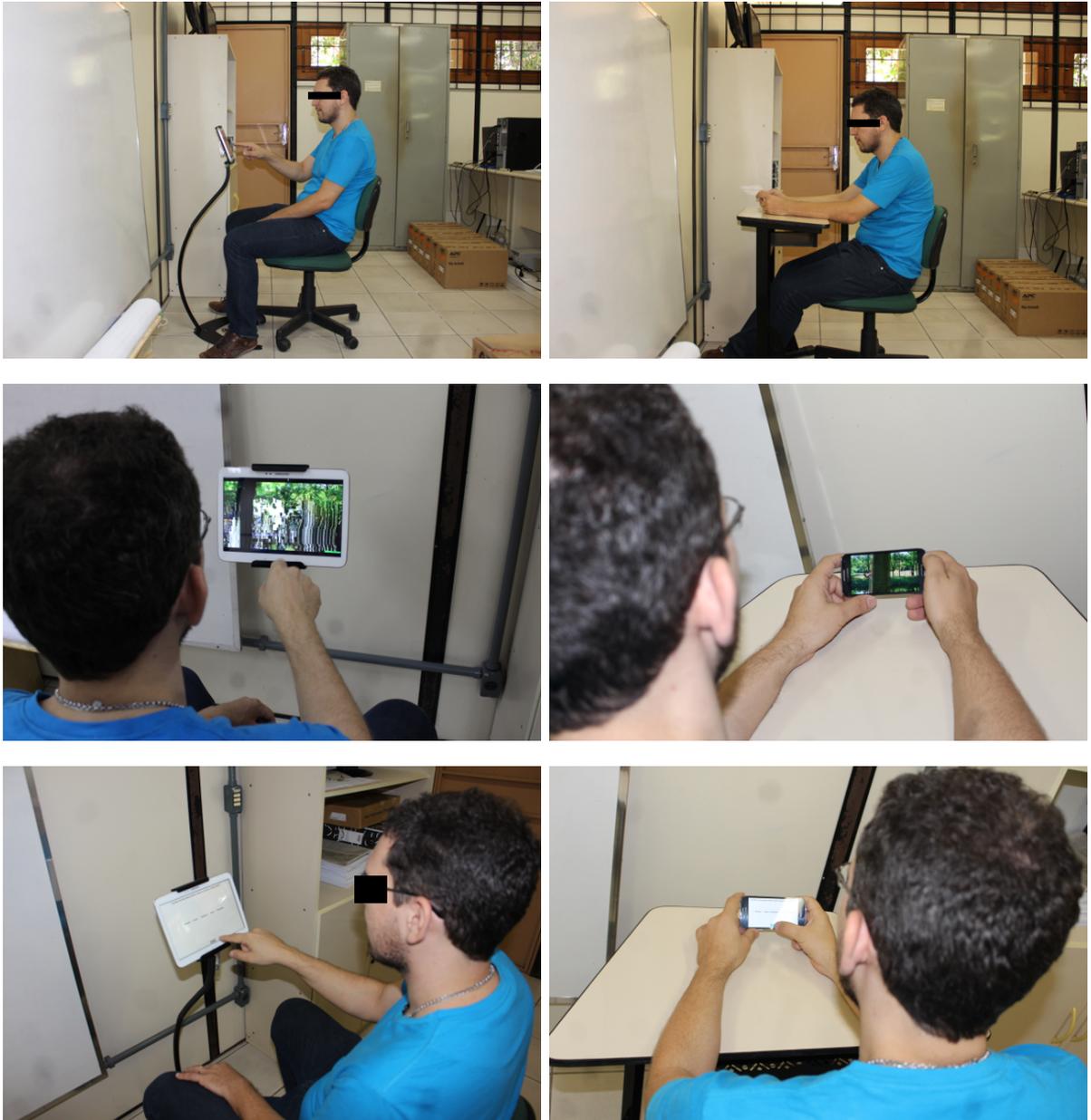


Figura 30: Ambiente utilizado nas avaliações subjetivas.

Tabela 24: Ordem de apresentação dos vídeos distorcidos (sem introdução de erros) nas avaliações subjetivas.

#	Vídeo	#	Vídeo
1	480p_okt2-854-480-25-1000-12-2	43	360p_okt2-640-360-25-300-12-2
2	720p_oldtowncross-1280-720-25-150-12-2	44	480p_oldtowncross-854-480-25-300-12-2
3	360p_ski2-640-360-25-2000-12-2	45	240p_ski2-426-240-25-50-12-2
4	360p_okt2-640-360-25-700-12-2	46	480p_okt2-854-480-25-300-12-2
5	720p_oldtowncross-1280-720-25-50-12-2	47	480p_oldtowncross-854-480-25-1000-12-2
6	360p_ski2-640-360-25-50-12-2	48	720p_ski2-1280-720-25-500-12-2
7	240p_okt2-426-240-25-150-12-2	49	720p_okt2-1280-720-25-300-12-2
8	240p_oldtowncross-426-240-25-1000-12-2	50	480p_oldtowncross-854-480-25-2000-12-2
9	360p_ski2-640-360-25-300-12-2	51	480p_ski2-854-480-25-50-12-2
10	240p_okt2-426-240-25-2000-12-2	52	720p_okt2-1280-720-25-1000-12-2
11	480p_oldtowncross-854-480-25-700-12-2	53	720p_oldtowncross-1280-720-25-700-12-2
12	360p_ski2-640-360-25-700-12-2	54	240p_ski2-426-240-25-500-12-2
13	360p_okt2-640-360-25-500-12-2	55	360p_okt2-640-360-25-150-12-2
14	720p_oldtowncross-1280-720-25-2000-12-2	56	360p_oldtowncross-640-360-25-300-12-2
15	240p_ski2-426-240-25-1000-12-2	57	480p_ski2-854-480-25-300-12-2
16	240p_okt2-426-240-25-1000-12-2	58	360p_okt2-640-360-25-1000-12-2
17	480p_oldtowncross-854-480-25-500-12-2	59	240p_oldtowncross-426-240-25-150-12-2
18	240p_ski2-426-240-25-150-12-2	60	240p_ski2-426-240-25-700-12-2
19	360p_okt2-640-360-25-50-12-2	61	240p_okt2-426-240-25-500-12-2
20	240p_oldtowncross-426-240-25-2000-12-2	62	240p_oldtowncross-426-240-25-300-12-2
21	720p_ski2-1280-720-25-150-12-2	63	360p_ski2-640-360-25-500-12-2
22	720p_okt2-1280-720-25-150-12-2	64	480p_okt2-854-480-25-2000-12-2
23	480p_oldtowncross-854-480-25-50-12-2	65	720p_oldtowncross-1280-720-25-500-12-2
24	720p_ski2-1280-720-25-50-12-2	66	720p_ski2-1280-720-25-700-12-2
25	240p_okt2-426-240-25-700-12-2	67	240p_okt2-426-240-25-50-12-2
26	240p_oldtowncross-426-240-25-500-12-2	68	360p_oldtowncross-640-360-25-500-12-2
27	480p_ski2-854-480-25-1000-12-2	69	720p_ski2-1280-720-25-300-12-2
28	720p_okt2-1280-720-25-50-12-2	70	720p_okt2-1280-720-25-700-12-2
29	360p_oldtowncross-640-360-25-150-12-2	71	720p_oldtowncross-1280-720-25-300-12-2
30	480p_ski2-854-480-25-700-12-2	72	240p_ski2-426-240-25-300-12-2
31	720p_okt2-1280-720-25-500-12-2	73	480p_okt2-854-480-25-150-12-2
32	480p_oldtowncross-854-480-25-150-12-2	74	360p_oldtowncross-640-360-25-1000-12-2
33	360p_ski2-640-360-25-150-12-2	75	720p_ski2-1280-720-25-2000-12-2
34	480p_okt2-854-480-25-50-12-2	76	720p_okt2-1280-720-25-2000-12-2
35	360p_oldtowncross-640-360-25-50-12-2	77	240p_oldtowncross-426-240-25-700-12-2
36	240p_ski2-426-240-25-2000-12-2	78	480p_ski2-854-480-25-2000-12-2
37	360p_okt2-640-360-25-2000-12-2	79	480p_okt2-854-480-25-500-12-2
38	360p_oldtowncross-640-360-25-2000-12-2	80	240p_oldtowncross-426-240-25-50-12-2

Continua na próxima página

GERENCIAMENTO DA QOE NA TRANSMISSÃO DE VÍDEOS EM REDES SEM FIO 85

Continuação da Tabela 24

#	Vídeo	#	Vídeo
39	480p_ski2-854-480-25-150-12-2	81	720p_ski2-1280-720-25-1000-12-2
40	480p_okt2-854-480-25-700-12-2	82	240p_okt2-426-240-25-300-12-2
41	720p_oldtowncross-1280-720-25-1000-12-2	83	360p_oldtowncross-640-360-25-700-12-2
42	480p_ski2-854-480-25-500-12-2	84	360p_ski2-640-360-25-1000-12-2

Tabela 25: Ordem de apresentação dos vídeos distorcidos (com introdução de erros) nas avaliações subjetivas.

#	Vídeo	#	Vídeo
1	360p_okt2-640-360-25-700-12-2-1-0-0.09	37	720p_okt2-1280-720-25-2000-12-2-1-0-0.03
2	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.12	38	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.09
3	720p_ski2-1280-720-25-2000-12-2-1-0-0.09	39	480p_ski2-854-480-25-1000-12-2-1-0-0.09
4	240p_okt2-426-240-25-700-12-2-1-0-0.05	40	240p_okt2-426-240-25-700-12-2-1-0-0.07
5	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.07	41	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.01
6	480p_ski2-854-480-25-1000-12-2-1-0-0.05	42	240p_ski2-426-240-25-700-12-2-1-0-0.12
7	480p_okt2-854-480-25-1000-12-2-1-0-0.03	43	360p_okt2-640-360-25-700-12-2-1-0-0.12
8	240p_oldtowncross-426-240-25-700-12-2-1-0-0.09	44	240p_oldtowncross-426-240-25-700-12-2-1-0-0.03
9	720p_ski2-1280-720-25-2000-12-2-1-0-0.01	45	480p_ski2-854-480-25-1000-12-2-1-0-0.01
10	240p_okt2-426-240-25-700-12-2-1-0-0.01	46	240p_okt2-426-240-25-700-12-2-1-0-0.12
11	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.09	47	360p_oldtowncross-640-360-25-700-12-2-1-0-0.12
12	480p_ski2-854-480-25-1000-12-2-1-0-0.03	48	240p_ski2-426-240-25-700-12-2-1-0-0.07
13	360p_okt2-640-360-25-700-12-2-1-0-0.05	49	720p_okt2-1280-720-25-2000-12-2-1-0-0.09
14	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.07	50	240p_oldtowncross-426-240-25-700-12-2-1-0-0.05
15	240p_ski2-426-240-25-700-12-2-1-0-0.09	51	360p_ski2-640-360-25-700-12-2-1-0-0.07
16	720p_okt2-1280-720-25-2000-12-2-1-0-0.12	52	480p_okt2-854-480-25-1000-12-2-1-0-0.05
17	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.05	53	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.05
18	360p_ski2-640-360-25-700-12-2-1-0-0.01	54	240p_ski2-426-240-25-700-12-2-1-0-0.01
19	360p_okt2-640-360-25-700-12-2-1-0-0.01	55	720p_okt2-1280-720-25-2000-12-2-1-0-0.01
20	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.03	56	360p_oldtowncross-640-360-25-700-12-2-1-0-0.07
21	720p_ski2-1280-720-25-2000-12-2-1-0-0.03	57	240p_ski2-426-240-25-700-12-2-1-0-0.03
22	720p_okt2-1280-720-25-2000-12-2-1-0-0.07	58	480p_okt2-854-480-25-1000-12-2-1-0-0.12
23	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.01	59	360p_oldtowncross-640-360-25-700-12-2-1-0-0.05
24	720p_ski2-1280-720-25-2000-12-2-1-0-0.07	60	360p_ski2-640-360-25-700-12-2-1-0-0.05
25	480p_okt2-854-480-25-1000-12-2-1-0-0.09	61	360p_okt2-640-360-25-700-12-2-1-0-0.03
26	480p_oldtowncross-854-480-25-1000-12-2-1-0-0.03	62	720p_oldtowncross-1280-720-25-2000-12-2-1-0-0.12

Continua na próxima página

Continuação da Tabela 25

#	Vídeo	#	Vídeo
27	720p_ski2-1280-720-25-2000-12-2-1-0-0.05	63	480p_ski2-854-480-25-1000-12-2-1-0-0.07
28	240p_okt2-426-240-25-700-12-2-1-0-0.09	64	240p_okt2-426-240-25-700-12-2-1-0-0.03
29	360p_oldtowncross-640-360-25-700-12-2-1-0-0.01	65	360p_oldtowncross-640-360-25-700-12-2-1-0-0.09
30	480p_ski2-854-480-25-1000-12-2-1-0-0.12	66	360p_ski2-640-360-25-700-12-2-1-0-0.03
31	480p_okt2-854-480-25-1000-12-2-1-0-0.01	67	720p_okt2-1280-720-25-2000-12-2-1-0-0.05
32	240p_oldtowncross-426-240-25-700-12-2-1-0-0.07	68	360p_oldtowncross-640-360-25-700-12-2-1-0-0.03
33	240p_ski2-426-240-25-700-12-2-1-0-0.05	69	360p_ski2-640-360-25-700-12-2-1-0-0.12
34	360p_okt2-640-360-25-700-12-2-1-0-0.07	70	480p_okt2-854-480-25-1000-12-2-1-0-0.07
35	240p_oldtowncross-426-240-25-700-12-2-1-0-0.12	71	240p_oldtowncross-426-240-25-700-12-2-1-0-0.01
36	720p_ski2-1280-720-25-2000-12-2-1-0-0.12	72	360p_ski2-640-360-25-700-12-2-1-0-0.09

4 Análise das Avaliações Subjetivas

Após a coleta das notas nas avaliações subjetivas para os cenários sem e com introdução de erros, calculou-se o MOS para cada vídeo de acordo com a equação 3.

$$\overline{MOS}_{video_{ij}} = \frac{1}{N} \sum_{k=0}^N MOS_{video_{ijk}} \quad (3)$$

onde:

$\overline{MOS}_{video_{ij}}$: média do MOS para a característica j , do vídeo i .

$MOS_{video_{ijk}}$: nota do avaliador k , para a característica j , do vídeo i .

N : quantidade de avaliações coletadas para o vídeo i .

Em seguida, obteve-se o intervalo de 95% de confiança para as notas obtidas para cada vídeo por meio da equação 4.

$$\left(\overline{MOS}_{video_{ij}} - 1,96 * \frac{\sigma_{MOS_{video_{ij}}}}{\sqrt{N}}, \overline{MOS}_{video_{ij}} + 1,96 * \frac{\sigma_{MOS_{video_{ij}}}}{\sqrt{N}} \right) \quad (4)$$

onde:

$\sigma_{MOS_{video_{ij}}}$: desvio padrão do MOS para a característica j , do vídeo i .

O resultado disso foi a criação de dois conjuntos de dados (sem e com introdução de erros) com a finalidade de elaborar o avaliador e o preditor de qualidade de vídeo propostos neste trabalho.

5 *Elaboração do Sistema de Gerenciamento da QoE*

A partir das informações coletadas durante a investigação sobre a influência da variação do GoP e do MTU na qualidade dos vídeos e as avaliações subjetivas dos vídeos distorcidos sem e com introdução de erros, propôs-se o sistema gerenciador em tempo real da QoE para dispositivos móveis, conforme ilustrado na Fig. 31. Ele é constituído por três componentes: *Coletador de Informações*, *Avaliador de Qualidade de Vídeo do Conteúdo*; e *Seletor do Formato de Vídeo do Conteúdo*.

Assim que o usuário escolhe um conteúdo, o sistema de gerenciamento em tempo real de QoE inicia suas atividades por meio do *Coletador de Informações*. Esse componente tem como papel coletar informações acerca do dispositivo móvel, dos formatos disponíveis do conteúdo escolhido pelo usuário, e da transmissão do conteúdo entre o servidor e o dispositivo móvel.

Caso o conteúdo seja reproduzido pela primeira vez, o *Seletor do Formato de Vídeo do Conteúdo* gera uma lista de formatos com suas respectivas notas de acordo com o MOS (entre 1 e 5) para o conteúdo escolhido. Essas notas são preditas por um preditor de qualidade de vídeo, o qual possui como entrada as informações coletadas anteriormente. Nesse primeiro momento, as informações sobre a comunicação entre o dispositivo móvel e o servidor de *streaming* não são utilizadas. O resultado desse processo é a escolha da maior nota de qualidade predita com seu respectivo formato de vídeo para o conteúdo desejado, que posteriormente será reproduzido no dispositivo móvel do cliente.

Durante a reprodução do conteúdo no dispositivo móvel, a cada 10 (dez) segundos, o *Coletador de Informações* é acionado. Após esse tempo, o *Avaliador de Qualidade de Vídeo do Conteúdo* verifica, por meio de outro preditor, se a qualidade de vídeo está satisfatória para o cliente, ou seja, se o MOS predito é pelo menos superior ao índice Bom. Além disso, o cliente, a qualquer momento, pode opinar sobre qualidade de vídeo do conteúdo exibido no dispositivo móvel.

Caso MOS predito ou a opinião do usuário seja inferior ao índice Bom ($MOS < 4$), é solicitado ao *Seletor do Formato de Vídeo do Conteúdo* um novo formato de vídeo a partir das novas informações coletadas. Neste caso, as informações sobre a comunicação são adicionadas ao cálculo da predição, pois podem ter ocorrido perdas de pacotes durante a transmissão do conteúdo.

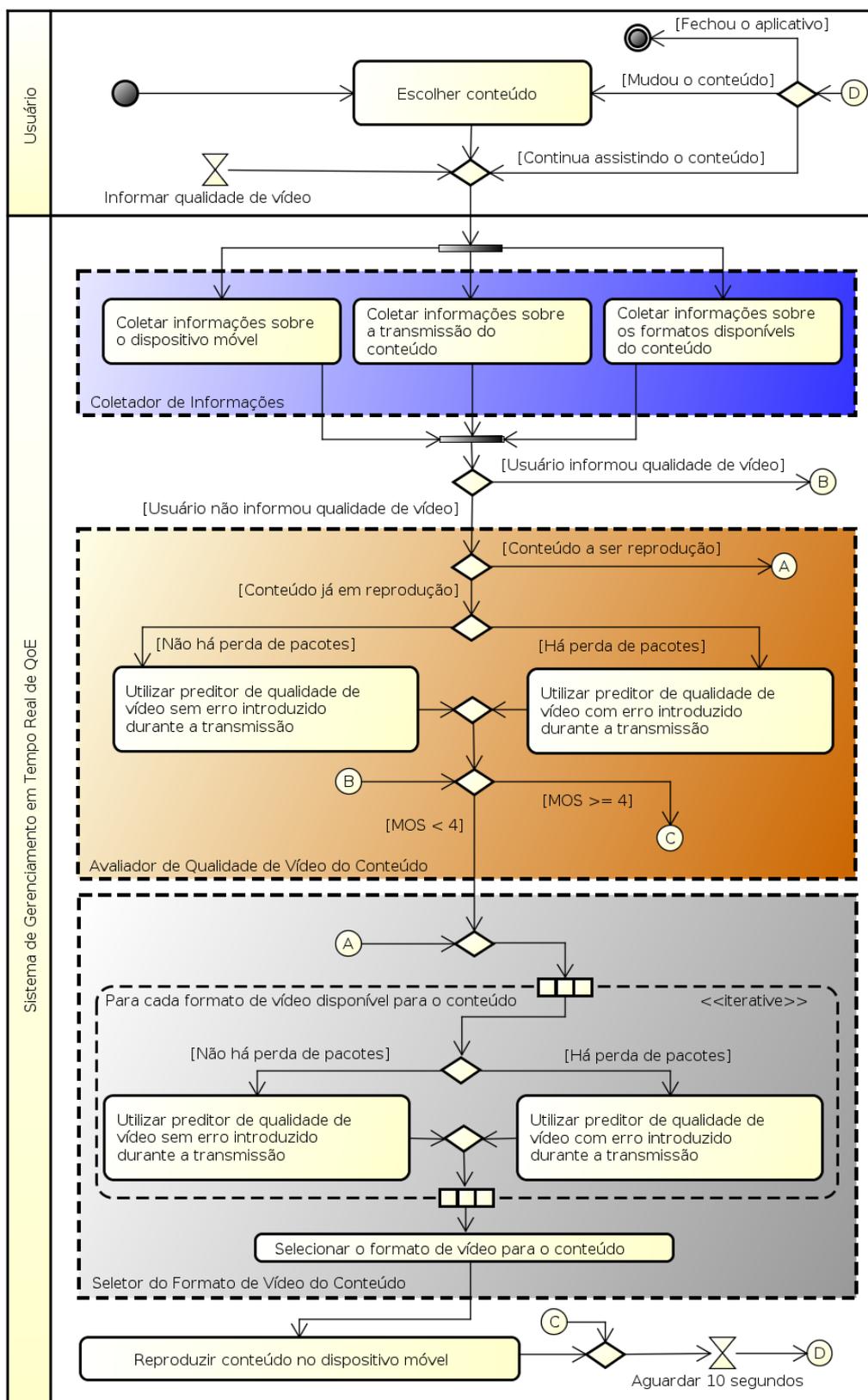


Figura 31: Sistema gerenciador de QoE elaborado neste trabalho.

Tanto no *Avaliador de Qualidade de Vídeo do Conteúdo* quanto no *Seleção do Formato de Vídeo do Conteúdo* são empregados preditores de qualidade. Foram definidos dois preditores: um para os vídeos distorcidos que tinham somente alteração da taxa de *bitrate*; e outro para os vídeos distorcidos com erros introduzidos durante a transmissão do conteúdo. Esses dois preditores foram concebidos de acordo com o fluxograma apresentado na Fig. 32.

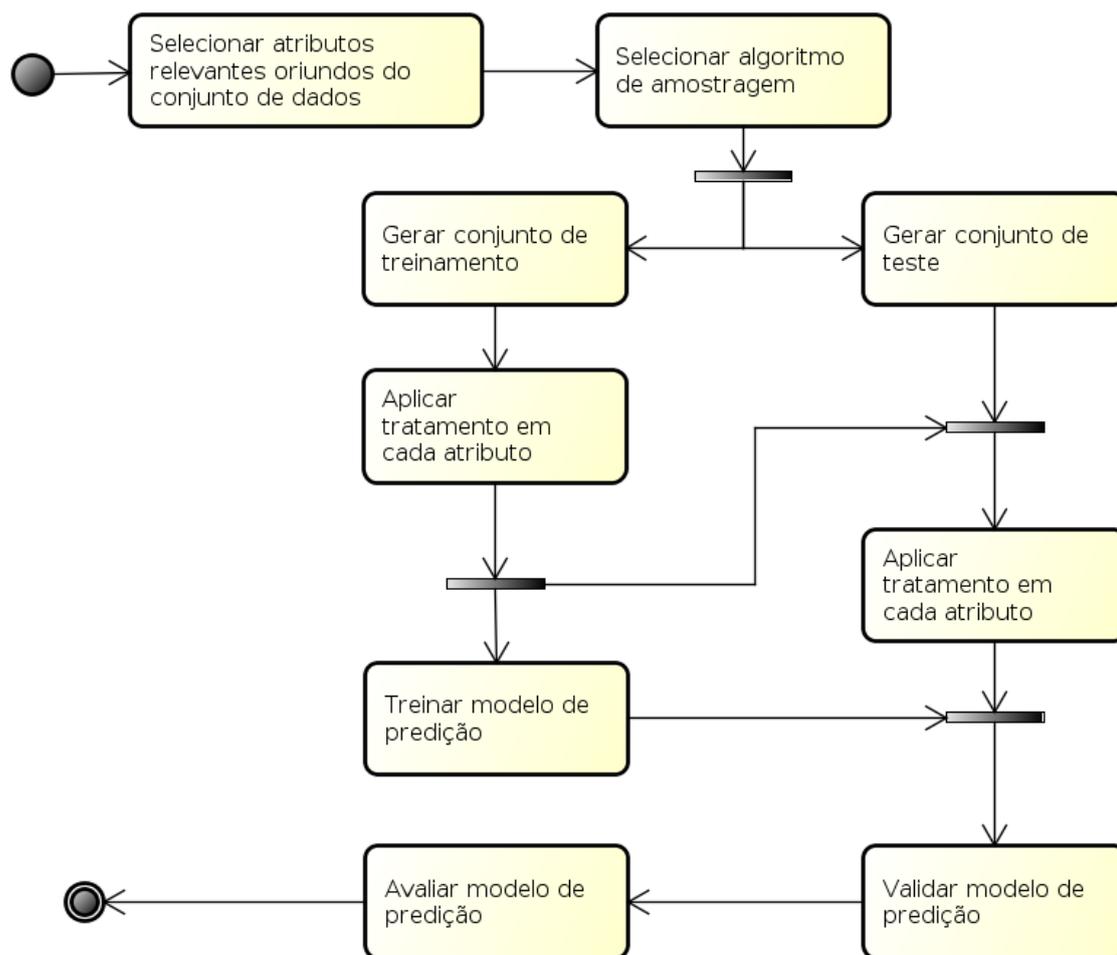


Figura 32: Fluxograma empregado para elaboração dos preditores utilizados no sistema proposto neste trabalho.

O conjunto de dados é representado pelas informações coletadas durante as avaliações subjetivas (sem e com introdução de erros nos vídeos distorcidos). Elas são constituídas por:

- Informações sobre o conteúdo: resolução espacial, informação espacial e temporal, *bitrate*, etc.;

- Informações sobre a transmissão: quantidade de pacotes gerados para os quadros do tipo I, P e B, média de pacotes gerador por segundo, etc.;
- Informações sobre o dispositivo: tamanho da tela e resolução de reprodução do vídeo;
- Média das notas de qualidade informadas para cada vídeo.

Uma das etapas mais importante é a seleção dos atributos utilizados para a criação do preditor. Devido a isso, foram avaliados grupos de atributos de forma a escolher quais atributos trariam o melhor resultado quando o modelo do preditor fosse avaliado. Para isso, definiram-se algumas estratégias: utilizar todos os atributos disponíveis; utilizar somente atributos relacionados ao conteúdo; utilizar somente atributos relacionados à transmissão; utilizar somente atributos relacionados ao dispositivo móvel; selecionar os atributos de acordo com a matriz de correlação; aplicar análise de componentes principais; e empregar alguns algoritmos disponibilizados no *Waikato Environment for Knowledge Analysis* (WEKA), uma ferramenta gratuita e de código aberto usada para minerar e transformar dados em informações úteis (WEKA, 2015).

Dois métodos de amostragem dos dados para criar os conjuntos de treinamento e de teste foram utilizados: *Random Subsampling* e *Cross Validation*. Para o primeiro método, utilizou-se 2/3 da base de dados para treinamento e 1/3 para teste. O resultado final foi calculado a partir da média após a execução de 10 rodadas. No *Cross Validation*, foram utilizados 10 divisões (*fold*). O resultado final (média) foi obtido após percorrer as 10 divisões.

Durante a pesquisa, observou-se que cada atributo possuía uma correlação com o MOS, mesmo que mínima. Diante desse fato, elaborou-se uma técnica de tratamento que consistiu em aplicar uma linha de tendência (exponencial, linear, logarítmico, polinomial de grau 2 ou potência) em cada atributo e selecionar a linha de tendência que obteve a maior correlação com o MOS. Essa técnica está descrita no Algoritmo 1.

Foram utilizados cinco algoritmos para treinamento dos preditores, conforme mostrado na Tabela 26. Além desses algoritmos, também foram empregadas as técnicas de árvore de decisão e vizinho mais próximo utilizando a distância euclidiana. Cada modelo de predição foi treinado e validado com as bases de treinamento e teste e avaliado de acordo com as métricas MSE e o coeficiente de Pearson.

	Entrada: Bases de treinamento e teste
	Saída: Aplicação da melhor linha de tendência em cada atributo na base de teste de acordo com o menor valor do MSE
1	início
2	para cada atributo x da base de treinamento faça
3	Encontrar a linha de tendência exponencial ($resExponencial$) para x de acordo com o MOS correspondente e calcular seu MSE ($mseExponencial$);
4	Encontrar a linha de tendência linear ($resLinear$) para x de acordo com o MOS correspondente e calcular seu MSE ($mseLinear$);
5	Encontrar a linha de tendência logarítmico ($resLogaritmico$) para x de acordo com o MOS correspondente e calcular seu MSE ($mseLogaritmico$);
6	Encontrar a linha de tendência polinomial de grau 2 ($resPolinomial$) para x de acordo com o MOS correspondente e calcular seu MSE ($msePolinomial$);
7	Encontrar a linha de tendência de potência ($resPotencia$) para x de acordo com o MOS correspondente e calcular seu MSE ($msePotencia$);
8	$minMSE = \min (mseExponencial, mseLinear, mseLogaritmico, msePolinomial, msePotencia);$
9	se $minMSE = mseExponencial$ então
10	Aplicar $resPotencia$ em x na base de teste;
11	fim
12	se $minMSE = mseLinear$ então
13	Aplicar $resLinear$ em x na base de teste;
14	fim
15	se $minMSE = mseLogaritmico$ então
16	Aplicar $resLogaritmico$ em x na base de teste;
17	fim
18	se $minMSE = msePolinomial$ então
19	Aplicar $resPolinomial$ em x na base de teste;
20	fim
21	se $minMSE = msePotencia$ então
22	Aplicar $resPotencia$ em x na base de teste;
23	fim
24	fim
25	fim

Algoritmo 1: Proposta de tratamento aplicando uma linha de tendência em cada atributo nas bases de treinamento e teste.

Tabela 26: Algoritmos utilizados para o treinamento dos preditores.

Algoritmo	Equação ¹
Regressão Linear	$y = a + bx$ (5)
Regressão Linear-Logarítmico	$y = a + b * \log(x)$ (6)
Regressão Logarítmico-Linear	$\log(y) = a + bx$ (7)
Regressão Logarítmico-Logarítmico	$\log(y) = a + b * \log(x)$ (8)
Regressão Logística	$y = 1/(1 + \exp(-(1 + b * x)))$ (9)

Dessa forma, obtiveram-se os preditores de qualidade de vídeo utilizados pelo sistema proposto neste trabalho. O diferencial desses preditores está no não uso do vídeo de referência para prever a qualidade do vídeo durante a sua reprodução no dispositivo móvel. Somente as informações sobre a transmissão, a codificação do conteúdo e o dispositivo móvel são levadas em consideração pelos preditores. Assim, não há uma sobrecarga no dispositivo móvel, maximizando o tempo de vida da bateria quando comparado ao caso de avaliar o conteúdo exibido no dispositivo do usuário.

Os preditores de qualidade de vídeo elaborados neste trabalho foram comparados, a partir das métricas MSE e coeficiente de Pearson, com alguns algoritmos utilizados na metodologia objetiva com referência completa. Foram eles: SSIM, *Multi Structural SIMilarity* (MSSIM), VSNR, VIF, *Pixel Based VIF* (VIFP), *Universal Quality Index* (UQI), IFC, *Noise Quality Measure* (NQM), *Weighted Signal-to-Noise Ratio* (WNSR), SNR e PSNR (MOORTHY *et al.*, 2012B).

4.3 Conclusões

Nesse capítulo foi apresentado o sistema de gerenciamento orientado a QoE em serviços de *streaming* de vídeos para dispositivos móveis. Esse sistema visa garantir a qualidade do vídeo durante a sua reprodução em um dispositivo móvel, a partir das análises feitas sobre o conteúdo, a transmissão

¹ a e b são valores a ser encontrados pelo algoritmo, x representa os atributos e y é o valor predito da qualidade do vídeo.

e o dispositivo móvel. No decorrer do capítulo, foi delimitado o problema a ser atacado e descrita em detalhes cada etapa da metodologia definida durante a elaboração do sistema proposto.

As ferramentas e o ambiente utilizados para a implementação desse gerenciador são mostrados no próximo capítulo. Também são apresentadas as análises realizadas durante a coleta das avaliações subjetivas assim como a definição dos preditores de qualidade utilizados pelo sistema proposto neste trabalho.

EXPERIMENTOS E RESULTADOS

*“Tudo o que um sonho precisa para ser realizado é
alguém que acredite que ele possa ser realizado”*

Roberto Shinyashiki

Nesse capítulo são apresentados os experimentos realizados com a finalidade de avaliar o sistema de gerenciamento da QoE para transmissão de vídeos para dispositivos móveis descrito anteriormente. Também são mostradas as ferramentas de *hardware* e *software* usadas para o desenvolvimento e a execução dos experimentos, além de discutir os resultados obtidos durante a pesquisa deste trabalho.

5.1 Ferramentas Utilizadas

O ambiente Linux foi usado como alicerce para a elaboração desta tese, desde o desenvolvimento dos aplicativos, realização das simulações, processamento e análise dos resultados e escrita do texto. Foram empregadas duas distribuições: Ubuntu Server 10.04 e Debian 7.0. Preferiu-se por trabalhar no ambiente Linux devido à sua total integração com as ferramentas de desenvolvimento empregadas.

Os aplicativos foram desenvolvidos utilizando diferentes linguagens de programação: C++ para automatização das simulações, *Tool Command Language* (TCL) para descrição das simulações e Java para coleta das notas (MOS) nas avaliações subjetivas. Para o desenvolvimento utilizando as linguagens C++ e Java, adotaram-se os ambientes multiplataforma e de código aberto

Code::Blocks 10.05 ([CODEBLOCKS, 2015](#)) e Eclipse Indigo ([ECLIPSE, 2015](#)), respectivamente. A modelagem dos fluxogramas contidos neste trabalho foi elaborada na ferramenta Astah Community 6.9 ([ASTAH, 2015](#)).

Os vídeos de referência e distorcidos foram criados com o auxílio das seguintes ferramentas: X264 0.133 ([VLC, 2015](#)), codificador de vídeos em formato H.264; FFmpeg 2.0.1 ([FFMPEG, 2015](#)), codificador e decodificador de áudio e vídeo em vários formatos; Yasm 1.2.0 ([YASM, 2015](#)), assembler recomendado para os softwares x264 e FFmpeg; e MP4Box 0.5.0 ([GPAC, 2015](#)), multiplexador para o container MPEG-4 (mp4).

Para simular a transmissão de vídeos em uma topologia de rede foram empregados o *Network Simulator* (NS) 2.34 ([NS, 2015](#)) e o EvalVid 2.7 ([EVALVID, 2015](#)). O NS é uma ferramenta baseada na teoria de eventos discretos que provê meios de simular comportamentos em vários tipos de redes e protocolos ([ISSARIYAKUL E HOSSAIN, 2011](#); [ALTMAN E JIMÉNEZ, 2012](#)). Já o EvalVid é um framework que possibilita a transmissão de vídeos no NS, além de possuir algumas ferramentas para avaliar a qualidade dos vídeos após a simulação ([KLAUE et al., 2003](#)).

A análise estatística e a implementação dos preditores de qualidade dos vídeos foram efetuadas no *software* Matlab 2012a 7.14. Foram empregados alguns algoritmos disponibilizados nesse *software*, como árvore de decisão (*classregtree*) e vizinho mais próximo (*knnclassify*). As tarefas implementadas e funções utilizadas são compatível com os *software* gratuitos Scilab ([SCILAB, 2015](#)) e GNU Octave ([OCTAVE, 2015](#)).

Essas atividades foram implementadas e realizadas em um servidor com processador Intel Xeon 8-Core X3470 com oito núcleos, memória RAM de 8 Gb DDR3-1333 e HD de 2 TB SATA 6Gb/s. Também foi utilizado um laptop Dell Inspiron N4050 com processador Intel i5 com 4 núcleos, memória RAM de 4 Gb DDR3-1333 e HD de 1 TB SATA.

5.2 Implementação dos Experimentos

A implementação, simulação e análise dos experimentos seguiu o fluxograma descrito na Fig. 33. Antes do início das simulações, todos os vídeos selecionados tanto para a investigação quanto para as avaliações subjetivas tiveram uma etapa de preparação. Essa etapa consistiu em codificar os vídeos selecionados (em formato YUV) em formato H.264 contido no container MPEG-4. Ao final dessa etapa, para cada vídeo codificado, foram gerados vídeos em

formato YUV para avaliação da qualidade pelos algoritmos da metodologia completa com referência completa. Os valores utilizados nos parâmetros são os que foram mostrados nas Tabelas 19 e 21.



Figura 33: Fluxograma da implementação dos experimentos.

Em seguida, no arquivo descritor da simulação, foram escritas todas as informações necessárias para a simulação do cenário desejado. A topologia da rede utilizada nos experimentos e sua representação no ambiente do NS e EvalVid são apresentadas nas Figs. 34 e 35, respectivamente. Os parâmetros de configuração aplicados nessa topologia são detalhados na Tabela 27. Quando foi investigada a influência da perda de pacotes na qualidade dos vídeos, ativou-se o modelo Gilbert-Elliot e foram acrescentados, no parâmetro de entrada do arquivo descritor, os valores de taxa de perda de pacotes.

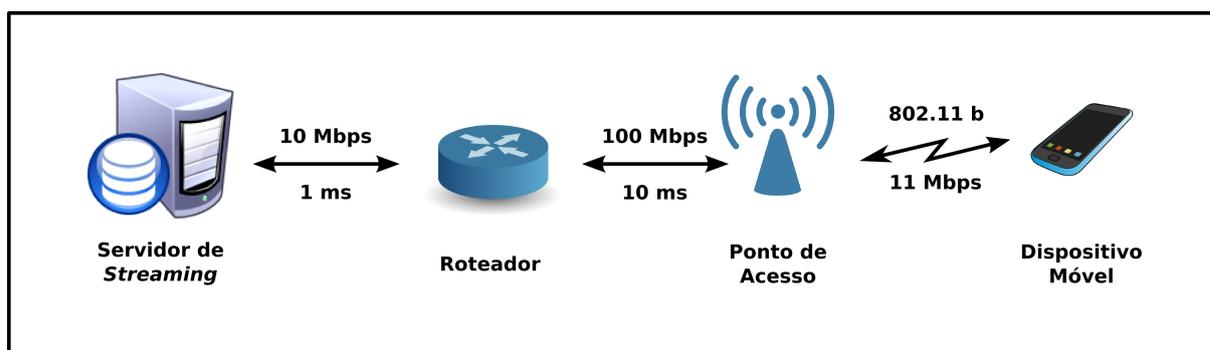


Figura 34: Topologia da rede utilizada nas simulações.

Feito isso, o arquivo descritor da simulação foi enviado para o EvalVid. Não foram adicionadas, durante a simulação, outras aplicações para trafegar na topologia, pois o intuito deste trabalho foi avaliar a influência da perda de

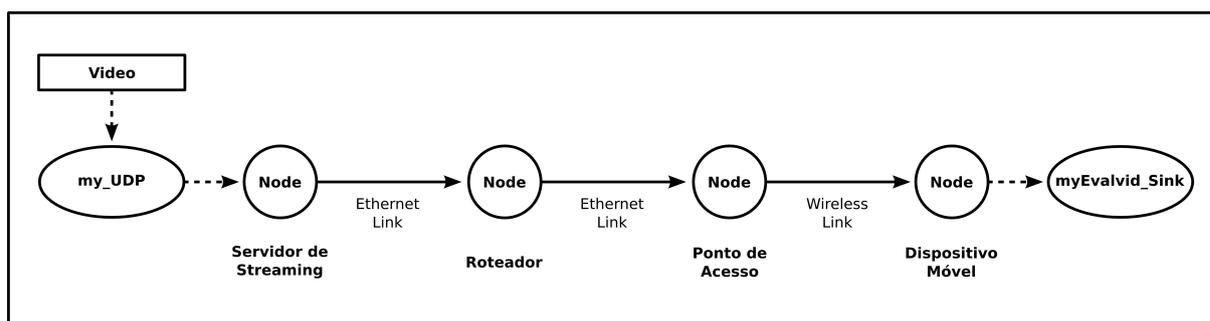


Figura 35: Componentes utilizados no NS e EvalVid para representar a topologia mostrada na Fig. 34.

Tabela 27: Parâmetros utilizados nas simulações.

Parâmetro	Valor
Rede Sem Fio	IEEE 802.11b
Frequência	2,4 GHz
Tipo da Antena	Omnidirecional
Ganho da Antena	12 dB
Altura da Antena	8 metros
Sombreamento	Desvio padrão de 4,4 dB

pacotes entre o ponto de acesso e o dispositivo móvel por meio do uso do modelo Gilbert-Elliot. Dessa forma, havia somente o tráfego de um vídeo por vez entre o servidor de *streaming* e o dispositivo móvel. Cada simulação teve duração, em média, de 60 segundos e foram coletadas diversas informações, tais como quantidade de pacotes enviados e recebidos, *throughput*, taxa de perda de cada tipo de quadro etc. Além disso, geraram-se os vídeos distorcidos (sem e com introdução de erros) que foram reproduzidos e avaliados nas avaliações subjetivas.

Para a coleta das notas durante as avaliações subjetivas, desenvolveu-se um aplicativo para o sistema operacional Android, conforme mostrado na Fig. 36. Na primeira tela (Fig. 36 (a)), eram informados a sequência de reprodução (treinamento ou avaliação), o tipo de dispositivo (*smartphone* ou *tablet*) e o nome do voluntário(a). Após a reprodução de cada vídeo, a tela de avaliação da qualidade do vídeo (Fig. 36 (b)) era exibida com 5 opções (péssimo, ruim, aceitável, bom ou excelente). Após a escolha, a pessoa pressionava o botão avaliar e era reproduzido outro vídeo da sequência.

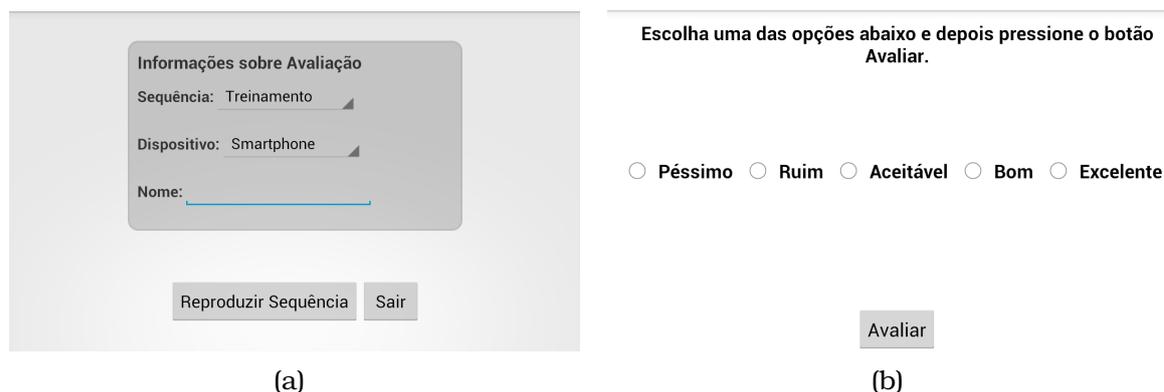


Figura 36: Aplicativo criado para coleta das notas nas avaliações subjetivas.

Após as fases de simulação e coleta das avaliações subjetivas, os resultados obtidos foram tabulados e analisados. Os dados coletados nas simulações e as notas coletadas nas avaliações subjetivas serviram como entrada e saída dos preditores de avaliação de qualidade dos vídeos, respectivamente. O treinamento, a validação e a avaliação desses preditores foram implementadas no Matlab.

5.3 Resultados Obtidos

Investigação sobre a influência da variação do GoP e do MTU na qualidade de vídeos

Essa investigação teve por finalidade encontrar os valores do GoP e MTU que seriam utilizados na criação dos vídeos distorcidos para as avaliações subjetivas. Para isso, utilizaram-se os vídeos descritos na Tabela 17. Foram averiguados dois cenários: variação do *bitrate* e introdução de perda de pacotes.

Os resultados obtidos foram similares tanto para a variação do *bitrate* quanto para a introdução de perda de pacotes nos 12 vídeos. Devido a grande quantidade de dados gerados para esses vídeos com os parâmetros apresentados anteriormente, optou-se, como exemplo, mostrar somente os dados coletados para os vídeos Akiyo, Tempete e Flower.

Quando analisou-se a influência do tamanho do GoP e MTU na qualidade do vídeo de acordo com o *bitrate* (ver Figs. 37 – 39), o melhor resultado encontrado foi 1024 e 18 para MTU e GoP, respectivamente.

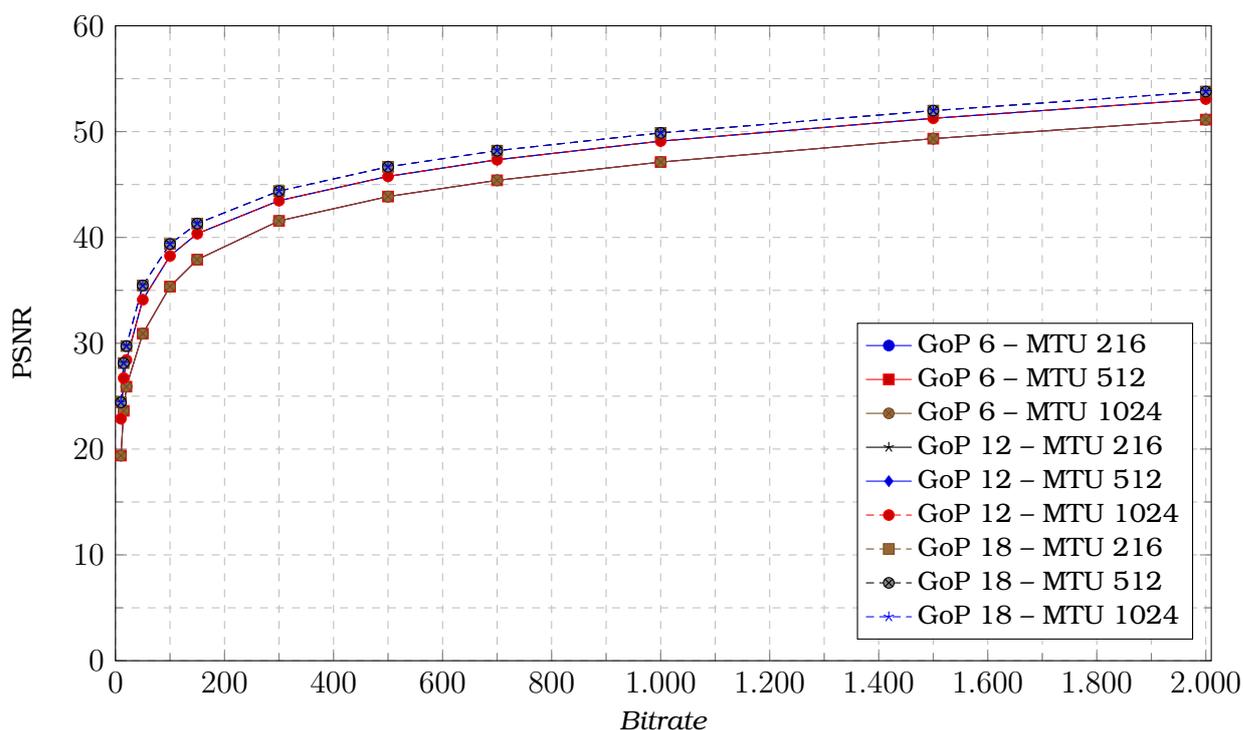


Figura 37: Influência do *bitrate* no vídeo Akiyo de acordo com o GoP e MTU.

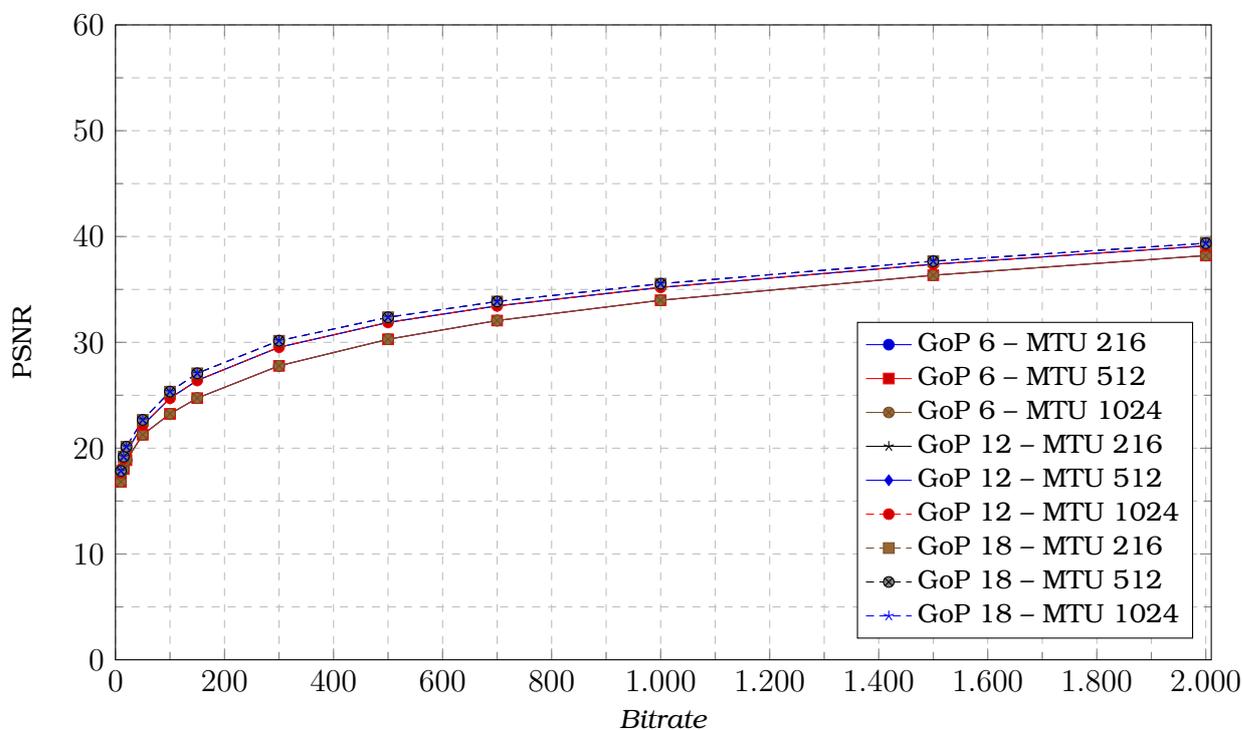


Figura 38: Influência do *bitrate* no vídeo Tempete de acordo com o GoP e MTU.

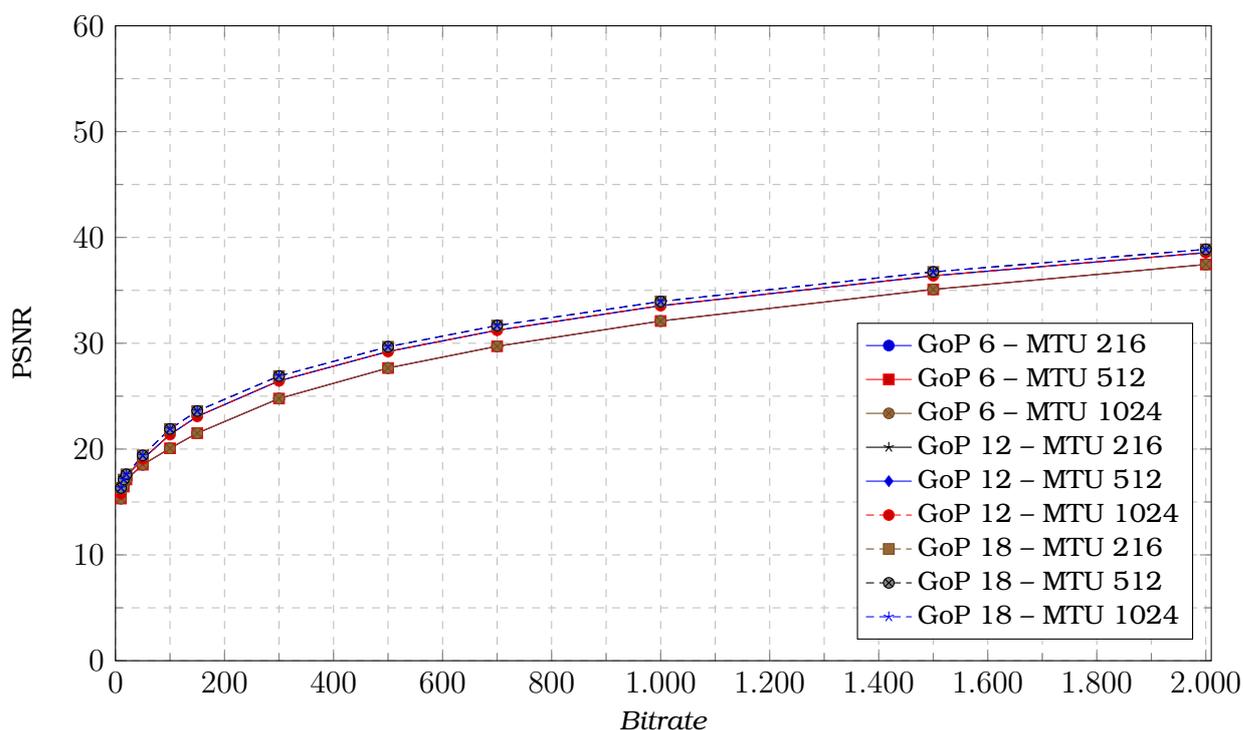


Figura 39: Influência do *bitrate* no vídeo Flower de acordo com o GoP e MTU.

No momento que introduzimos erros (perda de pacotes) nos vídeos Akiyo, Tempete e Flower com *bitrate* igual a 100 kbps, 500 kbps, 1000 kbps e 2000 kbps, os valores 1024 e 6 para MTU e GoP, respectivamente, foi, em média, o melhor resultado alcançado (ver Figs. 40 – 51).

Por outro lado, quando analisou-se a quantidade de dados transmitidos na rede durante sessenta minutos, por exemplo, observou-se que o MTU e GoP igual a 1024 e 12, respectivamente, apresentavam um custo benefício melhor quando eram levadas em consideração a influência da variação do *bitrate* e a taxa de perda de pacotes (ver Tabelas 28 – 30).

Após realizar essas três análises nos 12 vídeos, com diferentes *bitrates* e perdas de pacotes, concluiu-se que a melhor combinação era o uso de 1024 para MTU com GoP igual a 12, pois, em média, a qualidade dos vídeos era aceitável, de acordo com o PSNR, quando a taxa de perda de pacotes era alta, além de gerar menos dados na rede quando comparada, por exemplo, com os valores 1024 e 6 para MTU e GoP, respectivamente.

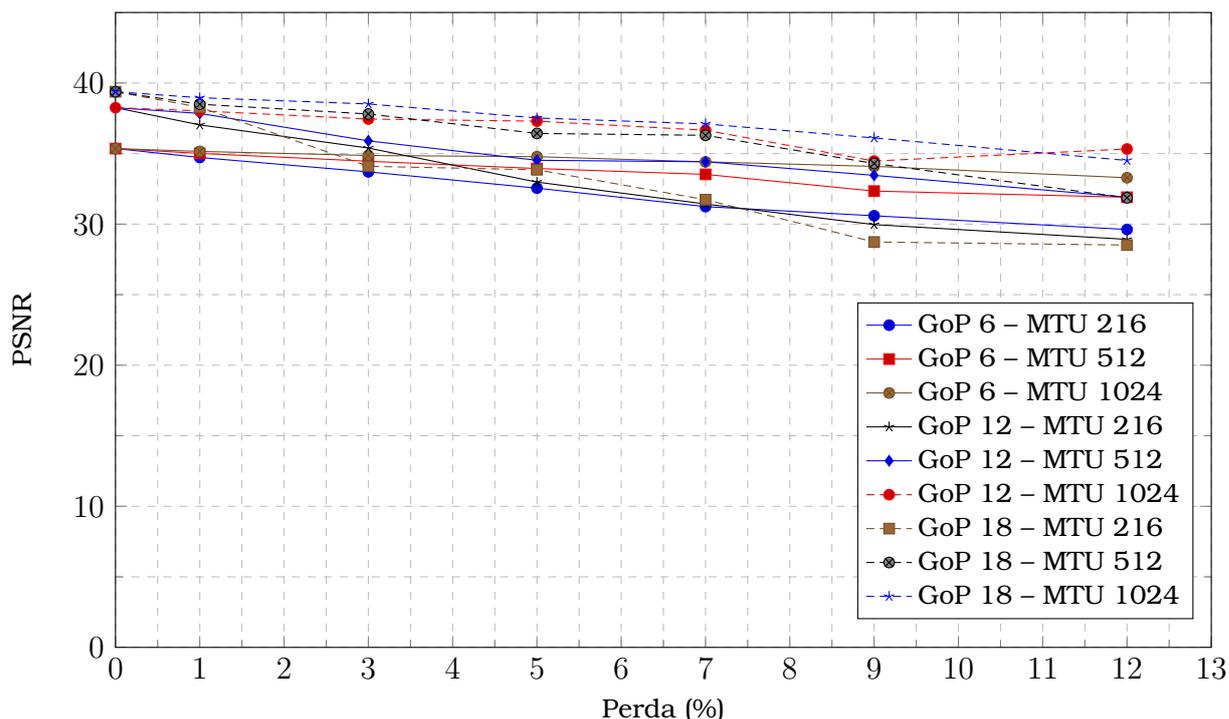


Figura 40: Influência da perda de pacotes no vídeo Akiyo com *bitrate* igual a 100 kbps de acordo com o GoP e MTU.

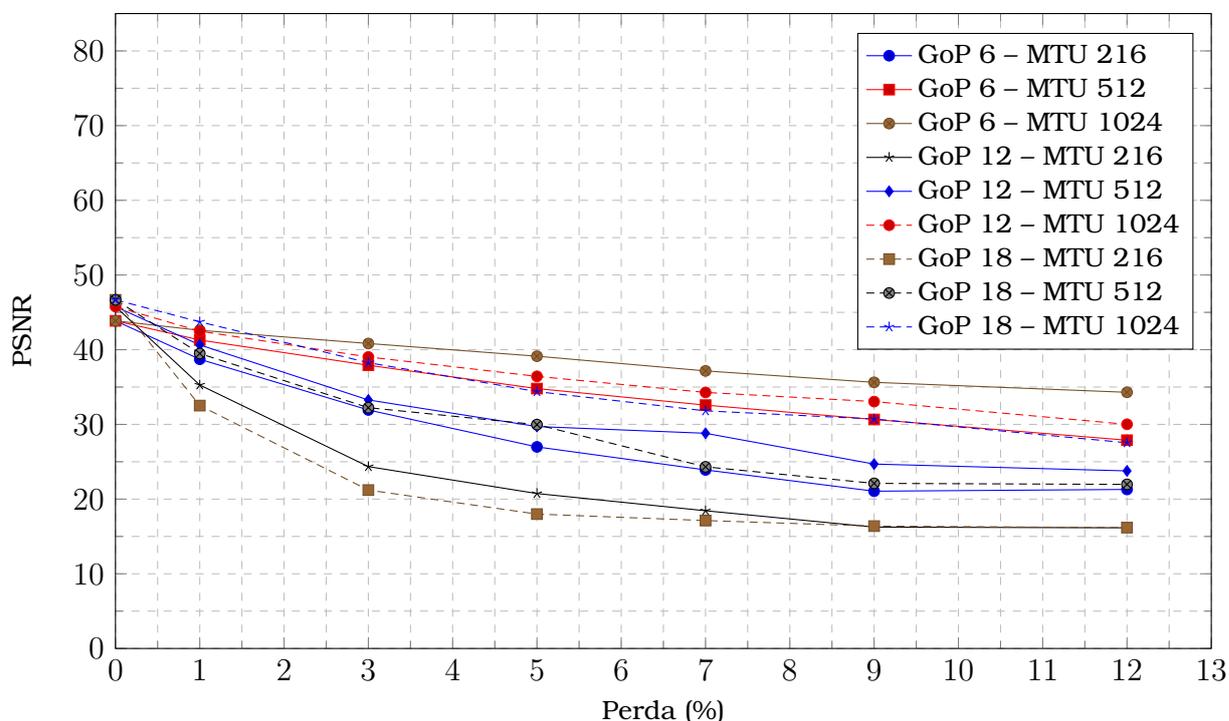


Figura 41: Influência da perda de pacotes no vídeo Akiyo com *bitrate* igual a 500 kbps de acordo com o GoP e MTU.

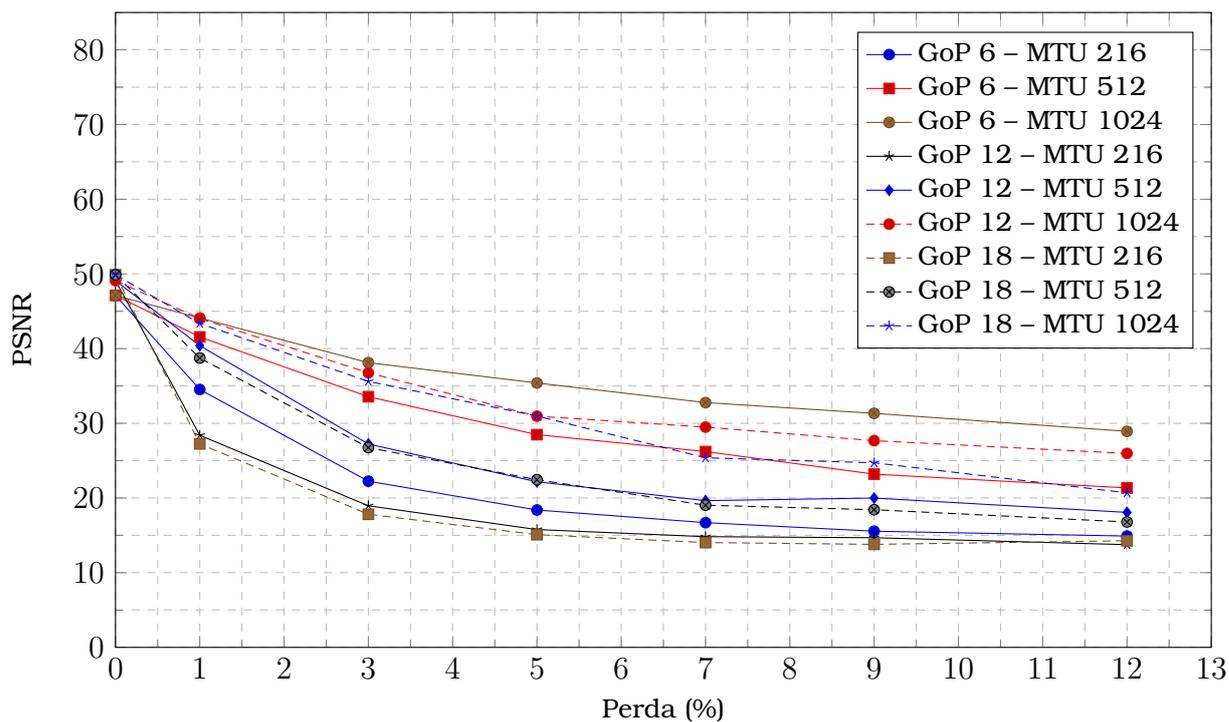


Figura 42: Influência da perda de pacotes no vídeo Akiyo com *bitrate* igual a 1000 kbps de acordo com o GoP e MTU.

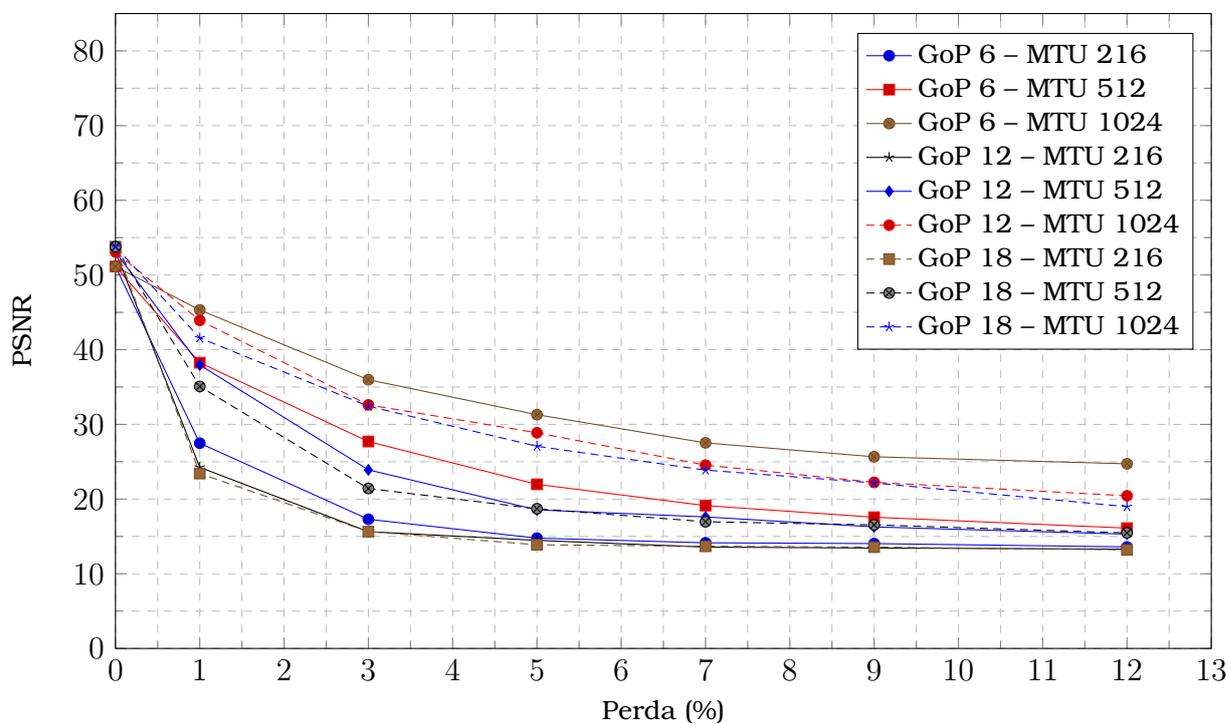


Figura 43: Influência da perda de pacotes no vídeo Akiyo com *bitrate* igual a 2000 kbps de acordo com o GoP e MTU.

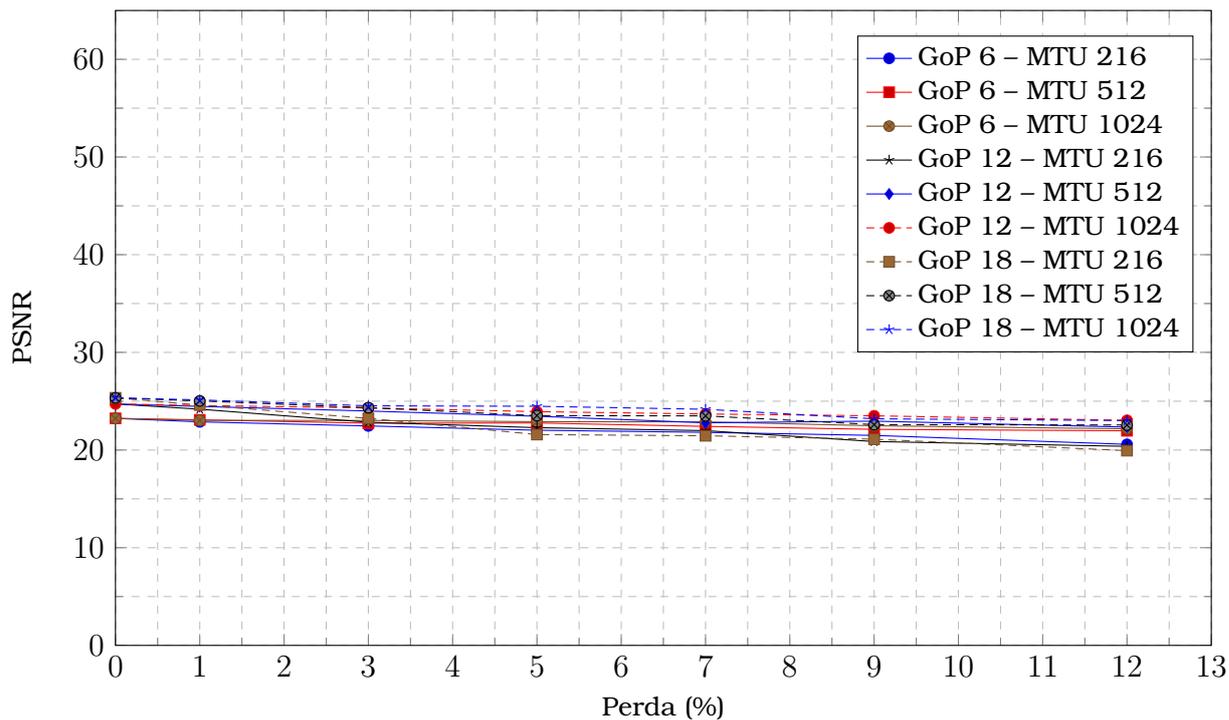


Figura 44: Influência da perda de pacotes no vídeo Tempete com *bitrate* igual a 100 kbps de acordo com o GoP e MTU.

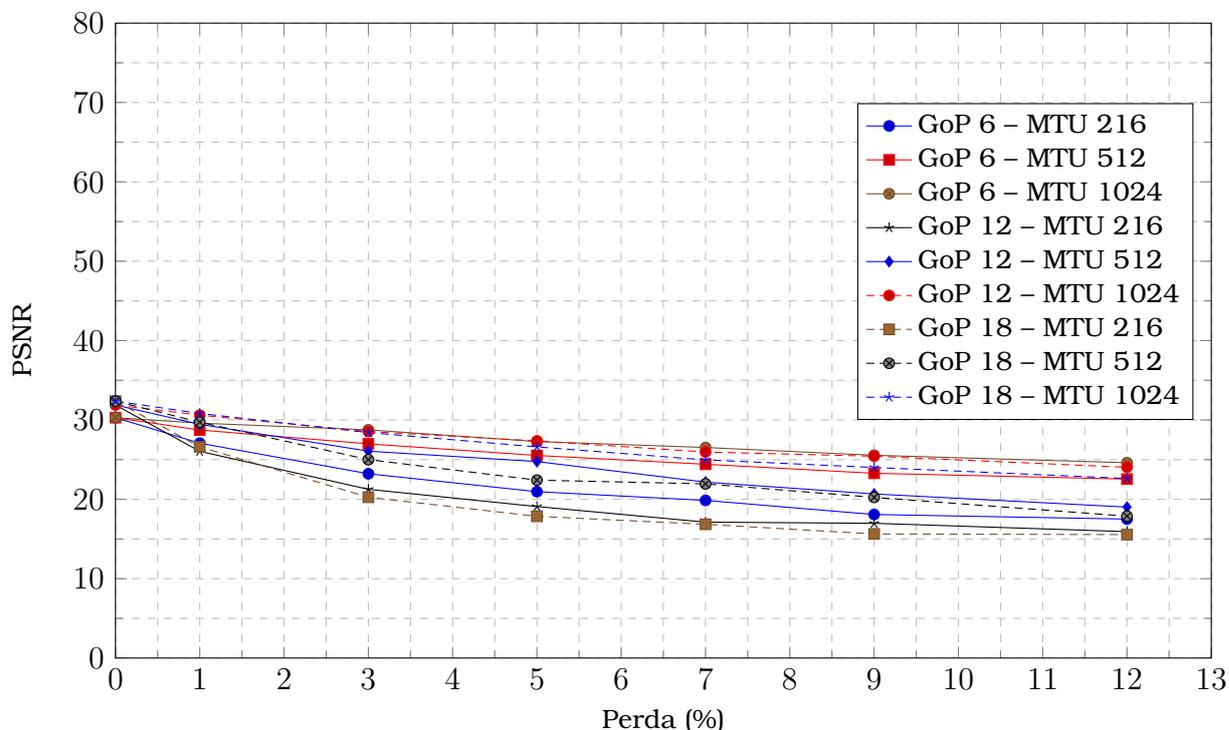


Figura 45: Influência da perda de pacotes no vídeo Tempete com *bitrate* igual a 500 kbps de acordo com o GoP e MTU.

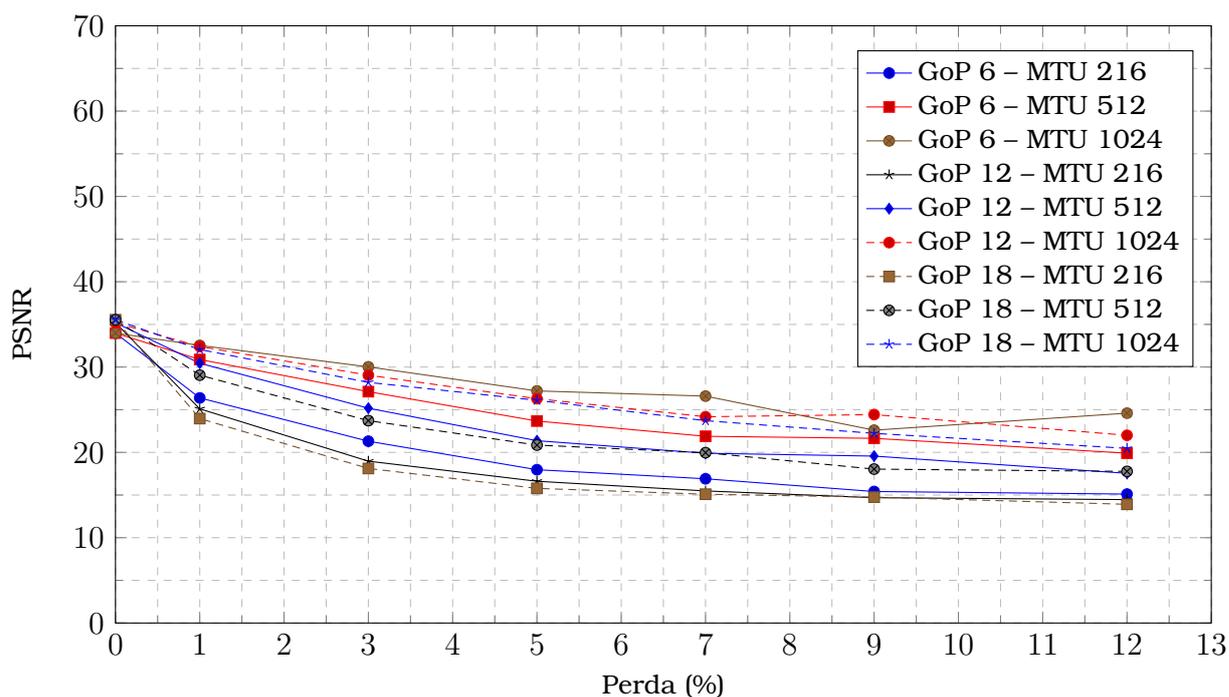


Figura 46: Influência da perda de pacotes no vídeo Tempete com *bitrate* igual a 1000 kbps de acordo com o GoP e MTU.

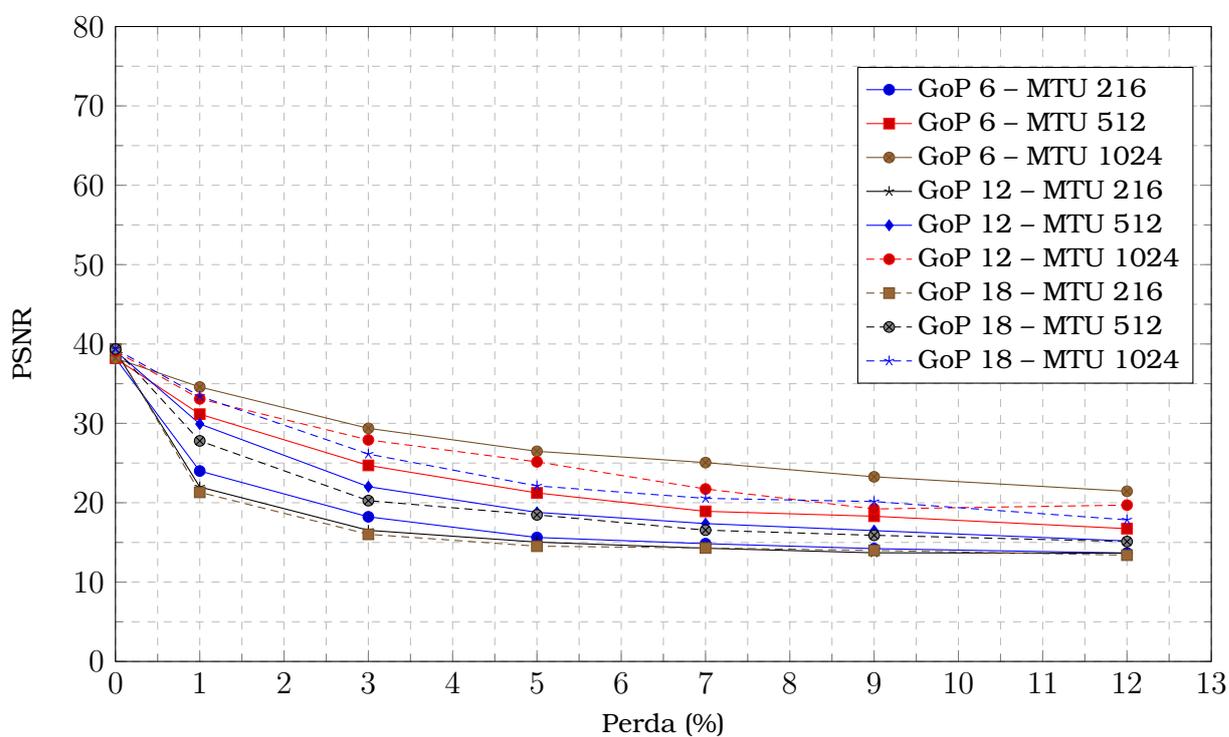


Figura 47: Influência da perda de pacotes no vídeo Tempete com *bitrate* igual a 2000 kbps de acordo com o GoP e MTU.

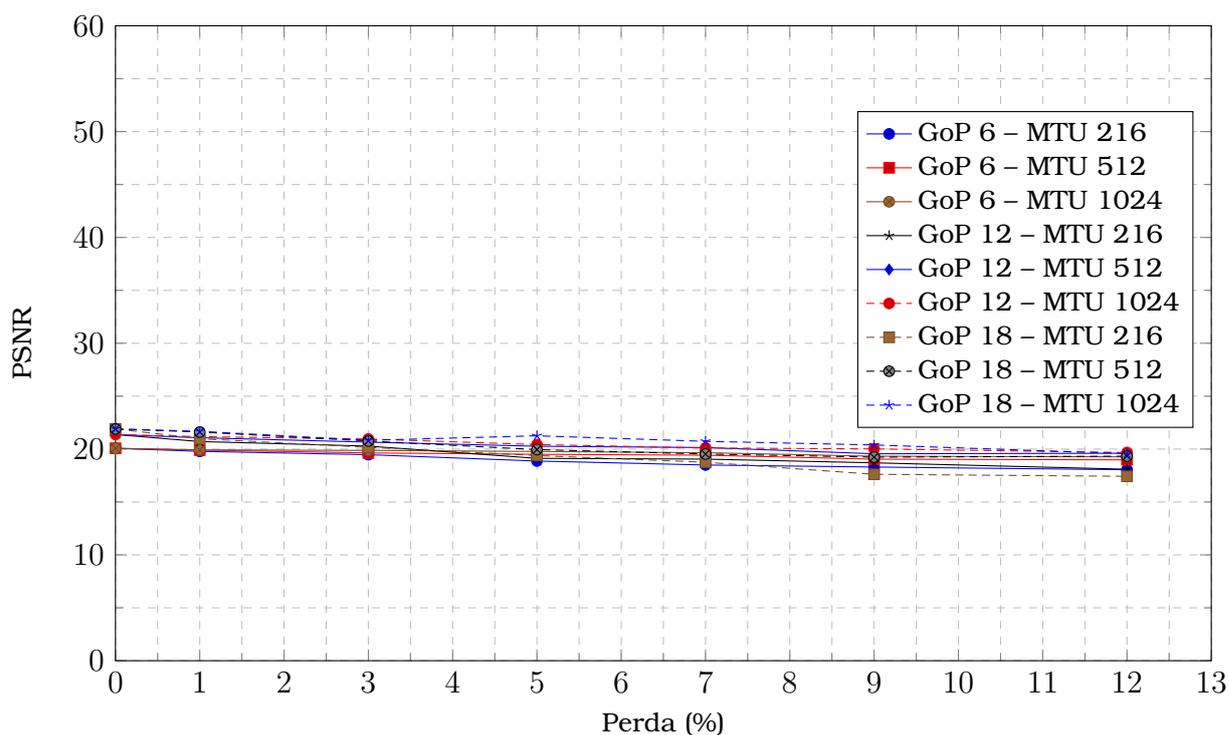


Figura 48: Influência da perda de pacotes no vídeo Flower com *bitrate* igual a 100 kbps de acordo com o GoP e MTU.

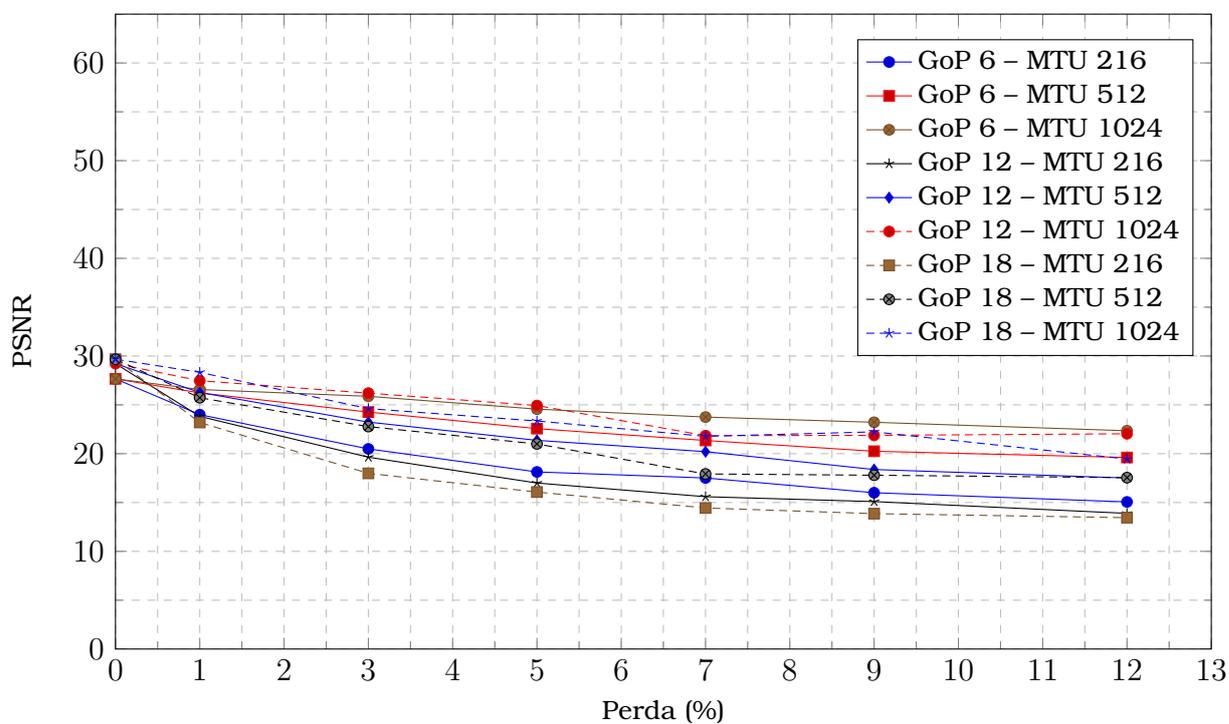


Figura 49: Influência da perda de pacotes no vídeo Flower com *bitrate* igual a 500 kbps de acordo com o GoP e MTU.

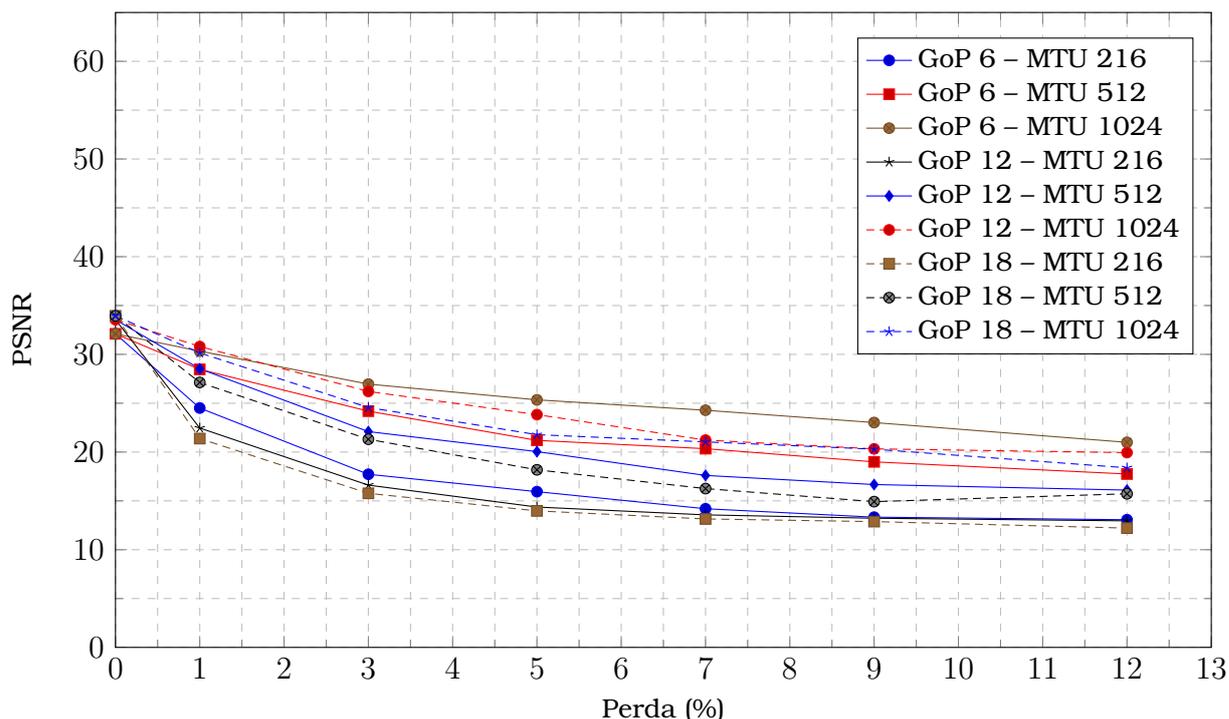


Figura 50: Influência da perda de pacotes no vídeo Flower com *bitrate* igual a 1000 kbps de acordo com o GoP e MTU.

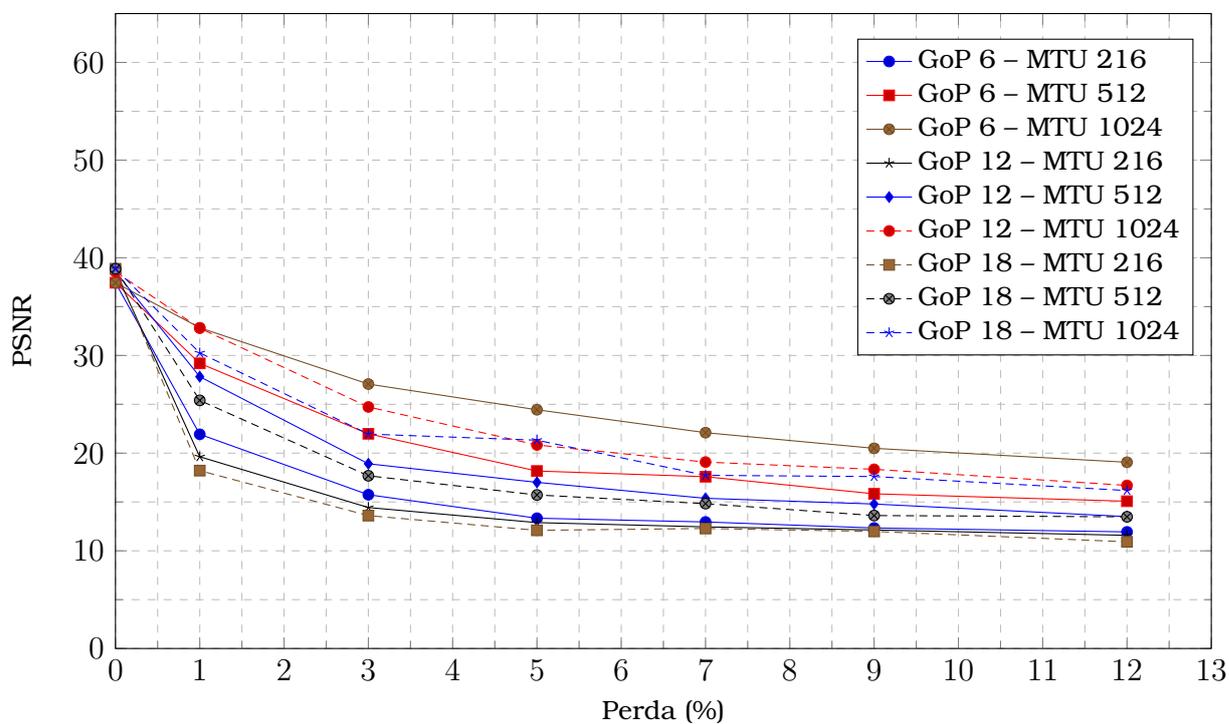


Figura 51: Influência da perda de pacotes no vídeo Flower com *bitrate* igual a 2000 kbps de acordo com o GoP e MTU.

Tabela 28: Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Akiyo de acordo com o *bitrate*, GoP e MTU.

Bitrate (kbps)	GoP	MTU	Dados trafegados na rede (MB)
100	6	216	67,93
		512	60,17
		1024	57,56
	12	216	64,39
		512	56,83
		1024	54,46
	18	216	62,92
		512	55,42
		1024	53,02
500	6	216	330,00
		512	278,14
		1024	263,58
	12	216	300,00
		512	266,22
		1024	252,66
	18	216	300,00
		512	258,90
		1024	245,71
1000	6	216	660,00
		512	570,00
		1024	540,00
	12	216	630,00
		512	540,00
		1024	510,00
	18	216	600,00
		512	510,00
		1024	480,00
2000	6	216	1260,00
		512	1080,00
		1024	1020,00
	12	216	1200,00
		512	1050,00
		1024	990,00
	18	216	1170,00
		512	1020,00
		1024	960,00

Tabela 29: Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Tempete de acordo com o *bitrate*, GoP e MTU.

Bitrate (kbps)	GoP	MTU	Dados trafegados na rede (MB)
100	6	216	70,21
		512	61,95
		1024	57,80
	12	216	68,17
		512	59,87
		1024	56,78
	18	216	67,39
		512	59,38
		1024	85,96
500	6	216	336,48
		512	290,78
		1024	275,10
	12	216	325,23
		512	282,20
		1024	266,34
	18	216	319,61
		512	277,35
		1024	262,05
1000	6	216	684,00
		512	576,00
		1024	540,00
	12	216	648,00
		512	565,20
		1024	536,40
	18	216	640,80
		512	540,00
		1024	525,60
2000	6	216	1332,00
		512	1152,00
		1024	1080,00
	12	216	1296,00
		512	1116,00
		1024	1044,00
	18	216	1260,00
		512	1080,00
		1024	1036,80

Tabela 30: Quantidade de dados trafegados durante sessenta minutos de transmissão do vídeo Flower de acordo com o *bitrate*, GoP e MTU.

Bitrate (kbps)	GoP	MTU	Dados trafegados na rede (MB)
100	6	216	69,05
		512	59,41
		1024	58,61
	12	216	67,32
		512	58,22
		1024	57,23
	18	216	65,81
		512	89,30
		1024	55,79
500	6	216	340,66
		512	294,15
		1024	278,68
	12	216	321,40
		512	277,84
		1024	263,32
	18	216	309,69
		512	267,61
		1024	253,48
1000	6	216	684,00
		512	576,00
		1024	576,00
	12	216	648,00
		512	540,00
		1024	540,00
	18	216	612,00
		512	536,40
		1024	504,00
2000	6	216	1332,00
		512	1152,00
		1024	1080,00
	12	216	1260,00
		512	1080,00
		1024	1044,00
	18	216	1224,00
		512	1062,00
		1024	1008,00

Pesquisa sobre a qualidade de vídeos provenientes da Internet transmitidos para dispositivos móveis

Durante o período de Março/2014 a Junho/2014, fez-se uma pesquisa sobre a qualidade de vídeos transmitidos para dispositivos móveis com o intuito de saber quais dispositivos os entrevistados costumam utilizar para assistir vídeos provenientes da Internet, além de perguntar se tinham disponibilidade para avaliar a qualidade de alguns vídeos reproduzidos em dispositivos móveis.

Obteve-se um total de 100 questionários respondidos, dos quais 71% eram do gênero masculino e 29% eram do gênero feminino. Na Fig. 52 é exposta a distribuição dos entrevistados de acordo com a escolaridade.

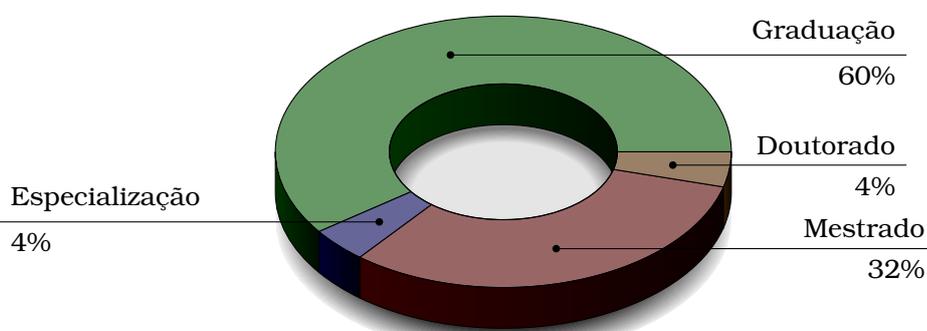


Figura 52: Distribuição dos entrevistados de acordo com a escolaridade.

Todos os entrevistados assistem vídeos provenientes da Internet. De acordo com a Fig. 53, a maioria prefere utilizar computador/laptop e *smartphone* para essa finalidade. Poucas pessoas utilizam o *tablet* para isso. O tempo médio, em minutos ao dia, em que elas assistem vídeos provenientes da Internet no computador/laptop é de 97,87. Quando *smartphones* e *tablets* são usados, o tempo médio cai para quase metade: 58,55 e 47,94 para *smartphone* e *tablet*, respectivamente. Apesar da televisão ser a terceira opção utilizada para essa finalidade, os entrevistados reservam, em média, 96,97 minutos do seu tempo para isso.

Quando perguntados sobre o tempo de espera para normalizar a reprodução do vídeo após uma falha, houve uma variação por tipo de serviço utilizado, conforme mostrada na Fig. 54. Quando o serviço é gratuito, a maioria dos entrevistados esperaria até 10 segundos (34%). Por outro lado, quando o serviço é pago, a maioria ou não esperaria (27%) ou esperaria até 5 segundos (26%).

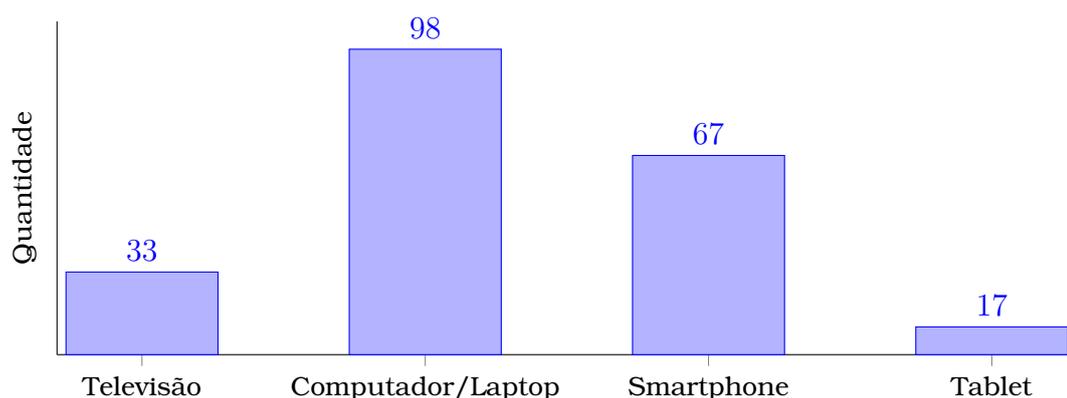


Figura 53: Distribuição de dispositivos utilizados para assistir vídeos provenientes da Internet.

Esses resultados mostram que as pessoas esperariam entre 5 e 10 segundos para normalizar o serviço, independente do tipo de serviço. Esse fato foi utilizado pelo sistema de gerenciamento da QoE proposto neste trabalho, como indicador de período em que a qualidade do vídeo reproduzido no dispositivo móvel seria avaliada.

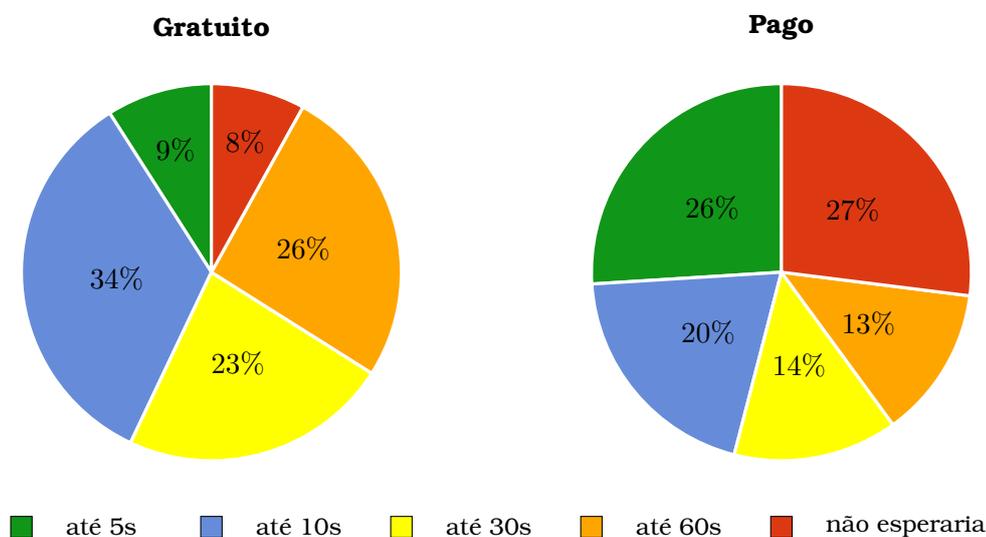


Figura 54: Distribuição de espera pela normalidade de exibição do vídeo após o surgimento de alguma falha por tipo de serviço.

Das 100 pessoas entrevistadas, 51 disponibilizaram do seu tempo para participar das avaliações subjetivas. Destas, somente 44 pessoas estavam aptas para avaliar a qualidade dos vídeos. Seis pessoas não passaram no teste de acuidade visual e uma pessoa não passou no teste de deficiência visual para cores, a qual foi aconselhada a procurar um médico para tratamento.

As 44 pessoas aptas foram divididas em dois grupos: metade avaliaram a qualidade dos vídeos reproduzidos no *smartphone* e a outra metade no *tablet*. Na primeira fase, elas deram notas de qualidade nos vídeos distorcidos sem introdução de erros, somente variação do *bitrate*. Após 24 horas (segunda fase), pelo menos, as pessoas voltaram para avaliar a qualidade de vídeos distorcidos com introdução de erros (perda de pacotes gerados pelo modelo Gilbert-Elliot). Nesse caso, as pessoas que deram nota para os vídeos exibidos em *smartphone* avaliaram a qualidade dos vídeos reproduzidos no *tablet* e vice-versa.

Avaliação subjetiva de vídeos sem introdução de erros

Durante as avaliações subjetivas, coletaram-se 3696 amostras (1848 para cada tipo de dispositivo móvel) para os vídeos distorcidos sem introdução de erros (apenas variação do *bitrate*). Um histograma dessas amostras de acordo com o índice de qualidade (MOS) para cada tipo de dispositivo móvel é mostrado na Fig. 55.

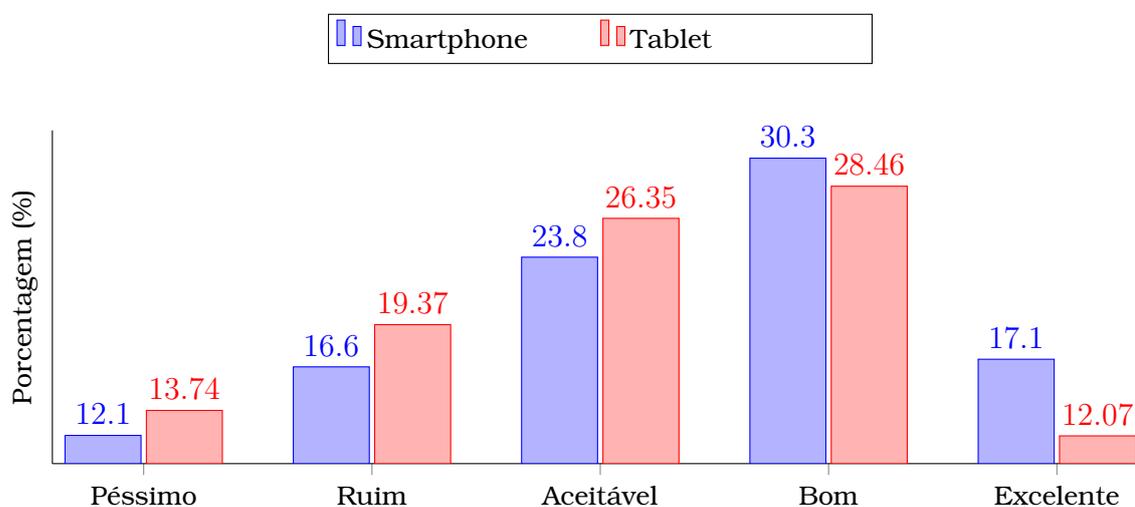


Figura 55: Histograma das avaliações subjetivas coletadas com vídeos distorcidos (sem introdução de erros).

Observa-se que, entre os índices péssimo e aceitável, o *tablet* possui uma quantidade maior de amostras e, entre os índices bom e excelente, o *smartphone* possui mais amostras. Isso é decorrência da qualidade percebida pelos voluntários diante do tamanho da tela do dispositivo móvel, que fica evidenciada na Fig. 56.

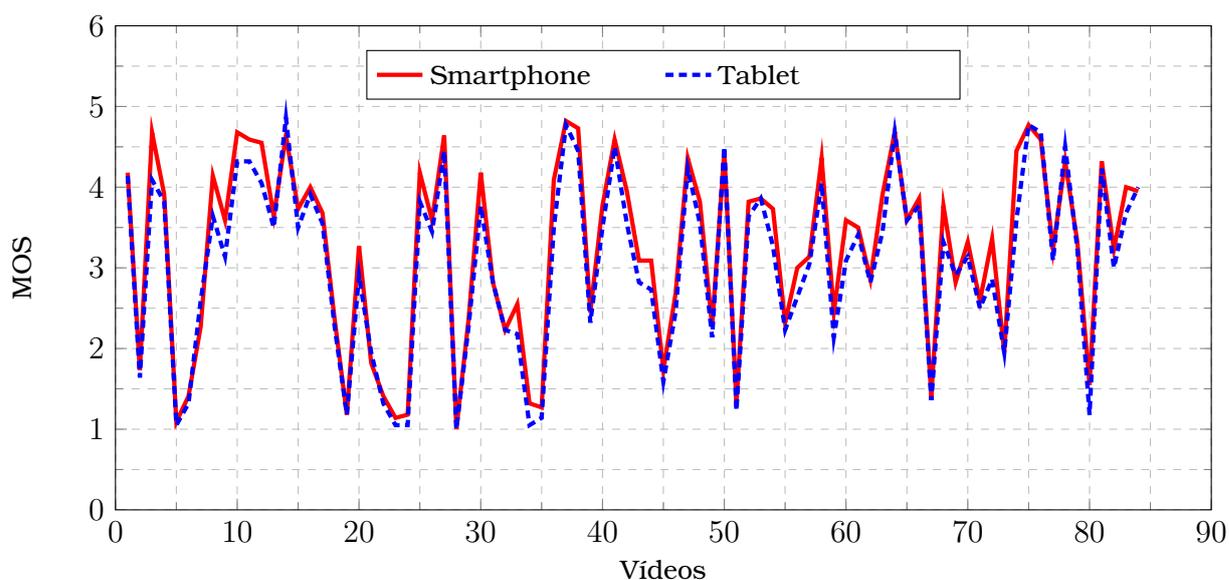


Figura 56: Distribuição das avaliações subjetivas em cada vídeo distorcido (sem introdução de erros).

Entre o sétimo e o décimo terceiro item da sequência reproduzida nos dispositivos móveis, os vídeos possuem resolução entre 240p e 480p e o *bitrate* varia entre 150 kbps e 2000 kbps. Quando esses vídeos foram exibidos no *tablet*, apesar de um *bitrate* igual a 2000 kbps, a qualidade percebida foi baixa quando comparada com o mesmo vídeo exibido no *smartphone*. Isso acontece devido ao processo de interpolação do conteúdo de vídeos com resolução baixa (240p, por exemplo) quando são reproduzidos em um dispositivo com tela grande. Assim, tem-se a impressão que o conteúdo está pouco nítido quando exibido no *tablet*, por exemplo.

O desvio padrão observado a partir das notas coletadas no *smartphone* variou entre 0 e 0,63 e para o *tablet* entre 0 e 0,68, mostrando que houve pouca variação das notas com relação a média final. O intervalo de confiança dessas notas para cada tipo de dispositivo móvel é mostrado nas Figs. 57 e 58.

Avaliação subjetiva de vídeos com introdução de erros

Nas avaliações subjetivas dos vídeos distorcidos com introdução de erros gerados a partir do modelo Gilbert-Elliot foram coletadas 3168 amostras (1584 para cada tipo de dispositivo móvel). Um histograma dessas amostras de acordo com o índice de qualidade (MOS) para cada tipo de dispositivo móvel é mostrado na Fig. 59.

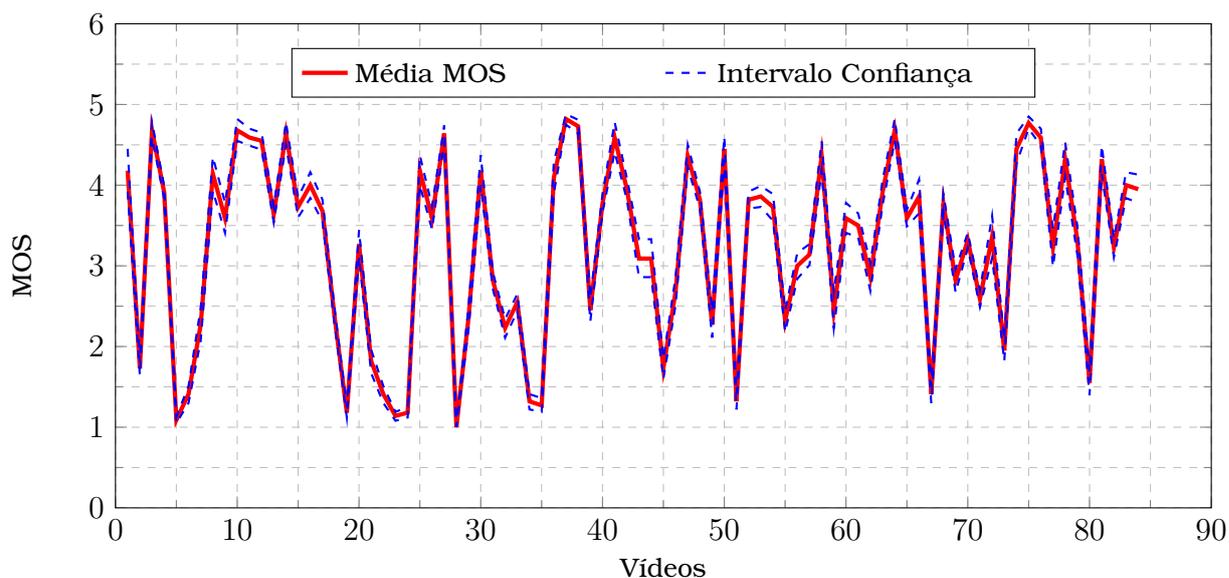


Figura 57: Distribuição das avaliações subjetivas dos vídeos distorcidos (sem introdução de erros) coletadas no *smartphone* com o seu intervalo de confiança.

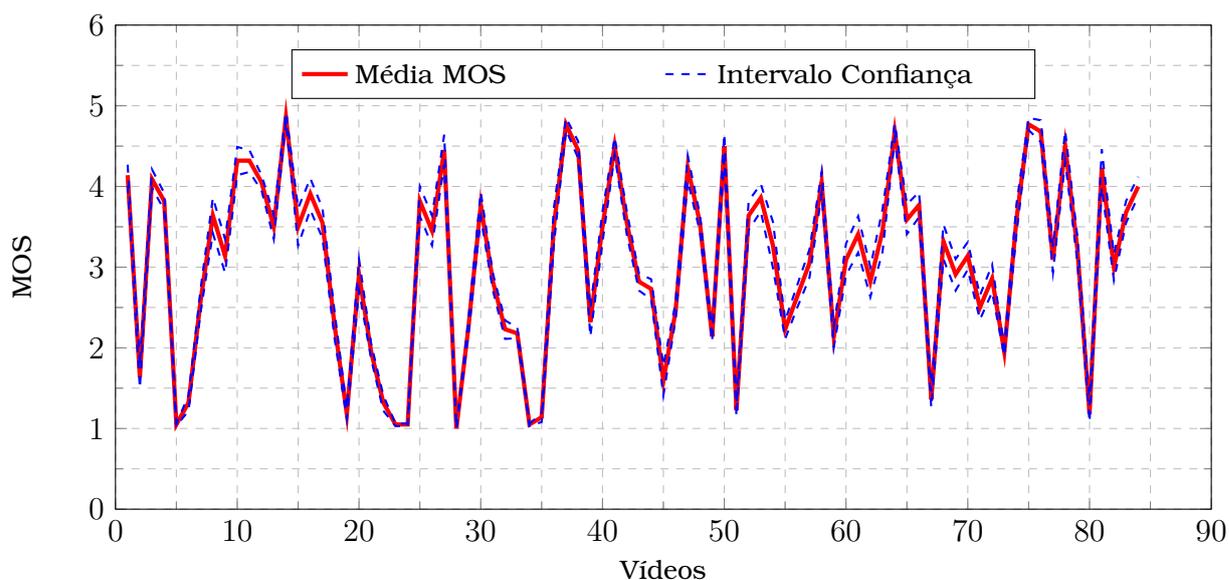


Figura 58: Distribuição das avaliações subjetivas dos vídeos distorcidos (sem introdução de erros) coletadas no *tablet* com o seu intervalo de confiança.

Nesse caso, não obtiveram-se notas para o índice Excelente, pois todos os vídeos distorcidos possuíam alguma falha (perda de pacote). Observa-se que, entre os índices péssimo e ruim, o *smartphone* possui uma quantidade maior de amostras e, entre os índices aceitável e bom, o *tablet* possui mais amostras. Aqui, também o tamanho da tela do dispositivo móvel influenciou na percepção da qualidade do vídeo, conforme ilustrado na Fig. 60.

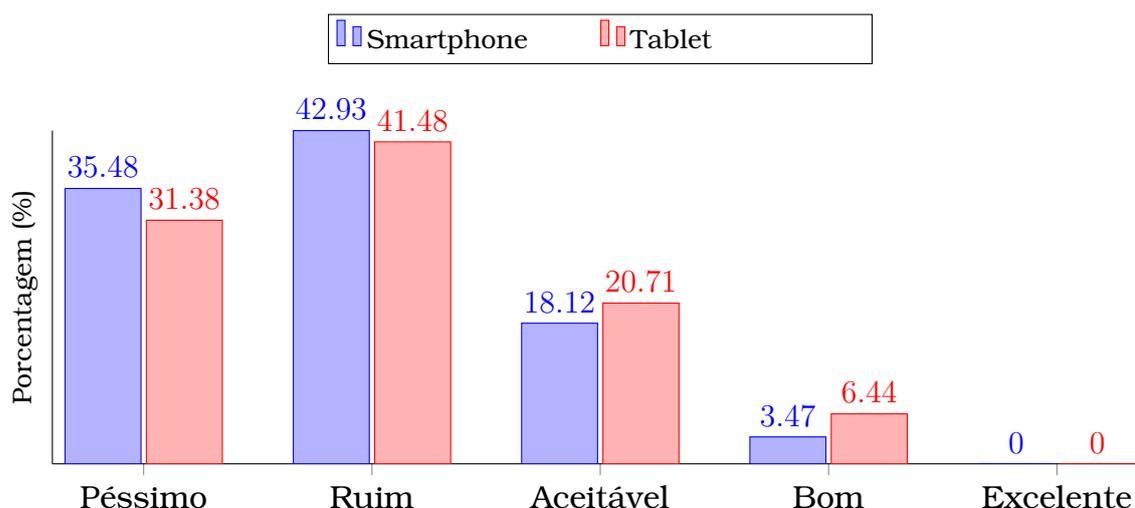


Figura 59: Histograma das avaliações subjetivas coletadas com vídeos distorcidos (com introdução de erros).

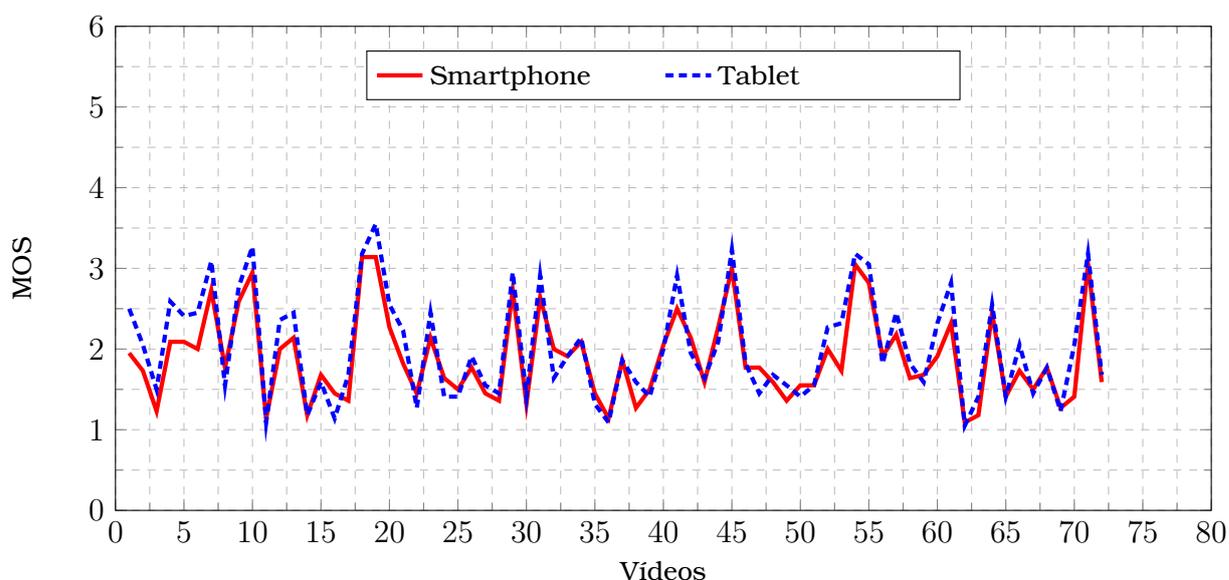


Figura 60: Distribuição das avaliações subjetivas em cada vídeo distorcido (com introdução de erros).

Analisando a distribuição da notas, constatou-se que a percepção da qualidade dos vídeos exibidos no *smartphone* é maior quando comparada com os mesmos reproduzidos no *tablet*, principalmente quando a resolução e o *bitrate* são superiores a 480p e 1000 kbps, respectivamente. Isso acontece pelo fato de que esses vídeos possuem muitas informações geradas e que, quando são perdidas durante a transmissão, são exibidas muitas falhas durante a reprodução.

O desvio padrão investigado a partir das notas coletadas no *smartphone* variou entre 0,09 e 0,75 e para o *tablet* entre 0,05 e 0,81, apresentando que houve pouca variação das notas com relação a média final. O intervalo de confiança dessas notas para cada tipo de dispositivo móvel é mostrado nas Figs. 61 e 62.

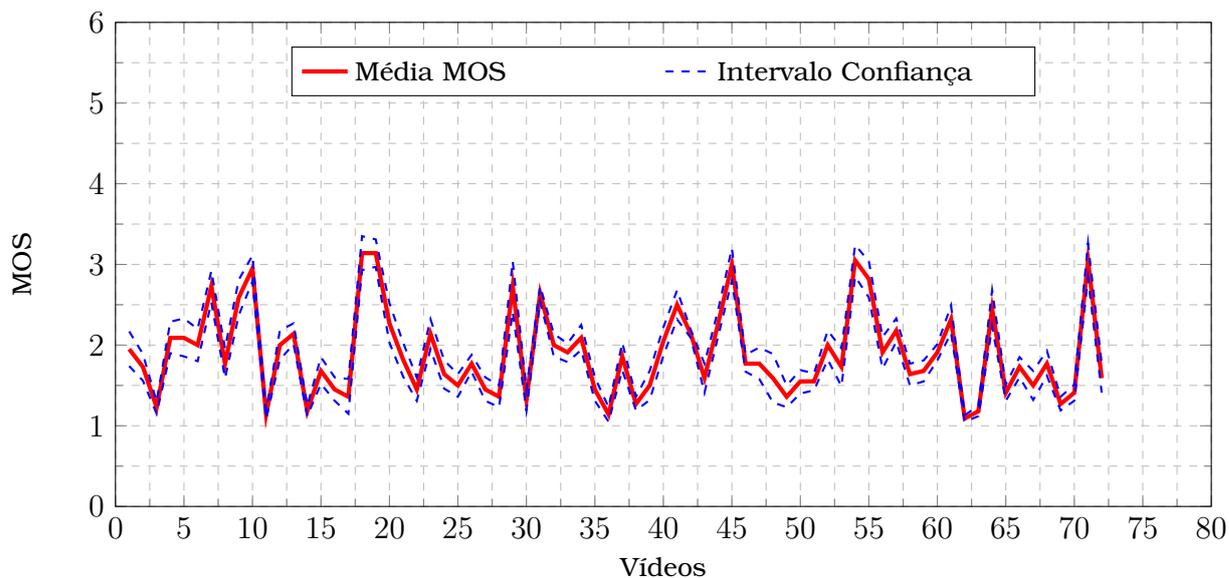


Figura 61: Distribuição das avaliações subjetivas dos vídeos distorcido (com introdução de erros) coletadas no smartphone com o seu intervalo de confiança.

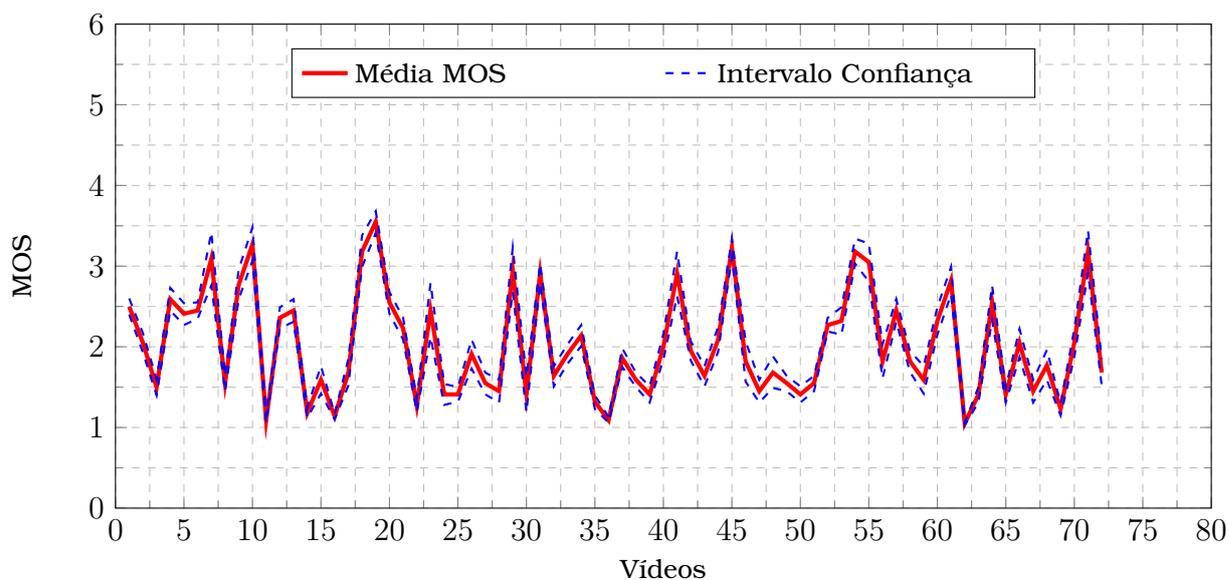


Figura 62: Distribuição das avaliações subjetivas dos vídeos distorcido (com introdução de erros) coletadas no tablet com o seu intervalo de confiança.

Modelo de predição de qualidade de vídeo sem erros introduzidos durante a transmissão do conteúdo

Com o conjunto de dados criado a partir das informações coletadas durante as simulações e as avaliações subjetivas, iniciou-se a elaboração dos preditores de qualidade. Neste primeiro momento, não foram levadas em consideração informações sobre perda de pacotes durante a transmissão do conteúdo para o dispositivo móvel. Assim, foram consideradas as seguintes informações:

- *Informações sobre o formato de vídeo do conteúdo:* (1) Resolução Espacial (Largura - pixels); (2) Resolução Espacial (Altura - pixels); (3) Informação Espacial; (4) Informação Temporal; (5) Bitrate; (6) Quantidade de Quadros do Tipo I; (7) Quantidade de Quadros do Tipo P; (8) Quantidade de Quadros do Tipo B; (9) % Quadros do Tipo I; (10) % Quadros do Tipo P; (11) % Quadros do Tipo B; (12) Média QP dos Quadros do Tipo I; (13) Média QP dos Quadros do Tipo P; (14) Média QP dos Quadros do Tipo B; (15) Média QP; (16) % intra block of size 8x8; e (17) % inter block of size 8x8.
- *Informações sobre a transmissão de vídeo do conteúdo:* (18) Quantidade de Pacotes I-Frame; (19) Quantidade de Pacotes P-Frame; (20) Quantidade de Pacotes B-Frame; (21) Quantidade Total de Pacotes; (22) Média Pacotes I-Frame por segundo; (23) Média Pacotes P-Frame por segundo; (24) Média Pacotes B-Frame por segundo; (25) Média Pacotes por segundo; (26) Tam. dos Quadros do Tipo I (bytes); (27) Tam. dos Quadros do Tipo P (bytes); (28) Tam. dos Quadros do Tipo B (bytes); (29) Média Tam. dos Quadros do Tipo I (bytes); (30) Média Tam. dos Quadros do Tipo P (bytes); (31) Média Tam. dos Quadros do Tipo B (bytes); e (32) Média Throughput (Mbits).
- *Informações sobre o dispositivo móvel:* (33) Resolução (Largura - pixels); (34) Resolução (Altura - pixels); e (35) Tamanho da Tela.

Com essas informações foram criados grupos de atributos com o intuito de selecionar os melhores de forma que os valores do preditor de qualidade dos vídeos distorcidos fossem mais próximos das notas obtidas nas avaliações subjetivas. Esses grupos são apresentados na Tabela 31.

Tabela 31: Agrupamento dos atributos utilizados para criação do preditor de qualidade de vídeos distorcidos sem introdução de erros.

Grupo	Atributos	Descrição
G1	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35	Todas os atributos
G2	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	Somente informações sobre o vídeo
G3	18 19 20 21 22 23 24 25 26 27 28 29 30 31 32	Somente informações sobre os pacotes
G4	33 34 35	Somente informações sobre o dispositivo móvel
G5 a G40	Vários atributos	Matriz de correlação entre os atributos
G41	3 4 12 13 14 15 16 17 18 22 26 28 29 30 31 33 34 35	Algoritmo BestFirst (Backward) do Weka
G42	3 12 13 14 15 16 17 22 25 29 31 35	Algoritmo BestFirst (Forward) do Weka
G43	3 12 13 14 15 16 17 22 29 31 32 35	Algoritmo BestFirst (Bi-Diretional) do Weka
G44	3 12 13 14 15 16 17 22 25 29 31 35	Algoritmo GreedyStepwise do Weka
G45	3 12 13 14 15 16 17 22 25 29 31 35	Algoritmo LinearForwardSelection (Forward selection) do Weka
G46	3 12 13 14 15 16 17 22 29 31 32 35	Algoritmo LinearForwardSelection (Floating forward selection) do Weka
G47	3 12 13 14 15 17 18 29 30 31 34	Algoritmo ScatterSearchV1 (Greed Combination) do Weka
G48	12 18 29 30 31 32 33	Algoritmo ScatterSearchV1 (Reduced Greed Combination) do Weka
G49	3 12 13 14 15 16 17 22 25 29 31 35	Algoritmo SubsetSizeForwardSelection do Weka
G50	7 Atributos	PCA

Em seguida, realizaram-se vários experimentos utilizando os algoritmos mostrados na Tabela 26 mais as técnicas de árvore de decisão e vizinho mais próximo. Esses algoritmos estão contidos na metodologia objetiva sem referência. Os melhores resultados obtidos com o uso da amostragem *random subsampling* para esses algoritmos são mostrados na Tabela 32. Nela são apresentados o nome do algoritmo, a forma de tratamento de dados (sem ou com o proposto neste trabalho), o melhor grupo de atributo e as métricas MSE e coeficiente de Pearson. Todos os algoritmos foram executados na mesma base de treinamento e teste criados para atender o *random subsampling*. O melhor resultado está destacado em azul. Na Fig. 63 são exibidos os resultados obtidos quando é utilizado a melhor combinação quando comparado com o MOS.

Tabela 32: Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem *random subsampling* para os vídeos distorcidos sem introdução de erros.

Algoritmo	Tratamento dos Dados	Atributos	MSE	Pearson
Regressão linear	Sem	G35	0,089 ± 0,027	0,966 ± 0,010
	Proposto neste trabalho	G36	0,056 ± 0,011	0,978 ± 0,005
Regressão linear-logarítmico	Sem	G23	0,073 ± 0,010	0,972 ± 0,006
	Proposto neste trabalho	G40	0,076 ± 0,019	0,970 ± 0,008
Regressão linear-logarítmico-linear	Sem	G2	0,094 ± 0,019	0,962 ± 0,008
	Proposto neste trabalho	G41	0,051 ± 0,007	0,980 ± 0,004
Regressão linear-logarítmico-logarítmico	Sem	G40	0,064 ± 0,012	0,975 ± 0,006
	Proposto neste trabalho	G41	0,057 ± 0,010	0,978 ± 0,004
Regressão logística	Sem	G37	0,060 ± 0,014	0,971 ± 0,007
	Proposto neste trabalho	G1	0,067 ± 0,013	0,974 ± 0,006
Árvore de decisão	Sem	G18	0,073 ± 0,017	0,971 ± 0,007
	Proposto neste trabalho	G36	0,079 ± 0,012	0,969 ± 0,005
Vizinho mais próximo (distância euclidiana)	Sem	G5	0,090 ± 0,019	0,964 ± 0,010
	Proposto neste trabalho	G1	0,094 ± 0,018	0,963 ± 0,008

Para efeito de comparação, também foram executados esses mesmos experimentos com algoritmos SSIM, MSSIM, VSNR, VIF, VIFP, UQI, IFC, NQM, WNSR, SNR e PSNR. Os resultados obtidos quando comparados com o MOS estão expostos nas Figs. 64 – 66.

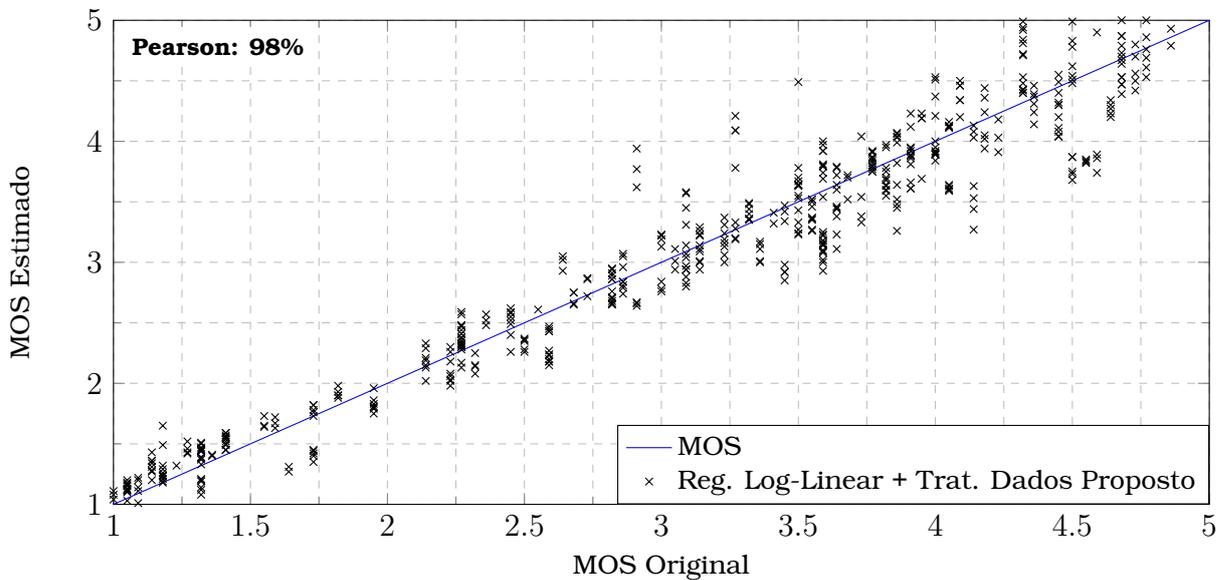


Figura 63: Resultados obtidos para regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos sem introdução de erros com o uso do método de amostragem *random subsampling*.

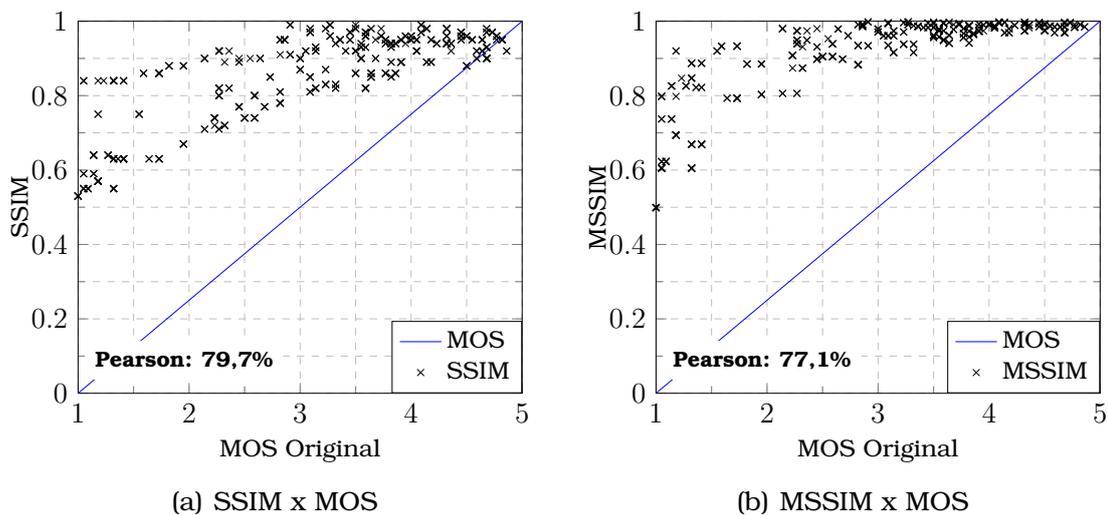


Figura 64: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos sem introdução de erros (1 de 3).

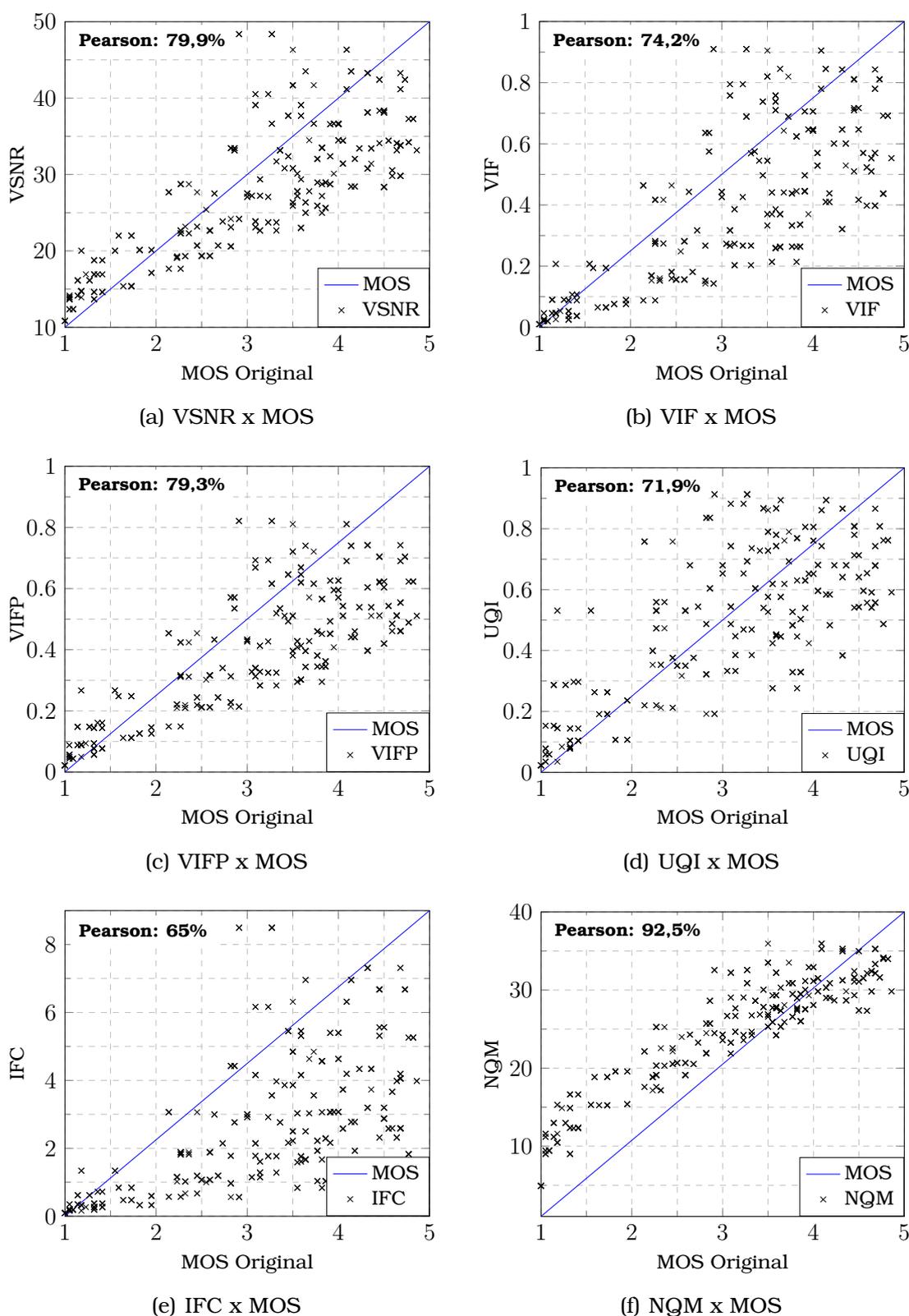


Figura 65: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos sem introdução de erros (2 de 3).

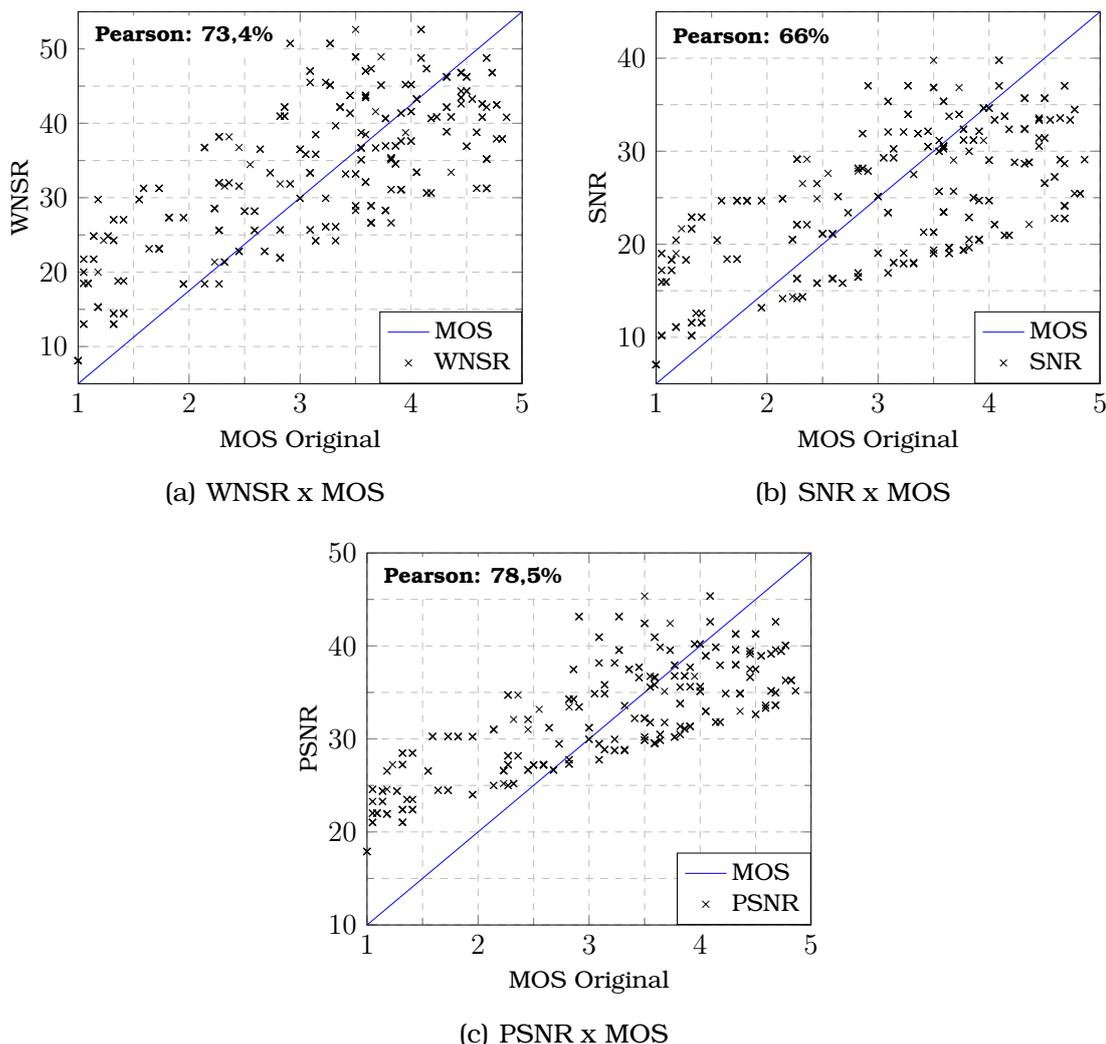


Figura 66: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos sem introdução de erros (3 de 3).

A partir dos resultados expostos nas Figs. 63 – 66, observou-se que o algoritmo de regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho possuiu maior correlação com as notas obtidas nas avaliações subjetivas (98%), enquanto que NQM obteve 92,5%. Além disso, o índice de erro (MSE) foi, em média, 0,051.

Quando o método de amostragem *cross validation* foi aplicado, novamente a melhor combinação foi o uso do algoritmo regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho, conforme mostrado destacado em azul na Tabela 33. Ele obteve, em média, erro igual a 0,045 e 98,2% de correlação de Pearson, enquanto o NQM, que foi o melhor algoritmo

da metodologia objetiva com referência completa, obteve 92,4% de correlação com as notas obtidas nas avaliações subjetivas.

Tabela 33: Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem *cross validation* para os vídeos distorcidos sem introdução de erros.

Algoritmo	Tratamento dos Dados	Atributos	MSE	Pearson
Regressão linear	Sem	G1	$0,062 \pm 0,012$	$0,976 \pm 0,004$
	Proposto neste trabalho	G1	$0,047 \pm 0,018$	$0,981 \pm 0,006$
Regressão linear-logarítmico	Sem	G1	$0,053 \pm 0,015$	$0,979 \pm 0,007$
	Proposto neste trabalho	G1	$0,059 \pm 0,020$	$0,977 \pm 0,007$
Regressão linear-logarítmico-linear	Sem	G1	$0,065 \pm 0,018$	$0,975 \pm 0,005$
	Proposto neste trabalho	G41	$0,045 \pm 0,017$	$0,982 \pm 0,006$
Regressão logarítmico-logarítmico	Sem	G1	$0,055 \pm 0,020$	$0,978 \pm 0,008$
	Proposto neste trabalho	G1	$0,051 \pm 0,018$	$0,980 \pm 0,006$
Regressão logística	Sem	G37	$0,051 \pm 0,017$	$0,980 \pm 0,007$
	Proposto neste trabalho	G1	$0,049 \pm 0,021$	$0,981 \pm 0,007$
Árvore de decisão	Sem	G5	$0,063 \pm 0,022$	$0,975 \pm 0,009$
	Proposto neste trabalho	G41	$0,065 \pm 0,025$	$0,973 \pm 0,010$
Vizinho mais próximo (distância euclidiana)	Sem	G41	$0,069 \pm 0,025$	$0,973 \pm 0,010$
	Proposto neste trabalho	G41	$0,064 \pm 0,023$	$0,975 \pm 0,009$

O resultado obtido quando é empregado o algoritmo regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho com o método de amostragem *cross validation* é exibido na Fig. 67. Também os algoritmos SSIM, MSSIM, VSNR, VIF, VIFP, UQI, IFC, NQM, WNSR, SNR e PSNR foram executados nesse mesmo método de amostragem e os resultados obtidos quando comparados com o MOS estão apresentados nas Figs. 68 – 70.

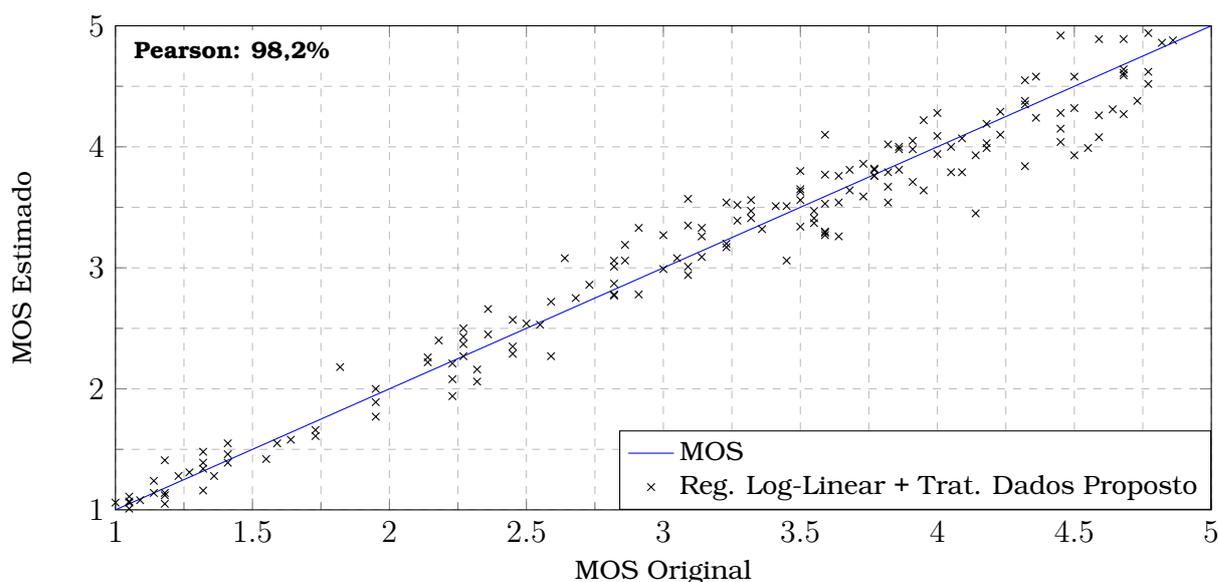


Figura 67: Resultados obtidos para regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos sem introdução de erros com o uso do método de amostragem *cross validation*.

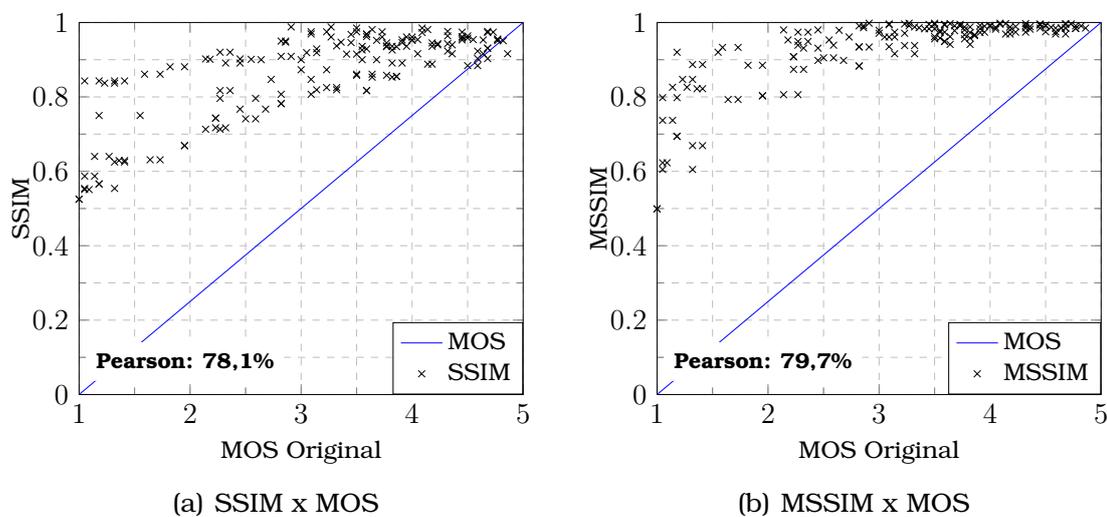


Figura 68: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *cross validation* para os vídeos distorcidos sem introdução de erros (1 de 3).

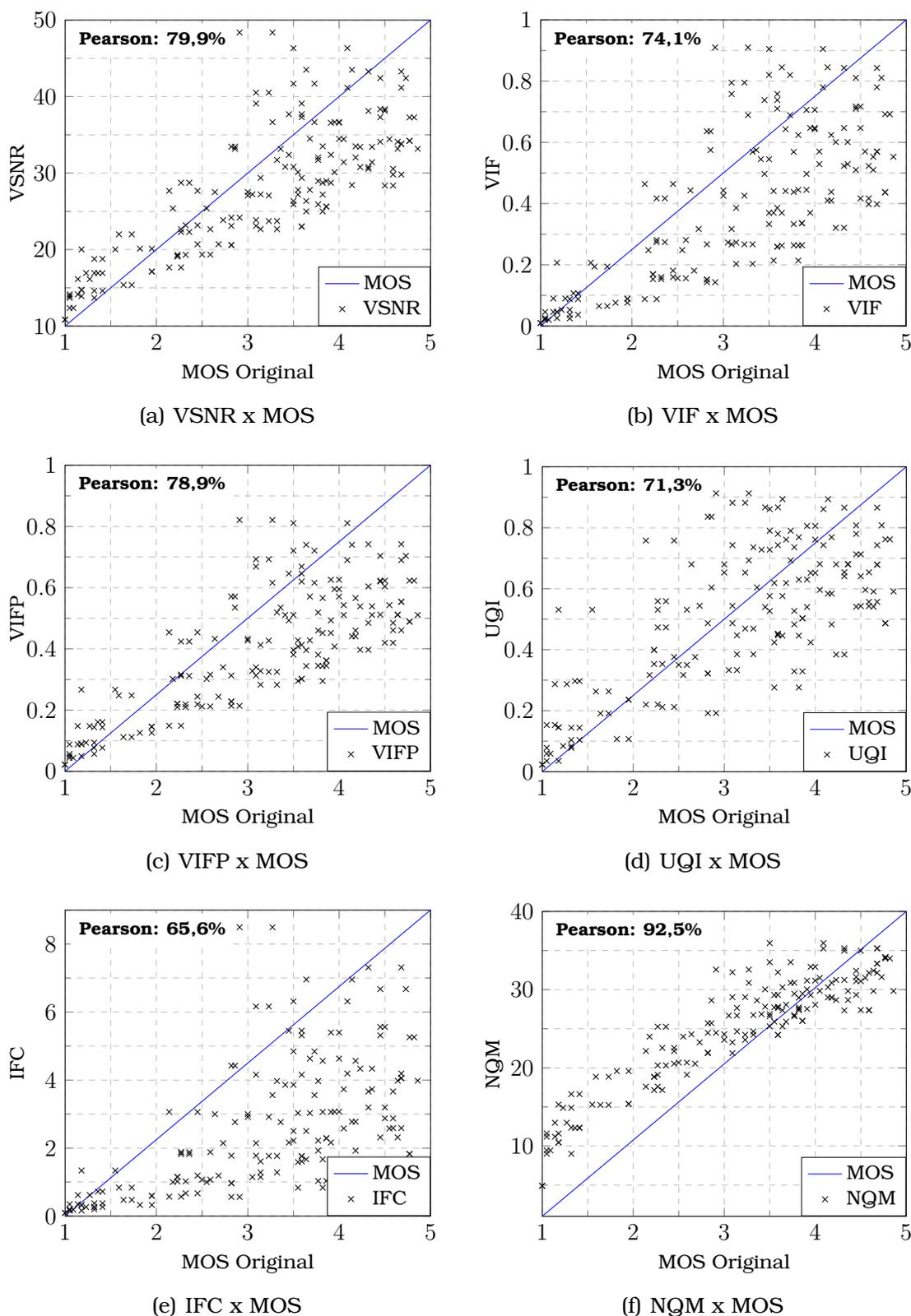


Figura 69: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *cross validation* para os vídeos distorcidos sem introdução de erros (2 de 3).

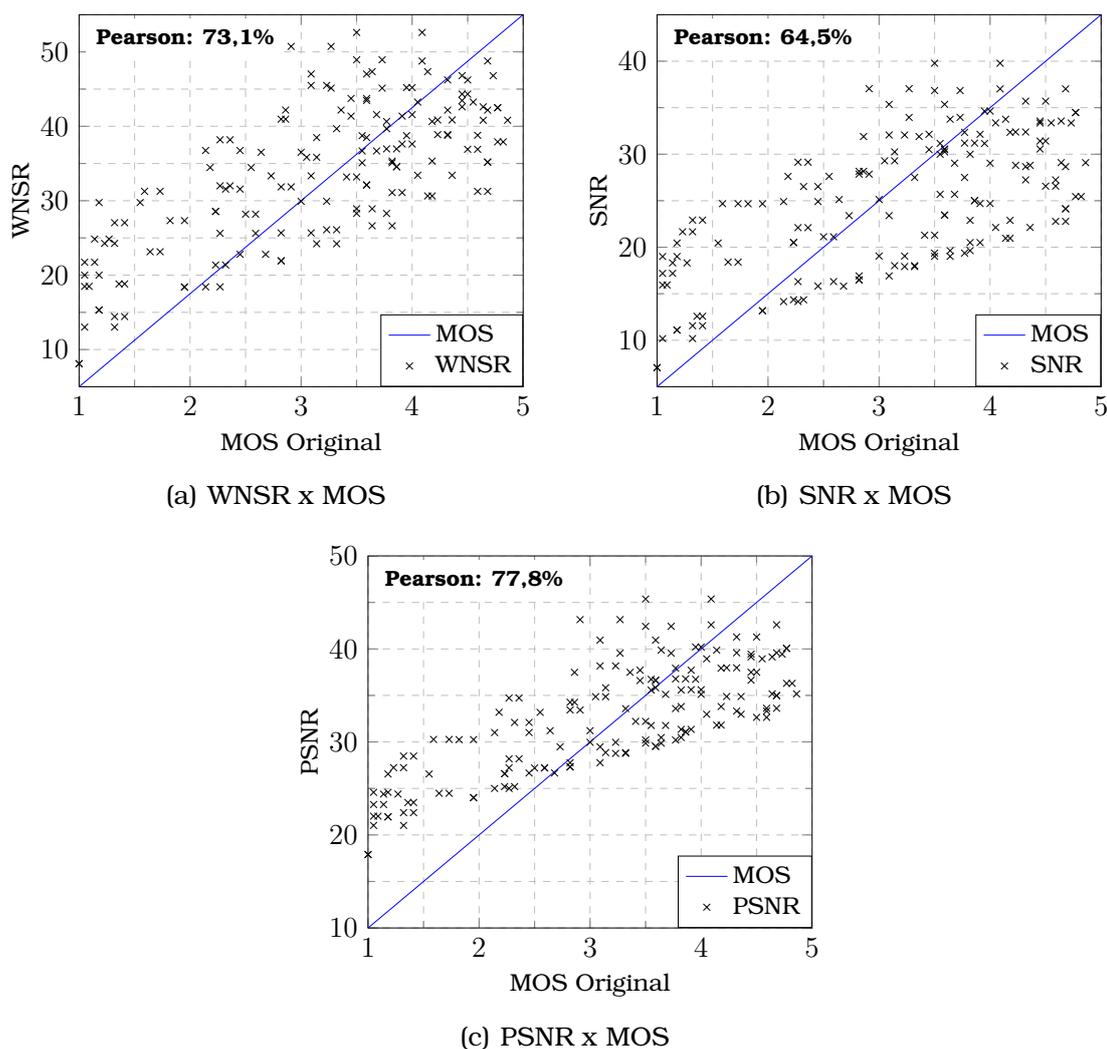


Figura 70: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *cross validation* para os vídeos distorcidos sem introdução de erros (3 de 3).

Portanto, o algoritmo definido para atuar no preditor que escolhe o vídeo no primeiro momento que o usuário seleciona o conteúdo desejado ou quando não possui perda de pacotes durante a transmissão do conteúdo foi a regressão logarítmico-linear mais o tratamento de dados proposto neste trabalho com o grupo de atributos G41.

Modelo de predição de qualidade de vídeo com erros introduzidos durante a transmissão do conteúdo

Neste segundo momento foram levadas em consideração informações sobre os vídeos distorcidos com introdução de erros (perda de pacotes). Os atributos empregados foram:

- *Informações sobre o formato de vídeo do conteúdo:* (1) Resolução Espacial (Largura - pixels); (2) Resolução Espacial (Altura - pixels); (3) Informação Espacial; (4) Informação Temporal; (5) Bitrate; (6) Quantidade de Quadros do Tipo I; (7) Quantidade de Quadros do Tipo P; (8) Quantidade de Quadros do Tipo B; (9) % Quadros do Tipo I; (10) % Quadros do Tipo P; (11) % Quadros do Tipo B; (12) Média QP dos Quadros do Tipo I; (13) Média QP dos Quadros do Tipo P; (14) Média QP dos Quadros do Tipo B; (15) Média QP; (16) % intra block of size 8x8; e (17) % inter block of size 8x8.
- *Informações sobre a transmissão de vídeo do conteúdo:* (18) Quantidade de Pacotes I-Frame; (19) Quantidade de Pacotes P-Frame; (20) Quantidade de Pacotes B-Frame; (21) Quantidade Total de Pacotes; (22) Média Pacotes I-Frame por segundo; (23) Média Pacotes P-Frame por segundo; (24) Média Pacotes B-Frame por segundo; (25) Média Pacotes por segundo; (26) Tam. dos Quadros do Tipo I (bytes); (27) Tam. dos Quadros do Tipo P (bytes); (28) Tam. dos Quadros do Tipo B (bytes); (29) Média Tam. dos Quadros do Tipo I (bytes); (30) Média Tam. dos Quadros do Tipo P (bytes); (31) Média Tam. dos Quadros do Tipo B (bytes); (32) Quantidade de Pacotes do Tipo I Recebida; (33) Quantidade de Pacotes do Tipo P Recebida; (34) Quantidade de Pacotes do Tipo B Recebida; (35) Quantidade Total de Pacotes Recebida; (36) Quantidade de Pacotes do Tipo I Perdida; (37) Quantidade de Pacotes do Tipo P Perdida; (38) Quantidade de Pacotes do Tipo B Perdida; (39) Quantidade Total de Pacotes Perdida; (40) % de Perda Pacotes I-Frame; (41) % de Perda Pacotes P-Frame; (42) % de Perda Pacotes B-Frame; (43) % de Perda Pacotes Geral; e (44) Média Throughput (Mbits);
- *Informações sobre o dispositivo móvel:* (45) Resolução (Largura - pixels); (46) Resolução (Altura - pixels); e (47) Tamanho da Tela.

Em seguida foram criados grupos de atributos com o propósito de escolher os melhores atributos de forma que os valores do preditor de qualidade dos vídeos distorcidos com introdução de erros fossem mais próximo das notas obtidas nas avaliações subjetivas. Na Tabela 34 são descritos esses grupos.

Tabela 34: Agrupamento dos atributos utilizados para criação do preditor de qualidade de vídeos distorcidos com introdução de erros.

Grupo	Atributos	Descrição
G1	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Todas os atributos
G2	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	Somente informações sobre o vídeo
G3	18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44	Somente informações sobre os pacotes
G4	45 46 47	Somente informações sobre o dispositivo móvel
G5 a G2500	Vários atributos	Matriz de correlação entre os atributos
G2501	3 4 5 13 17 18 22 23 24 25 26 28 29 36 37 38 39 40 41 42 43 44 45 46 47	Algoritmo BestFirst (Backward) do Weka
G2502	9 10 11 36 43 44 45	Algoritmo BestFirst (Forward) do Weka
G2503	9 10 11 36 43 44 45	Algoritmo BestFirst (Bi-Diretional) do Weka
G2504	9 10 11 36 43 44 45	Algoritmo GreedyStepwise do Weka
G2505	9 10 11 36 43 44 45	Algoritmo LinearForwardSelection (Forward selection) do Weka
G2506	9 10 11 36 43 44 45	Algoritmo LinearForwardSelection (Floating forward selection) do Weka
G2507	36 43 44 45	Algoritmo ScatterSearchV1 (Greed Combination) do Weka
G2508	6 40 43 44 45	Algoritmo ScatterSearchV1 (Reduced Greed Combination) do Weka
G2509	9 10 11 36 43 44 45	Algoritmo SubsetSizeForwardSelection do Weka
G2510	7 Atributos	PCA

O mesmo procedimento realizado para a elaboração do preditor para vídeos com introdução de erros foi aplicado para os vídeos com introdução de erros. Não foram empregados os algoritmos de regressão linear-logarítmico e logarítmico-logarítmico porque haviam valores iguais a zero (como % de perda de pacotes para I-Frame, por exemplo). Na Tabela 35 são mostrados os melhores resultados obtidos com o uso da amostragem *random subsampling* e o melhor resultado está destacado em azul. Na Fig. 71 são exibidos os resultados obtidos para a melhor combinação quando comparado com o MOS.

Tabela 35: Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem *random subsampling* para vídeos com introdução de erros.

Algoritmo	Tratamento dos Dados	Atributos	MSE	PEARSON
Regressão linear	Sem	G2508	$0,127 \pm 0,007$	$0,813 \pm 0,041$
	Proposto neste trabalho	G1958	$0,072 \pm 0,014$	$0,893 \pm 0,037$
Regressão logarítmico-linear	Sem	G2508	$0,110 \pm 0,011$	$0,845 \pm 0,038$
	Proposto neste trabalho	G2068	$0,069 \pm 0,009$	$0,900 \pm 0,028$
Regressão logística	Sem	G2508	$0,119 \pm 0,007$	$0,825 \pm 0,041$
	Proposto neste trabalho	G1958	$0,069 \pm 0,014$	$0,897 \pm 0,036$
Árvore de decisão	Sem	G2242	$0,088 \pm 0,029$	$0,875 \pm 0,029$
	Proposto neste trabalho	G773	$0,086 \pm 0,027$	$0,874 \pm 0,052$
Vizinho mais próximo (distância Euclidiana)	Sem	G2484	$0,139 \pm 0,029$	$0,802 \pm 0,044$
	Proposto neste trabalho	G1477	$0,092 \pm 0,015$	$0,869 \pm 0,029$

Também foram executados esses mesmos experimentos com algoritmos SSIM, MSSIM, VSNR, VIF, VIFP, UQI, IFC, NQM, WNSR, SNR e PSNR quando utilizados os vídeos com introdução de erros. Os resultados obtidos comparados com o MOS estão expostos nas Figs. 72 – 74.

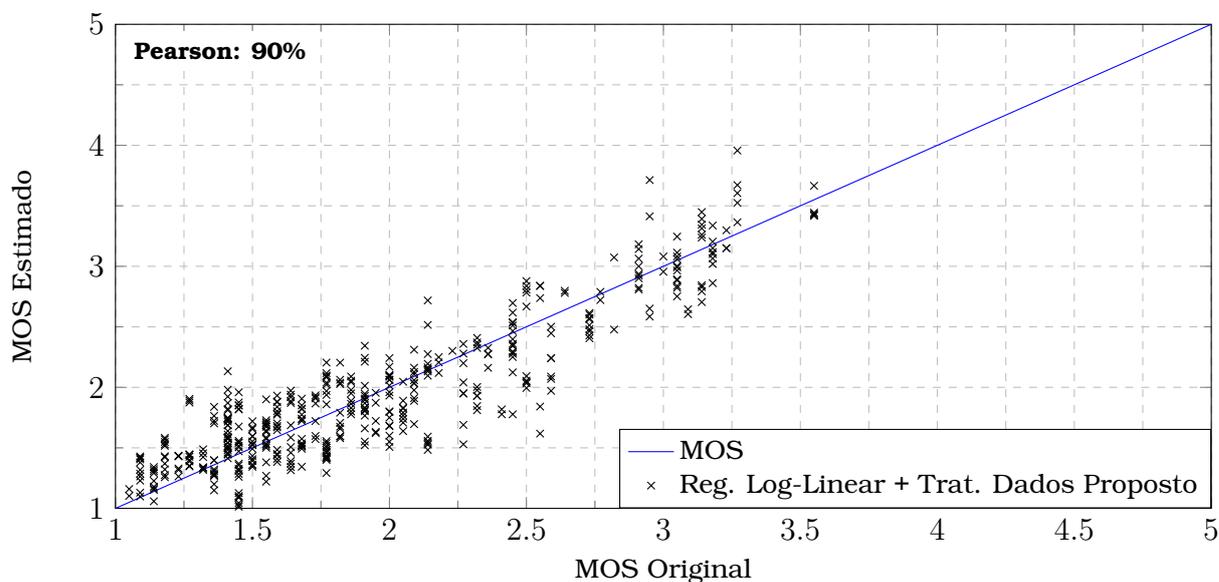


Figura 71: Resultados obtidos para regressão logarítmico-linear mais tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos com introdução de erros com o uso do método de amostragem *random subsampling*.

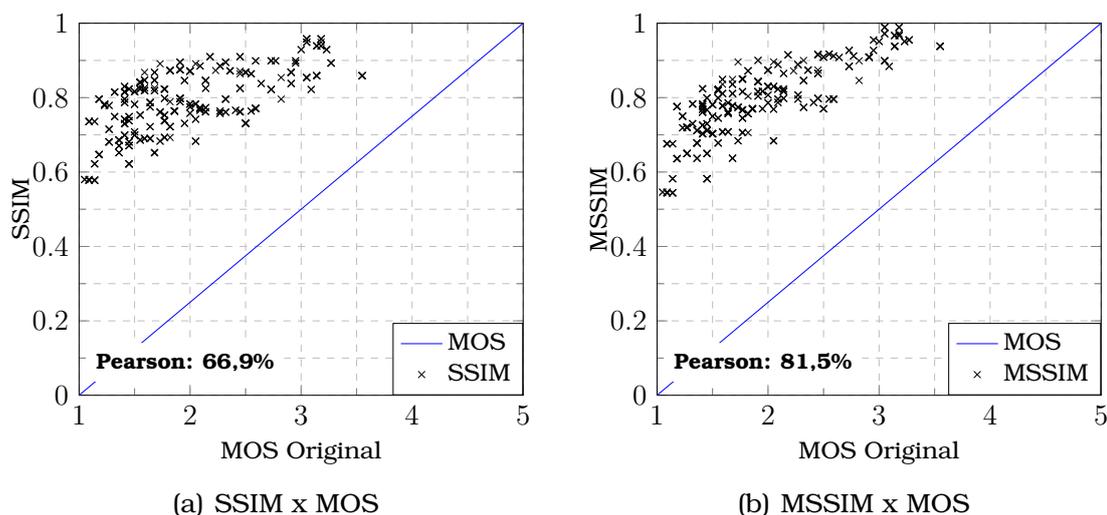


Figura 72: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos com introdução de erros (1 de 3).

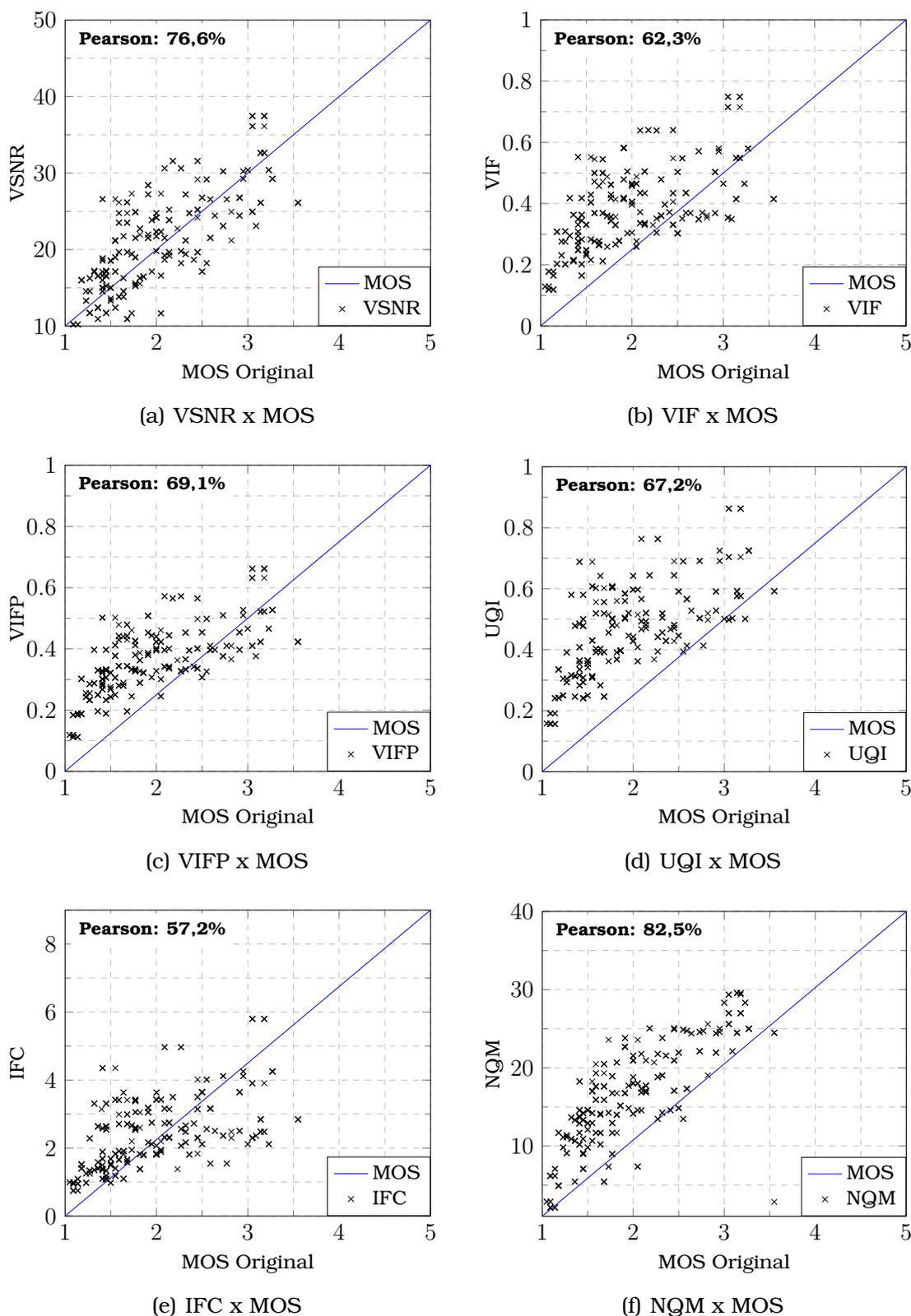


Figura 73: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos com introdução de erros (2 de 3).

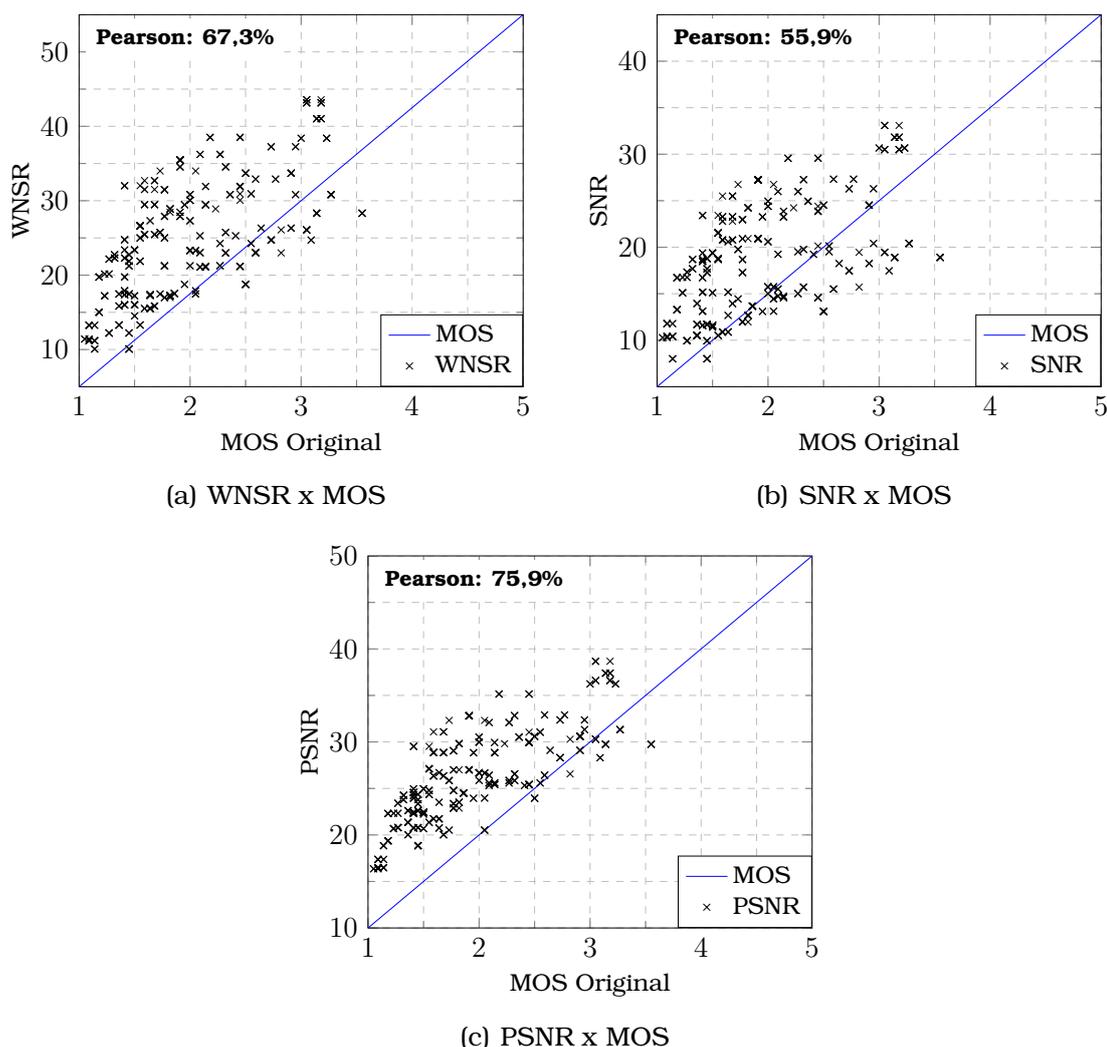


Figura 74: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *random subsampling* para os vídeos distorcidos com introdução de erros (3 de 3).

A partir dos resultados expostos nas Figs. 71 – 74, observou-se que o algoritmo de regressão logarítmico-linear mais o tratamento dos dados proposto neste trabalho atingiu maior correlação com as notas obtidas nas avaliações subjetivas (90%), enquanto que NQM obteve 82,5%. Além disso, o índice de erro (MSE) foi, em média, 0,069.

Quando o método de amostragem *cross validation* foi utilizado, novamente a melhor combinação foi o uso do algoritmo árvore de decisão mais o tratamento dos dados proposto neste trabalho, conforme mostrado destacado em azul na Tabela 36. Ele obteve, em média, erro igual a 0,068 e 92,9% de correlação de Pearson (ver Fig. 75), enquanto o NQM obteve 82,6%. Nas

Figs. 76 e 77 são apresentados os resultados dos algoritmos SSIM, MSSIM, VSNR, VIF, VIFP, UQI, IFC, NQM, WNSR, SNR e PSNR para esse cenário.

Tabela 36: Melhores resultados obtidos com os algoritmos da metodologia objetiva sem referência utilizando a amostragem *cross validation* para vídeos com introdução de erros.

Algoritmo	Tratamento dos Dados	Atributos	MSE	PEARSON
Regressão linear	Sem	G2508	$0,123 \pm 0,027$	$0,820 \pm 0,053$
	Proposto neste trabalho	G2054	$0,069 \pm 0,018$	$0,903 \pm 0,029$
Regressão logarítmico-linear	Sem	G1477	$0,109 \pm 0,034$	$0,837 \pm 0,066$
	Proposto neste trabalho	G1970	$0,065 \pm 0,020$	$0,913 \pm 0,027$
Regressão logística	Sem	G2508	$0,116 \pm 0,025$	$0,829 \pm 0,050$
	Proposto neste trabalho	G1970	$0,066 \pm 0,017$	$0,908 \pm 0,027$
Árvore de decisão	Sem	G888	$0,057 \pm 0,001$	$0,929 \pm 0,035$
	Proposto neste trabalho	G795	$0,058 \pm 0,022$	$0,929 \pm 0,029$
Vizinho mais próximo (distância euclidiana)	Sem	G3	$0,084 \pm 0,050$	$0,895 \pm 0,070$
	Proposto neste trabalho	G2062	$0,063 \pm 0,027$	$0,920 \pm 0,045$

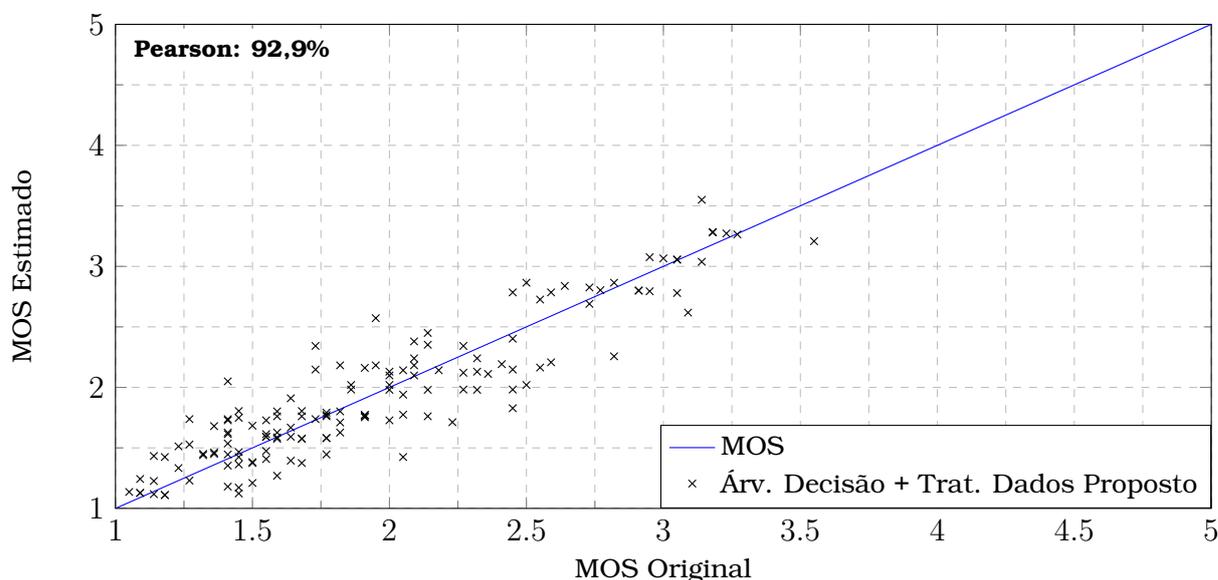


Figura 75: Resultados obtidos para regressão logarítmico-linear mais tratamento de dados proposto neste trabalho como preditor de qualidade de vídeos com introdução de erros com o uso do método de amostragem *cross validation*.

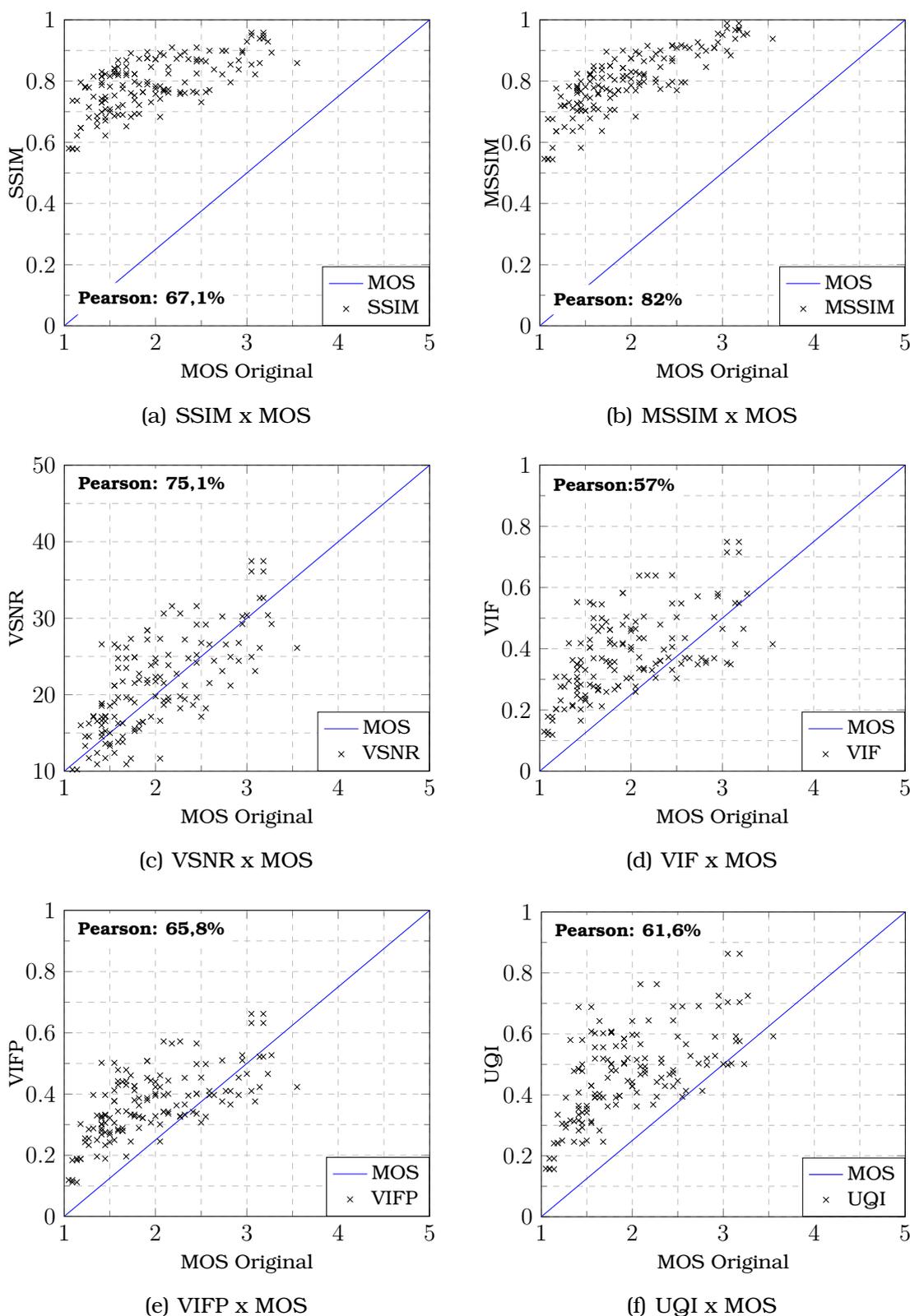


Figura 76: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *cross validation* para os vídeos distorcidos com introdução de erros (1 de 2).

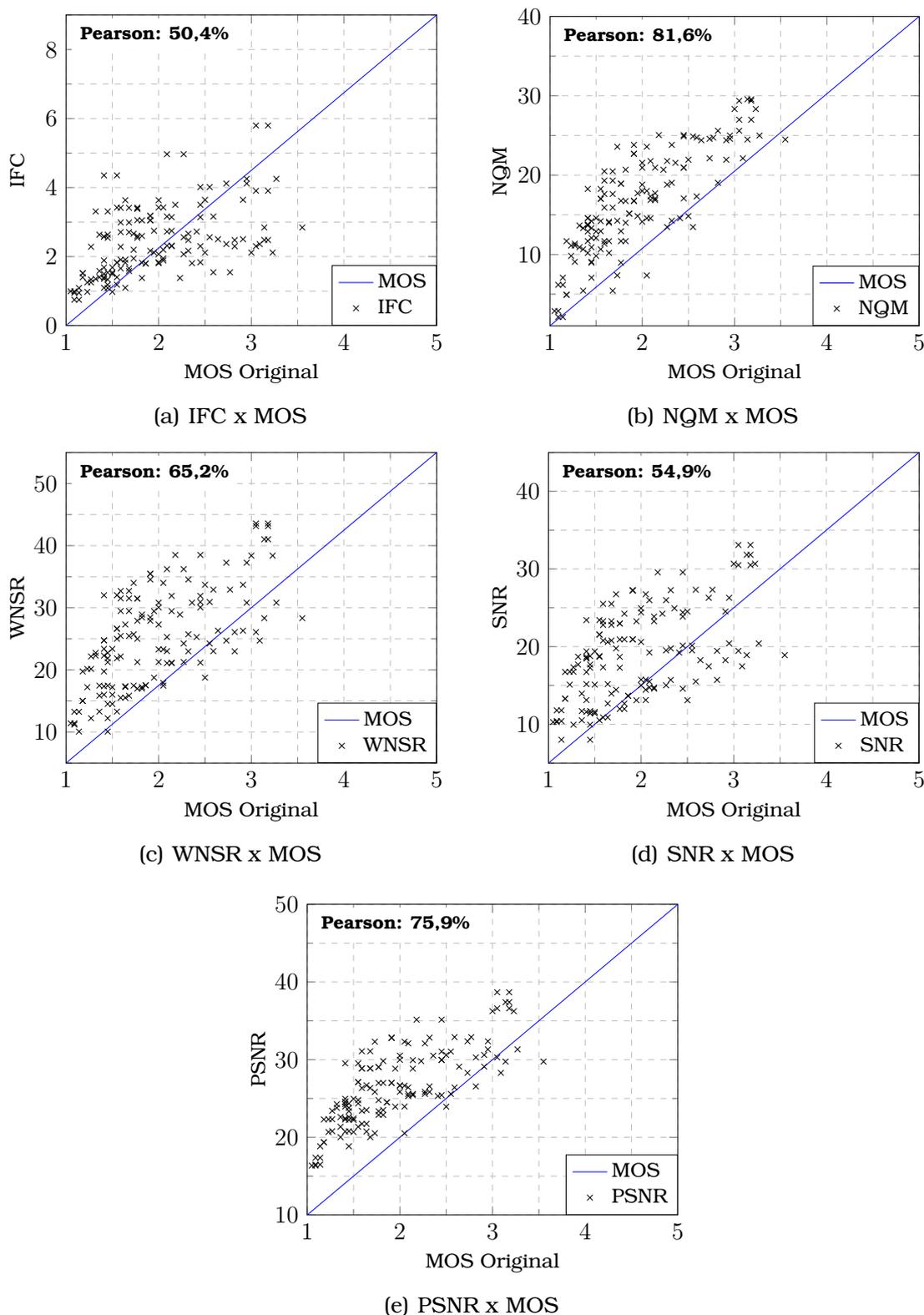


Figura 77: Comparativo de qualidade obtida pelos algoritmos da metodologia objetiva com referência completa executados no método de amostragem *cross validation* para os vídeos distorcidos com introdução de erros (2 de 2).

Portanto, caso o método de amostragem *random subsampling* seja empregado, opta-se por utilizar o algoritmo regressão logarítmico-linear mais tratamento dos dados proposto neste trabalho para G2068 (atributos 2, 3, 4, 5, 15, 16, 17, 21, 25, 29, 30, 31, 32, 33, 34, 36, 37, 38, 43, 44, 45, 46 e 47). Quando for usado o método de amostragem *cross validation*, adota-se o algoritmo árvore de decisão mais tratamento dos dados proposto neste trabalho para G795 (atributos 2, 3, 4, 5, 15, 16, 17, 18, 19, 20, 29, 30, 31, 32, 33, 34, 36, 37, 38, 44, 45, 46 e 47).

Sistema desenvolvido para o gerenciamento de QoE durante a transmissão de vídeos para dispositivos móveis

Para demonstrar o funcionamento do sistema descrito na Fig. 31, foram empregados os resultados obtidos com o uso da amostragem *random subsampling*. Assim, a relação entre os componentes que compõem o sistema proposto neste trabalho é apresentada na Fig. 78 e os atributos utilizados pelos preditores são mostrados nas Fig. 79 e 80.

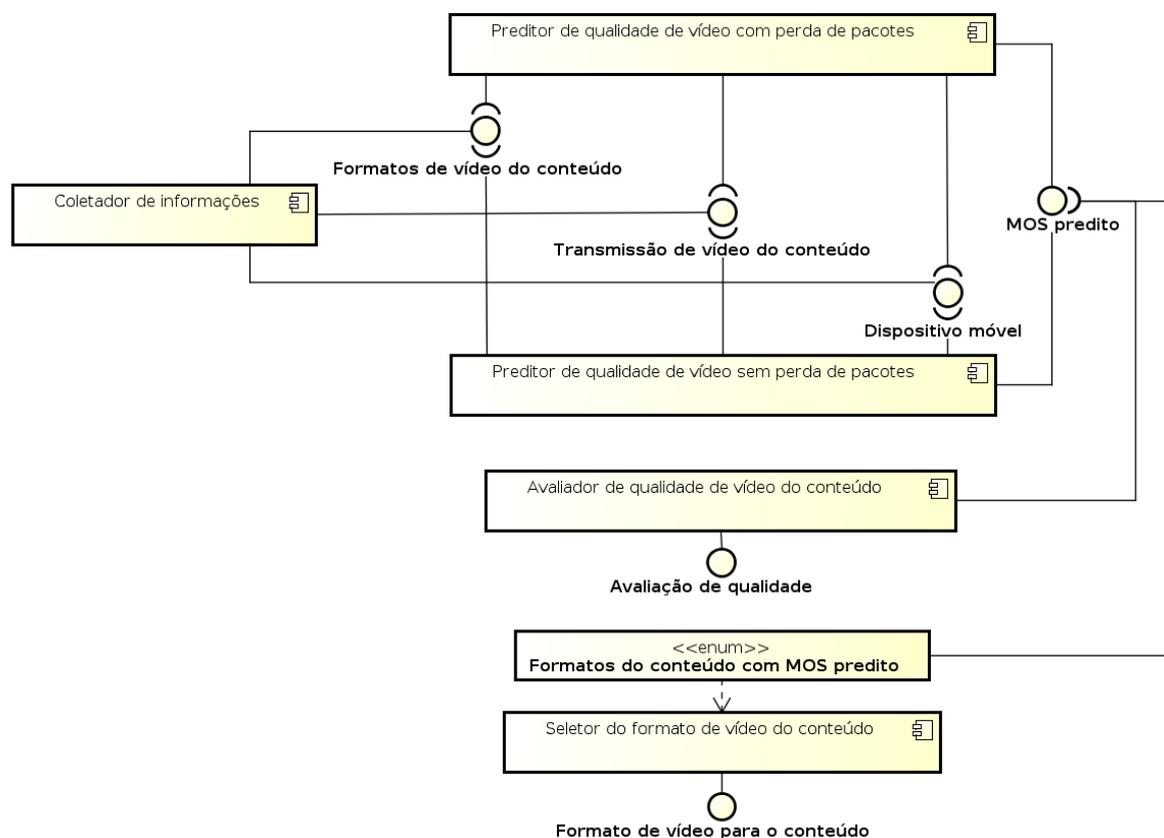


Figura 78: Relação entre os componentes que compõem o sistema proposto.

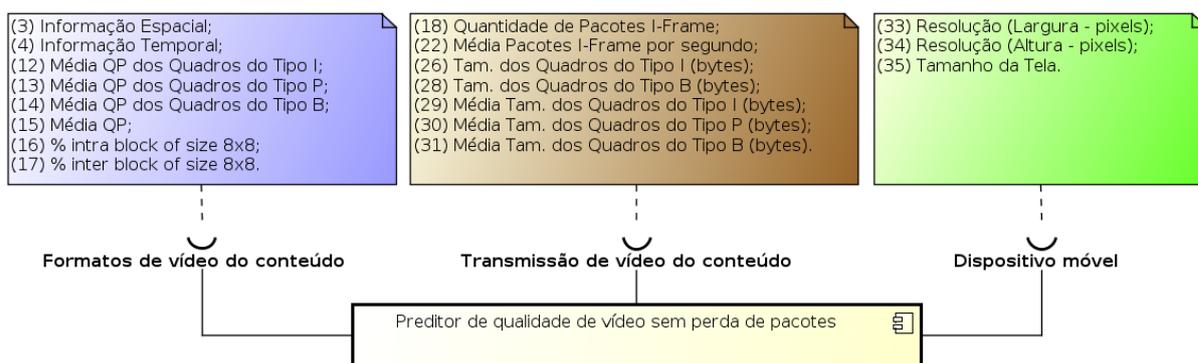


Figura 79: Informações utilizadas pelo preditor de qualidade de vídeo sem erros introduzidos durante a transmissão.

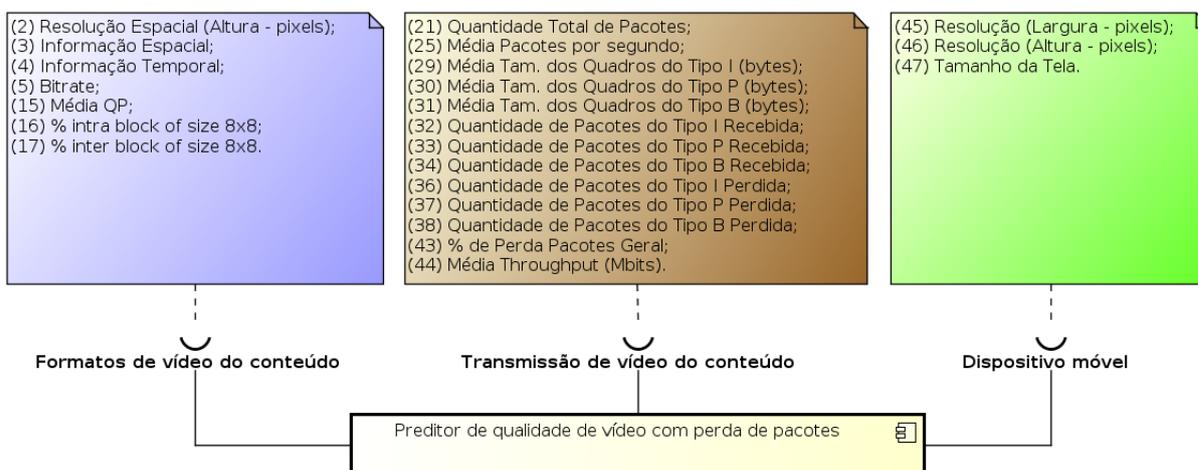


Figura 80: Informações utilizadas pelo preditor de qualidade de vídeo com erros introduzidos durante a transmissão.

Nesse contexto, formularam-se dois cenários de experimentos simulados no NS2 com EvalVid. No primeiro cenário, entre os instantes 15 e 30 segundos, foi ativado o modelo Gilbert-Elliot para gerar alguns erros durante a transmissão do conteúdo, fixando uma taxa de perda de 3%. O resultado da simulação desse cenário é mostrado na Fig. 81. O gerenciador de QoE atua a cada 10 segundos, observando as informações oriundas da transmissão, dos formatos disponíveis do conteúdo e do dispositivo móvel. No instante 20 segundos, o gerenciador de QoE percebe que a qualidade de vídeo do conteúdo é menor que 4. Então, o seletor do formato de vídeo do conteúdo é acionado para escolher a melhor codificação de vídeo diante das informações disponíveis. Após isso, no instante 40 segundos, a qualidade de vídeo do conteúdo é estabilizada.

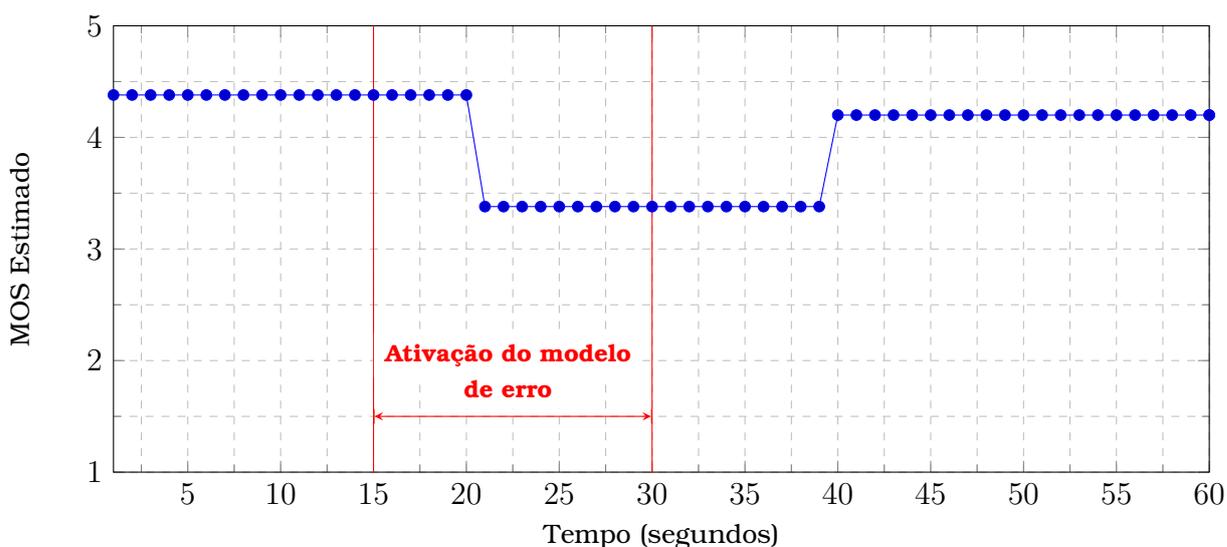


Figura 81: Gerenciador de QoE proposto atuando para o primeiro cenário.

Para o segundo e último cenário, após 10 segundos de exibição do conteúdo, ativou-se modelo Gilbert-Elliot com taxas de erros aleatória entre 1% e 9%. A cada 10 segundos era escolhida aleatoriamente a taxa de erros. O resultado desse cenário é apresentado na Fig. 82. Semelhante ao cenário anterior, o gerenciador de QoE atuou enquanto a qualidade de vídeo do conteúdo exibido no dispositivo do usuário era abaixo de Bom (MOS < 4). Dessa forma, a cada interação do gerenciador de QoE, a qualidade de vídeo melhorava aos poucos até chegar ao índice de qualidade superior a Bom.

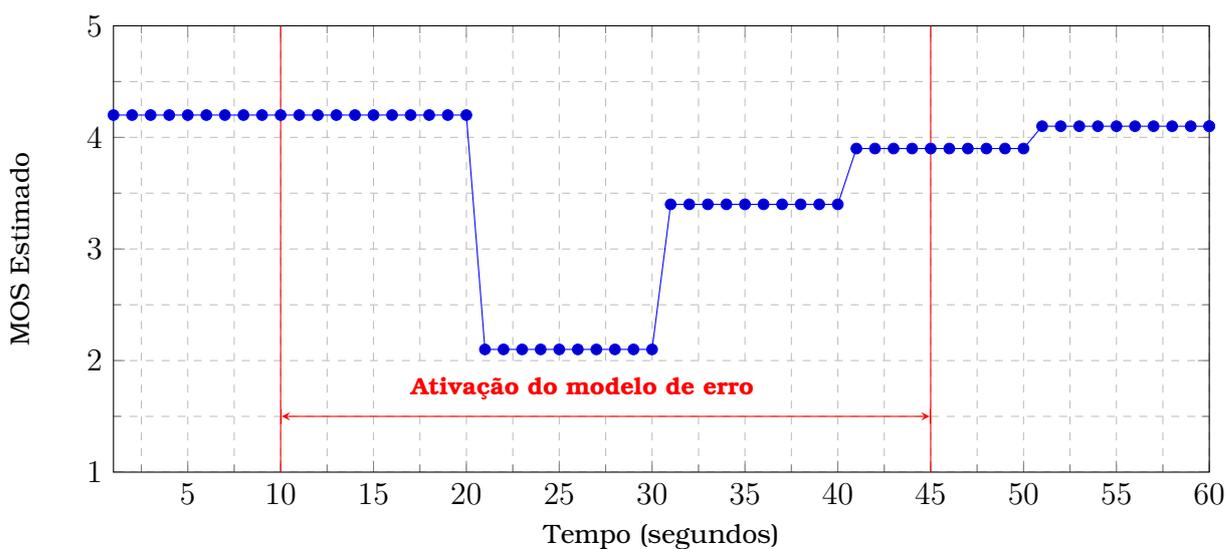


Figura 82: Gerenciador de QoE proposto atuando para o segundo cenário.

5.4 Conclusões

Nesse capítulo foram apresentadas as implementações do sistema de gerenciamento de qualidade de QoE durante a transmissão de vídeos para dispositivos móveis. Também foram mostrados os resultados obtidos nas avaliações subjetivas e nos treinamentos dos preditores de qualidade de vídeo, os quais foram integrados no sistema proposto neste trabalho.

No próximo capítulo são expostas as considerações finais deste trabalho, analisando o que foi desenvolvido e quais as perspectivas de trabalhos futuros.

CONCLUSÕES

“Seu tempo é limitado, então não o desperdice vivendo a vida de outra pessoa. Não fique preso pelo dogma – que é viver pelos resultados do que outras pessoas pensam. Não deixe o ruído da opinião dos outros afogar a sua voz interior. E o mais importante, tenha a coragem de seguir seu coração e sua intuição. Eles de alguma forma já sabem o que você realmente quer se tornar. Tudo o mais é secundário”

Steve Jobs

Favorecida pela melhoria na infraestrutura da Internet e popularização de *tablets*, *smartphones* e *smart TVs*, o uso de serviços de *streaming* de vídeos tem crescido rapidamente, permitindo que as pessoas possam assistir conteúdos em qualquer lugar, a qualquer hora e em qualquer dispositivo. Entretanto, nada disso tem valor algum se o provedor não garante a entrega dos conteúdos com qualidade ao consumidor.

Enquanto na qualidade de serviço são mensuradas características provenientes da comunicação entre os servidores e os clientes, a qualidade de experiência representa o nível de qualidade percebida até o usuário final, na qual há muitos desafios para serem resolvidos, especialmente no que diz respeito a dispositivos móveis.

Assim, neste trabalho apresentou-se a concepção e o desenvolvimento de um gerenciador de qualidade de experiência orientada à transmissão de vídeos para dispositivos móveis. Ele analisa informações oriundas da transmissão e dos vídeos disponíveis no provedor assim como as características do dispositivo móvel, com a finalidade de oferecer a codificação do conteúdo mais

adequada às condições atuais do ambiente em que o dispositivo móvel está situado.

Outra característica desse gerenciador é a possibilidade da interação do usuário com relação à qualidade do vídeo percebida em seu dispositivo móvel. Desse modo, os preditores de qualidade do vídeo utilizados pelo gerenciador podem ser treinados de tempos em tempos, aprimorando assim as métricas MSE e coeficiente de Pearson.

As vantagens do uso desse gerenciador são inúmeras. Dentre elas estão: a não requisição da opinião do usuário o tempo todo, apenas quando ele quiser; o ajuste automatizado da qualidade do conteúdo exibido no dispositivo móvel de acordo com as informações analisadas durante a transmissão; a não realização da predição da qualidade do vídeo no dispositivo móvel, preservando o tempo restante da bateria; e a não sobrecarga da rede com a transmissão de pacotes excedentes devido a uma má escolha da codificação do conteúdo.

Além disso, o gerenciador foi desenvolvido em módulos, ou seja, seus componentes podem ser alterados independentemente de modo que cada módulo não compromete o funcionamento dos demais. Isso possibilita a adaptação desse gerenciador perante a evolução dos dispositivos móveis e das técnicas utilizadas para a codificação dos vídeos. Também motiva o uso de outros algoritmos e técnicas de inteligência computação para prever a qualidade dos vídeos.

6.1 Resultados Obtidos

No percorrer da pesquisa, observou-se a necessidade de criar um gerenciador que administrasse a qualidade de vídeos transmitidos para dispositivos móveis. Um diferencial desse gerenciador está no não uso do vídeo de referência para mensurar a qualidade do vídeo no dispositivo móvel, ou seja, são apenas utilizadas informações sobre a transmissão, o vídeo e as características do dispositivo móvel para prever a qualidade do vídeo. Dessa forma, não há uma sobrecarga no dispositivo móvel maximizando o tempo útil da bateria.

No entanto, houve a necessidade de criar um conjunto de dados para treinar o gerenciador. Essa criação teve a participação de 44 pessoas que deram notas de qualidade para alguns vídeos distorcidos (sem e com introdução de erros) exibidos em um *tablet* e *smartphone*. Um fato interessante observado

a partir dessas notas foi a percepção das pessoas com relação a qualidade dos vídeos distorcidos nos dispositivos móveis. Alguns vídeos que foram exibidos no *smartphone* tiveram notas inferiores quando exibidos no *tablet* e vice-versa. Essa informação não é observada quando o PSNR é utilizado para mensurar a qualidade de vídeos porque nele necessita-se que o vídeo de referência e distorcido tenham a mesma resolução espacial.

A partir dos experimentos apresentados no Capítulo 5, o PSNR teve uma correlação de Pearson igual a 78,5% com o conjunto de dados criados para os vídeos sem introdução de erros, enquanto que o algoritmo de regressão logarítmica-linear mais o tratamento de dados proposto neste trabalho obteve 98% de correlação (uma melhora de 24,8%). Quando o conjunto de vídeos com introdução de erros (perda de pacotes) foi utilizado nos treinamentos, o PSNR obteve 75,9% de coeficiente de Pearson, enquanto que o algoritmo de regressão logística mais o tratamento de dados proposto neste trabalho obteve 89% de correlação (uma melhora de 17,3%). Esses índices foram alcançados após vários experimentos para escolher quais atributos eram mais representativos para cada tipo do vídeo distorcido (sem ou com introdução de erros) e foram descritos no decorrer do Capítulo 5.

A estratégia elaborada para o sistema é relevante, pois leva em consideração as características do dispositivo móvel. Isso é importante por que, como foi observado nas avaliações subjetivas, a percepção à falhas pelo usuário é maior quando são reproduzidos vídeos de alta definição em *smartphones*. A solução proposta impacta pouco no dispositivo do usuário, pois não são feitas análises quadro-a-quadro do vídeo para predizer a qualidade, atualmente. Além disso, o usuário pode avaliar a qualidade do vídeo recebido em seu dispositivo sobre o vídeo exibido no dispositivo móvel a qualquer momento e quando ele quiser. Assim, o algoritmo empregado nos preditores pode ser incrementado com essas novas informações.

Durante o desenvolvimento do presente trabalho, uma das principais dificuldades encontradas foi a mudança tecnológica constante dos dispositivos móveis. A cada três meses, em média, eram lançados novos *smartphones* e *tables* no mercado, dificultando a escolha dos dispositivos para as avaliações subjetivas. A Samsung, por exemplo, lançou um novo dispositivo móvel, em média, a cada 10 dias (VENTURA, 2015). Mesmo com uma pequena variação do tamanho das telas desses dispositivos, as especificações de *hardware* mudavam bastante, principalmente com relação à reprodução de vídeos de várias resoluções.

Uma das fases de suma importância deste trabalho foi a criação de um conjunto de dados sobre a qualidade de vídeos distorcidos exibidos em dispositivos móveis. Esses dados foram obtidos a partir da coleta de notas informadas por pessoas. Isso foi uma tarefa árdua por causa da disponibilidade das pessoas para participar, além do tempo exigido de pelo menos 30 minutos para avaliar a qualidade dos vídeos.

A escolha de quais informações seriam utilizadas pelos preditores de qualidade foi desafiador, pois haviam muitas possibilidades de combinações até encontrar um resultado que minimizasse a métrica MSE e maximizasse o índice de Pearson. Além disso, dependendo da técnica de inteligência computacional utilizada, o tempo necessário para treinar os preditores de qualidade levava horas.

6.2 Propostas de Continuidade

A fim de melhorar as técnicas utilizadas para mensurar e melhorar a qualidade percebida pelos usuários em um serviço de *streaming* de vídeo, uma série de questões ainda permanece. A avaliação da qualidade de vídeo ainda enfrenta desafios importantes que devem ser abordados.

O gerenciador descrito neste trabalho monitora informações provenientes do ambiente, levando em consideração informações sobre a transmissão, os conteúdos e o dispositivo móvel. Como sugestão de expansão, pode-se incorporar ao gerenciador o papel de escolha das melhores rotas entre o servidor de *streaming* e o cliente para diminuir o *delay* e *jitter* durante a recepção dos pacotes. Isso pode ser possível com a aplicação de técnicas de gerência da rede, como o *Multiprotocol Label Switching* (MPLS), uma técnica utilizada para gerenciar o endereçamento de rota dos pacotes sobre a rede IP, na própria rede do provedor de serviço.

Com o uso do MPLS, pode-se especificar classes para cada tipo de pacote trafegado na rede, visando melhorar a qualidade dos serviços utilizados pelos usuários. No entanto, existe uma grande quantidade e variedade de conteúdos que trafegam na rede, além da heterogeneidade de tecnologias e protocolos de comunicação presentes na infraestrutura.

Por outro lado, novos serviços têm sido criados, satisfazendo as crescentes demandas por sistemas mais rápidos, potentes e com maior capacidade de armazenamento e comunicação (BRAGA *et al.*, 2006). Uma consequência disso foi o aumento exponencial da complexidade dos sistemas de computa-

ção, tornando assim o seu gerenciamento complexo (AGOULMINE, 2011). Uma forma de automatizar esse processo é empregar conceitos de computação autônoma (*autonomic computing*), uma entidade computacional que é capaz de se auto-gerenciar, dado um conjunto de objetivos pré-definidos por seus criadores ou administradores (SAMAAN E KARMOUCH, 2009).

Com a propagação dos dispositivos móveis, é de suma importância levar em consideração as características desses dispositivos durante a concepção do preditor de qualidade de vídeo. Esses dispositivos têm limitações, quando comparado com outros dispositivos encontrados em uma residência, tais como o consumo de bateria, dimensão das telas pequenas (entre 3” e 12”), armazenamento limitado, e conectividade (PARK E JEONG, 2009). As metodologias utilizadas para avaliar a QoE em dispositivos móveis são adaptações de padrões do ITU-T e não são consideradas essas peculiaridades. Além disso, a comunicação sem fio, ambiente dinâmico e a cobertura do serviço podem influenciar na avaliação da QoE nestes cenários.

Devido às características heterogêneas e dinâmicas das redes e dos dispositivos móveis, a entrega de conteúdos de vídeo adaptativo tem sido utilizada. Informações sobre a rede e o dispositivo são coletadas e usadas para selecionar uma estratégia de compressão que maximiza a qualidade do vídeo (LEE *et al.*, 2012). No entanto, informações sobre o ambiente em que o usuário está inserido (casa, ônibus, praça etc.), bem como o estado comportamental cognitivo e psicológico dos usuários podem ser analisados neste contexto para determinar a QoE relacionada com o serviço de *streaming* de vídeo (MITRA *et al.*, 2014). Também podem ser agregadas informações sobre o perfil e gosto de cada usuário para aprimorar a QoE nesse tipo de serviço, além do uso de outros sensores como parâmetros para QoE (TIMMERER *et al.*, 2014).

Quando são aplicadas técnicas de mensuração objetiva da qualidade do vídeo, geralmente todo o conteúdo exibido no dispositivo do cliente é analisado. No entanto, pode-se utilizar o conceito de regiões de interesse (*regions of interest* – ROI). Assim, por meio da análise de objetos e da variação de luminância em regiões específicas em movimento, pode-se empregar ROI para incrementar o QoE em um serviço de *streaming* de vídeo (CAVIEDES, 2012).

Recentemente, os vídeos estão migrando para conteúdos em 3DTV (AS3DTV) e resoluções de alta definição (4K-UHD) e super HDTV (SHDTV) (CAVIEDES, 2012). Alguns serviços de *streaming* de vídeos, como YouTube, Netflix, Hulu Plus e Amazon Prime, já distribuem vídeos com resolução 4K-

UHD. Novos algoritmos para gerar conteúdos com essas resoluções, como o *high efficiency video coding* (HEVC), também estão sendo desenvolvidos (BAE *et al.*, 2013). A resolução, a distância entre os olhos da pessoa que está assistindo o conteúdo e as dimensões da tela do dispositivo podem ser considerados durante a criação de vídeos com conteúdo 3D, bem como na criação de novas técnicas para avaliar a QoE (BANITALEBI-DEHKORDI *et al.*, 2013).

Outras questões com relação à garantia de QoE para serviços de *streaming* de vídeos também estão em aberto, tais como:

- Qual o tamanho ideal de GoP para geração de vídeos? Qual a relação do GoP com a resolução espacial do vídeo? Existe alguma correlação do GoP com as informações espacial e temporal?
- Qual seria a relação do tipo de codificação de vídeo com o consumo de bateria do dispositivo móvel?
- Será que o uso de novas tendências, como computação nas nuvens, *Big Data* e *Fast Data*, podem ajudar na escalabilidade da gestão da QoE nesse tipo de serviço?
- Existe alguma diferença de percepção de qualidade de vídeos com resoluções HD, alta definição (4K-UHD) e super HDTV em dispositivos móveis? Qual a relação de perda de pacotes ocorridos durante a transmissão para essas novas resoluções?
- Qual o impacto da qualidade percebida do áudio do conteúdo com relação ao vídeo do conteúdo e vice-versa? Até que ponto um interfere no outro?
- Que circunstâncias físicas e emocionais realmente interferem na percepção da QoE pelo usuário nesse tipo de serviço?
- O uso de redes definidas por software (*software defined networks* – SDN) pode melhorar na gestão da QoE? Isso pode ajudar a criar cenários de conteúdos direcionados para cada tipo de dispositivo?
- Existem outras características de vídeo, da infraestrutura, do ambiente ou do dispositivo que podem ser utilizados para criar modelos de predição de qualidade de vídeo? Qual seria a relação entre elas? Quais trariam informações relevantes para elaboração de um gestor de QoE?

6.3 Considerações Finais

Um novo cenário está se formando com a popularização dos serviços de *streaming* de vídeos pela Internet. Fato este que nesses serviços só eram disponibilizados filmes e seriados produzidos por grandes produtoras de cinema e televisão. Agora, empresas como Netflix e Amazon estão investindo na criação de seus próprios conteúdos levando em consideração o perfil de seus clientes. Recentemente, elas vem recebendo indicações e prêmios em eventos de grande prestígio, como o Globo de Ouro, por esses conteúdos originais, além de alcançarem sucesso de crítica e público.

Assuntos referentes à qualidade em serviços de *streaming* de vídeos, principalmente a experiência usufruída pelos usuários, ainda é uma fonte a ser explorada e pesquisada. Assim, o objetivo geral deste trabalho que compreendeu na concepção e no desenvolvimento de um gerenciador da qualidade de um serviço de *streaming* de vídeos, capaz de ajustar automaticamente a qualidade do vídeo recebido pelo dispositivo móvel, obteve com êxito a finalidade de maximizar a experiência percebida pelo usuário.

Para isso, por intermédio dos objetivos específicos, propiciaram-se a investigação sobre o impacto da qualidade do vídeo percebida pelas pessoas em dispositivos móveis (*tablet* e *smartphone*), o estudo sobre a relação entre as informações oriundas do conteúdo, da transmissão e do dispositivo móvel e a idealização dos preditores de qualidade que, de acordo com a metodologia proposta neste trabalho, acarretou na obtenção do objetivo geral.

Os resultados obtidos no decorrer dos experimentos mostraram que o gerenciador desenvolvido pode ser empregado em um serviço de *streaming* de vídeos para dispositivos móveis, mesmo com a constante evolução dos mesmos. Além disso, o gerenciador pode ser aperfeiçoado a partir da colaboração dos usuários sobre a qualidade percebida por eles. Dessa forma, o gerenciador provê o formato do conteúdo diante das informações monitoradas sobre o ambiente proporcionando, assim, um serviço com qualidade para o usuário.

REFERÊNCIAS BIBLIOGRÁFICAS

4KHEVC. *4K-HEVC Video Subjective Quality Database*. Disponível em: <http://medialab.sjtu.edu.cn/resources/resources_subdataset.html>. Acesso em: 5 Abr. 2015.

ACUVUE. *The Snellen Test for Visual Acuity*. Disponível em: <<http://www.acuvue.co.za/sites/default/files/content/pdf/eyechart.pdf>>. Acesso em: 5 Abr. 2015.

ADOBE. *Adobe Media Server*. Disponível em: <<http://www.adobe.com/br/products/adobe-media-server-family.html>>. Acesso em: 5 Abr. 2015.

AGOULMINE, N. Introduction to Autonomic Concepts Applied to Future Self-Managed Networks. AGOULMINE, N. (Org.). *Autonomic Network Management Principles: From Concepts to Applications*. San Diego: Elsevier, 2011. p. 1–26.

AKHSHABI, S.; BEGEN, A. C.; DOVROLIS, C. An experimental evaluation of rate-adaptation algorithms in adaptive streaming over HTTP. *Proceedings of the second annual ACM conference on Multimedia systems*, San Jose (Estados Unidos), p. 157–168, Fev. 2011.

ALTMAN, E.; JIMÉNEZ, T. *NS Simulator for Beginners*. California: Morgan cLaypool publishers, 2012. 185 p.

AMIRSHAHI, S. A.; LARABI, M. Spatial-temporal Video Quality Metric based on an estimation of QoE. *Proceedings of the Third International Workshop on Quality of Multimedia Experience (QoMEX)*, Mechelen (Bélgica), p. 84–89, Set. 2011.

ANDROUTSELLIS-THEOTOKIS, S.; SPINELLIS, D. A survey of peer-to-peer content distribution technologies. *ACM Computing Surveys*, v. 36, n. 4, p. 335–371, Dez. 2004.

- ASTAH. *Astah Community – An add-on tool for your UML*. Disponível em: <<http://astah.net/editions/community>>. Acesso em: 5 Abr. 2015.
- BAE, S.-H.; KIM, J.; KIM, M.; CHO, S.; CHOI, J. S. Assessments of Subjective Video Quality on HEVC-Encoded 4K-UHD Video for Beyond-HDTV Broadcasting Services. *IEEE Transactions on Broadcasting*, v. 59, n. 2, p. 209–222, Jun. 2013.
- BANITALEBI-DEHKORDI, A.; POURAZAD, M.; NASIOPOULOS, P. 3D video quality metric for mobile applications. *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver (Canadá), p. 3731–3735, Maio 2013.
- BARAKOVIĆ, S.; BARAKOVIĆ, J.; BAJRIĆ, H. QoE Dimensions and QoE Measurement of NGN Services. *Proceedings of the 18th Telecommunications Forum (TELFOR)*, Belgrade (Sérvia), p. 1–5, Nov. 2010.
- BEGEN, A.; AKGUL, T.; BAUGHER, M. Watching Video over the Web: Part 1: Streaming Protocols. *IEEE Internet Computing*, v. 15, n. 2, p. 54–63, Mar./Abr. 2011.
- BOVIK, A. C. *The Essential Guide to Video Processing*. San Diego: Academic Press, 2009. 778 p.
- BRAGA, T. R. M.; SILVA, F. A.; RUIZ, L. B.; ASSUNÇÃO, H. P. Redes Autônomicas. *XXIV Simpósio Brasileiro de Redes de Computadores*, Curitiba (Brasil), p. 1–49, Maio-Jun. 2006.
- BRANDÃO, T.; QUELUZ, M. P. No-reference perceptual quality metric for H.264/AVC encoded video. *Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Scottsdale (Estados Unidos), p. 1–6, Jan. 2010a.
- BRANDÃO, T.; QUELUZ, M. P. No-Reference Quality Assessment of H.264/AVC Encoded Video. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 20, n. 11, p. 1437–1447, Nov. 2010c.
- BRANDÃO, T.; ROQUE, L.; QUELUZ, M. P. Quality assessment of H.264/AVC encoded video. *Proc. of Conf. on Telecommunications (ConfTele)*, Sta. Maria da Feira (Portugal), p. 1–4, Abr. 2010b.

- BROOKS, P.; HESTNES, B. User Measures of Quality of Experience: Why Being Objective and Quantitative is Important. *IEEE Network*, v. 24, n. 2, p. 8–13, Mar./Abr. 2010.
- CALLET, P. L.; MÖLLER, S.; PERKIS, A. *Qualinet White Paper on Definitions of Quality of Experience Output from the fifth Qualinet meeting, Version 1.2, March 12, 2013*. Qualinet. Switzerland, 2013. 24 p.
- CARISSIMI, A. D. S.; ROCHOL, J.; GRANVILLE, L. Z. *Redes de Computadores*. Porto Alegre: Bookman, 2009. 392 p.
- CATELLIER, A.; PINSON, M.; INGRAM, W.; WEBSTER, A. Impact of mobile devices and usage location on perceived multimedia quality. *Proceedings of the Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley (Austrália), p. 39–44, Jul. 2012.
- CAVIEDES, J. The Evolution of Video Processing Technology and Its Main Drivers. *Proceedings of the IEEE*, v. 100, n. 4, p. 872–877, Abr. 2012.
- CERQUEIRA, E.; AGUIAR, E.; RIKER, A.; ABELEM, A.; MU, M. Video quality estimator for wireless mesh networks. *Proceedings of the IEEE 20th International Workshop on Quality of Service (IWQoS)*, Coimbra (Portugal), p. 1–9, Jun. 2012a.
- CERQUEIRA, E.; NETO, A.; CURADO, M.; RIKER, A.; IMMICH, R.; BARROS, H.; AGUIAR, E.; ABELEM, A. A parametric QoE video quality estimator for Wireless Networks. *Proceedings of the IEEE Globecom Workshops (GC Wkshps) (2012)*, Anaheim (Estados Unidos), p. 1–9, Dez. 2012b.
- CHANDLER, D.; HEMAMI, S. VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images. *IEEE Transactions on Image Processing*, v. 16, n. 9, p. 2284–2298, Set. 2007.
- CISCO. *Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. White Paper, May 29, 2013*. Cisco Systems, Inc. San Jose, 2013. 15 p.
- CODEBLOCKS. *Code::Blocks – The open source, cross platform, free C, C++ and Fortran IDE*. Disponível em: <<http://www.codeblocks.org/>>. Acesso em: 5 Abr. 2015.
- COMER, D. E. *Internetworking with TCP/IP: Principles, Protocols, and Architecture (Vol. 1)*. 6 ed. New Jersey: Pearson Prentice Hall, 2014. 733 p.

CRAWLEY, E.; NAIR, R.; RAJAGOPALAN, B.; SANDICK, H. *A framework for QoS-based routing in the Internet*. IETF RFC 2386. California, 1998. 37 p.

DARWIN. *Darwin Streaming Server*. Disponível em: <<http://dss.macosforge.org/>>. Acesso em: 5 Abr. 2015.

DE MOOR, K.; KETYKO, I.; JOSEPH, W.; DERYCKERE, T.; DE MAREZ, L.; MARTENS, L.; VERLEYE, G. Proposed Framework for Evaluating Quality of Experience in a Mobile, Testbed-oriented Living Lab Setting. *Mobile Networks and Applications*, v. 15, n. 3, p. 378–391, Jun. 2010.

DE SIMONE, F.; NACCARI, M.; TAGLIASACCHI, M.; DUFAUX, F.; TUBARO, S.; EBRAHIMI, T. Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel. *International Workshop on Quality of Multimedia Experience (QoMEX)*, San Diego (Estados Unidos), p. 204–209, Jul. 2009.

DE SIMONE, F.; TAGLIASACCHI, M.; NACCARI, M.; TUBARO, S.; EBRAHIMI, T. A H.264/AVC video database for the evaluation of quality metrics. *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Dallas (Estados Unidos), p. 2430–2433, Mar. 2010.

ECLIPSE. *Eclipse Indigo*. Disponível em: <<https://eclipse.org/indigo/>>. Acesso em: 5 Abr. 2015.

EPFL. *3D Video Quality Assessment Database*. Disponível em: <<http://mmspg.epfl.ch/3dvqa>>. Acesso em: 5 Abr. 2015.

EPFL-POLIMI. *EPFL-PoliMI video quality assessment database*. Disponível em: <<http://vqa.como.polimi.it/index.htm>>. Acesso em: 5 Abr. 2015.

EVALVID. *EvalVid – A Video Quality Evaluation Tool-set*. Disponível em: <<http://www2.tkn.tu-berlin.de/research/evalvid/>>. Acesso em: 5 Abr. 2015.

FACELI, K.; LORENA, A. C.; GAMA, J.; DE CARVALHO, A. C. P. L. F. *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*. Rio de Janeiro: LTC, 2012. 394 p.

FARIAS, M. C. Q. Video Quality Metrics. RANGO, F. D. (Org.). *Digital Video*. Vukovar: InTech, 2010. p. 329–358.

FARIAS, M. C. Q.; CARVALHO, M. M.; KUSSABA, H. T. M.; NORONHA, B. H. A. A hybrid metric for digital video quality assessment. *Proceedings of the IEEE*

International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Nuremberg (Alemanha), p. 1–6, Jun. 2011.

FERRO, E.; POTORTI, F. Bluetooth and Wi-Fi wireless protocols: a survey and a comparison. *IEEE Wireless Communications*, v. 12, n. 1, p. 12–26, Fev. 2005.

FFMPEG. *FFMPEG*. Disponível em: <<http://www.ffmpeg.org/>>. Acesso em: 5 Abr. 2015.

FFSERVER. *FFserver*. Disponível em: <<http://www.ffmpeg.org/ffserver.html>>. Acesso em: 5 Abr. 2015.

FIEDLER, M.; HOSSFELD, T.; TRAN-GIA, P. A Generic Quantitative Relationship between Quality of Experience and Quality of Service. *IEEE Network*, v. 24, n. 2, p. 36–41, Mar./Abr. 2010.

FU-ZHENG, Y.; XIN-DAI, W.; YI-LIN, C.; SHUAI, W. A no-reference video quality assessment method based on digital watermark. *Proceedings of the 14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Beijing (China), p. 2707–2710, Set. 2003.

GHAREEB, M.; VIHO, C. Hybrid QoE Assessment Is Well-Suited for Multiple Description Coding Video Streaming in Overlay Networks. *Proceedings of the Eighth Annual Communication Networks and Services Research Conference (CNSR)*, Montreal (Canadá), p. 327–333, Maio 2010.

GOLDMANN, L.; DE SIMONE, F.; EBRAHIMI, T. A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. *Proc. SPIE*, v. 7526, p. 75260S–75260S–11, Jan. 2010.

GPAC. *GPAC – Multimedia Open Source Project*. Disponível em: <<http://gpac.wp.mines-telecom.fr/mp4box/>>. Acesso em: 5 Abr. 2015.

GROSS, J.; KLAUE, J.; KARL, H.; WOLISZ, A. Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming. *Computer Communications*, v. 27, n. 11, p. 1044–1055, Jul. 2004.

GROUP, A. . T. W. *Triple-play services quality of experience (QoE) requirements*. Disponível em: <<http://www.broadband-forum.org/technical/download/TR-126.pdf>>. Acesso em: 5 Abr. 2015.

- GSMARENA. *GSMarena.com - GSM phone reviews, news, opinions, votes, manuals and more*. Disponível em: <<http://www.gsmarena.com/>>. Acesso em: 5 Abr. 2015.
- GUNAWAN, I. P.; GHANBARI, M. Efficient Reduced-Reference Video Quality Meter. *IEEE Trans. on Broadcasting*, v. 54, n. 3, p. 669–679, Set. 2008.
- HASSLINGER, G.; HOHLFELD, O. The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet. *Proceedings of 14th GI/ITG Conference - Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB)*, Dortmund (Alemanha), p. 1–15, Mar. 2008.
- HEMAMI, S. S.; REIBMAN, A. R. No-reference image and video quality estimation: Applications and human-motivated design. *Signal Processing: Image Communication*, v. 25, n. 7, p. 469–481, Ago. 2010.
- HWANG, J.-N. Wireless MediaNets: application-driven next-generation wireless IP networks. *Multimedia Systems*, v. 13, n. 3, p. 1–35, Nov. 2010.
- IDC. *Smartphone OS Market Share, Q4 2014*. Disponível em: <<http://www.idc.com/prodserv/smartphone-os-market-share.jsp>>. Acesso em: 5 Abr. 2015.
- IRCCYN/IVC. *IRCCyN/IVC 1080i Database*. Disponível em: <<http://www.irccyn.ec-nantes.fr/spip.php?article541>>. Acesso em: 5 Abr. 2015a.
- IRCCYN/IVC. *IRCCyN/IVC H264 AVC vs SVC VGA Video database*. Disponível em: <<http://130.66.64.103/spip.php?article783>>. Acesso em: 5 Abr. 2015b.
- IRCCYN/IVC. *IRCCyN/IVC Influence Content Video VGA database*. Disponível em: <<http://130.66.64.103/spip.php?article771>>. Acesso em: 5 Abr. 2015c.
- ISHIHARA, S. *Tests for Colour-Blindness (24 plates)*. 5 ed. Tóquio: Kanehara Co. LTD, 1972. 35 p.
- ISSARIYAKUL, T.; HOSSAIN, E. *Introduction to Network Simulator NS2*. 2 ed. New York: Springer, 2011. 527 p.
- IT-IST. *IT-IST Lisbon H.264/MPEG-2 Video Quality Database*. Disponível em: <http://amalia.img.lx.it.pt/tgsb/H264_test/>. Acesso em: 5 Abr. 2015.

INTERNATIONAL TELECOMMUNICATION UNION – ITU-T. *Method for objective measurements of perceived audio quality*. RECOMMENDATION ITU-R BS.1387-1. Geneva, 1999. 100 p.

INTERNATIONAL TELECOMMUNICATION UNION – ITU-T. *Methodology for the subjective assessment of the quality of television pictures*. ITU-T Recommendation BT 500-11. Geneva, 2002. 48 p.

INTERNATIONAL TELECOMMUNICATION UNION – ITU-T. *Opinion model for video-telephony applications*. Recommendation ITU-T G.1070. Geneva, 2007. 28 p.

INTERNATIONAL TELECOMMUNICATION UNION – ITU-T. *Definitions of terms related to quality of service*. RECOMMENDATION ITU-T E.800. Geneva, 2008. 30 p.

INTERNATIONAL TELECOMMUNICATION UNION – ITU-T. *Subjective video quality assessment methods for multimedia applications*. ITU-T Recommendation P-910. Geneva, 2009. 37 p.

KAWANO, T.; YAMAGISHI, K.; WATANABE, K.; OKAMOTO, J. No reference video-quality-assessment model for video streaming services. *Proceedings of the 18th International Packet Video Workshop (PV)*, Hong Kong (China), p. 158–164, Dez. 2010.

KEIMEL, C.; HABIGT, J.; HABIGT, T.; ROTHBUCHER, M.; DIEPOLD, K. Visual quality of current coding technologies at high definition IPTV bitrates. *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Saint Malo (França), p. 390–393, Out. 2010.

KEIMEL, C.; REDL, A.; DIEPOLD, K. The TUM high definition video datasets. *Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley (Austrália), p. 97–102, Jul. 2012.

KHAN, A.; SUN, L.; IFEACHOR, E. An ANFIS-Based Hybrid Video Quality Prediction Model for Video Streaming over Wireless Networks. *Proceedings of the Second International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST)*, Cardiff (País de Gales), p. 357–362, Set. 2008.

KIM, H. J.; CHOI, S. G. A study on a QoS/QoE correlation model for QoE evaluation on IPTV service. *The 12th International Conference on Advanced*

Communication Technology (ICACT), Seoul (Coréia do Sul), p. 1377–1382, Fev. 2010.

KIM, Y.-H.; HAN, J.; KIM, H.; SHIN, J. Novel no-reference video quality assessment metric with estimation of dynamic range distortion. *Proceedings of the 12th International Conference on Advanced Communication Technology (ICACT)*, Phoenix Park (Irlanda), p. 1689–1692, Fev. 2010.

KLAUE, J.; RATHKE, B.; WOLISZ, A. Ultra High Definition HEVC DASH Data Set. *Proceedings of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, Illinois (Estados Unidos), p. 255–272, Set. 2003.

KORHONEN, J.; BURINI, N.; JUNYONG, Y.; NADERNEJAD, E. How to evaluate objective video quality metrics reliably. *Proceedings of the Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, Yarra Valley (Austrália), p. 57–62, Jul. 2012.

KUMAR, A. *Mobile Broadcasting with WiMAX: Principles, Technology, and Applications*. United Kingdom: Elsevier, 2008. 619 p.

KUMAR, A. *Implementing Mobile TV: ATSC Mobile DTV, MediaFLO, DVB-H/SH, DMB, WiMAX, 3G Systems, and Rich Media Applications*. 2 ed. Burlington: Elsevier, 2010. 658 p.

KURAN, M. S.; TUGCU, T. A survey on emerging broadband wireless access technologies. *Computer Networks*, v. 51, n. 11, p. 3013–3046, Ago. 2007.

KUROSE, J. F.; ROSS, K. W. *Redes de Computadores e a Internet: Uma Nova Abordagem*. São Paulo: Pearson, 2003. 548 p.

LAGHARI, K. R.; KHAN, I.; CRESPI, N. Quantitative and qualitative assessment of QoE for multimedia services in wireless environment. *Proceedings of the 4th Workshop on Mobile Video (MoVid)*, Chapel Hill (Estados Unidos), p. 7–12, Fev. 2012.

LE FEUVRE, J.; THIESSE, J.-M.; PARMENTIER, M.; RAULET, M.; DAGUET, C. Ultra High Definition HEVC DASH Data Set. *Proceedings of the 5th ACM Multimedia Systems Conference*, Singapore (Singapura), p. 7–12, Mar. 2014.

LEE, J.-S.; DE SIMONE, F.; EBRAHIMI, T.; RAMZAN, N.; IZQUIERDO, E. Quality assessment of multidimensional video scalability. *IEEE Communications Magazine*, v. 50, n. 4, p. 38–46, Abr. 2012.

LG. *LG Smart TV*. Disponível em: <<http://www.lge.com/br/tvs-smart>>. Acesso em: 5 Abr. 2015.

LIN, X.; TIAN, X.; CHEN, Y. No-reference video quality assessment based on region of interest. *Proceedings of the 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, Yichang (China), p. 1924–1927, Abr. 2012.

LIVE. *LIVE Video Quality Database*. Disponível em: <http://live.ece.utexas.edu/research/quality/live_video.html>. Acesso em: 5 Abr. 2015a.

LIVE. *LIVE Mobile Video Quality Database*. Disponível em: <http://live.ece.utexas.edu/research/quality/live_mobile_video.html>. Acesso em: 5 Abr. 2015b.

LIVE555. *LIVE555 Streaming Media*. Disponível em: <<http://www.live555.com/liveMedia/>>. Acesso em: 5 Abr. 2015.

MA, L.; LI, S.; NGAN, K. N. Reduced-Reference Video Quality Assessment of Compressed Video Sequences. *IEEE Trans. on Circuits and Systems for Video Technology*, v. 22, n. 10, p. 1441–1456, Out. 2012.

MINISTÉRIO DA SAÚDE. *Projeto Olhar Brasil: triagem de acuidade visual: manual de orientação*. Série A. Normas e Manuais Técnicos. Brasília, 2008. 25 p.

MITRA, K.; ZASLAVSKY, A.; AHLUND, C. Context-Aware QoE Modelling, Measurement and Prediction in Mobile Computing Systems. *IEEE Transactions on Mobile Computing*, v. 99, n. PrePrints, p. 1–14, 2014.

MOHAMED, S.; RUBINO, G. A study of real-time packet video quality using random neural networks. *IEEE Trans. on Circuits and Systems for Video Technology*, v. 12, n. 12, p. 1071–1083, Dez. 2002.

MOORTHY, A. K.; CHOI, L. K.; DE VECIANA, G.; BOVIK, A. C. Subjective Analysis of Video Quality on Mobile Devices. *Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Scottsdale (Estados Unidos), p. 1–6, Jan. 2012a.

MOORTHY, A.; CHOI, L. K.; BOVIK, A.; DE VECIANA, G. Video Quality Assessment on Mobile Devices: Subjective, Behavioral and Objective Studies. *IEEE*

Journal of Selected Topics in Signal Processing, v. 6, n. 6, p. 652–671, Out. 2012b.

MSAKNI, H. G.; YOUSSEF, H. Is QoE estimation based on QoS parameters sufficient for video quality assessment? *Proceedings of the 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Sardinia (Itália), p. 538–544, Jul. 2013.

NS. *Network Simulation 2*. Disponível em: <<http://sourceforge.net/projects/nsnam/>>. Acesso em: 5 Abr. 2015.

OCTAVE. *GNU Octave*. Disponível em: <<https://www.gnu.org/software/octave/>>. Acesso em: 5 Abr. 2015.

O'DRISCOLL, G. *Next Generation IPTV Services and Technologies*. New Jersey: Wiley, 2008. 512 p.

PANASONIC. *Panasonic Smart TV*. Disponível em: <<http://www.panasonic.com/br/consumidor/tv/viera-led.html>>. Acesso em: 5 Abr. 2015.

PARK, S.; JEONG, S.-H. Mobile IPTV: Approaches, Challenges, Standards, and QoS Support. *IEEE Internet Computing*, v. 13, n. 3, p. 23–31, Maio/Jun. 2009.

PAUL, S. *Digital Video Distribution in Broadband, Television, Mobile and Converged Networks - Trends, Challenges and Solutions*. United Kingdom: John Wiley Sons Ltd, 2011. 385 p.

PEI, Z.; DENG, Z.; YANG, B.; CHENG, X. Application-oriented wireless sensor network communication protocols and hardware platforms: A survey. *IEEE International Conference on Industrial Technology*, Chengdu (China), p. 1–6, Abr. 2008.

PETERSON, L. L.; DAVIE, B. S. *Redes de Computadores: Uma Abordagem de Sistemas*. São Paulo: Elsevier, 2004. 588 p.

PHILIPS. *Philips Smart TV*. Disponível em: <<http://www.philips.com.br/c/televisao/29614/cat/>>. Acesso em: 5 Abr. 2015.

PIAMRAT, K.; KSENTINI, A.; VIHO, C.; BONNIN, J. QoE-Aware Admission Control for Multimedia Applications in IEEE 802.11 Wireless Networks. *Proceedings of the IEEE 68th Vehicular Technology Conference (VTC)*, Calgary (Canadá), p. 1–5, Set. 2008.

PITREY, Y.; BARKOWSKY, M.; LE CALLET, P.; PÉPION, R. Subjective quality assessment of MPEG-4 Scalable Video Coding in a mobile scenario. *2nd European Workshop on Visual Information Processing (EUVIP)*, Paris (França), p. 86–91, Jul. 2010.

PITREY, Y.; BARKOWSKY, M.; PÉPION, R.; LE CALLET, P.; HLAVACS, H. Influence of the source content and encoding configuration on the perceived quality for scalable video coding. *Proc. SPIE*, v. 8291, p. 82911K–1–82911K–8, Fev. 2012.

PROVISU. *Snellen Test*. Disponível em: <http://www.provisu.ch/Age/Snellenchart_en.pdf>. Acesso em: 5 Abr. 2015.

PULIPAKA, A.; SEELING, P.; REISSLEIN, M.; KARAM, L. Traffic and Statistical Multiplexing Characterization of 3-D Video Representation Formats. *IEEE Transactions on Broadcasting*, v. 59, n. 2, p. 382–389, Jun. 2013.

PÉCHARD, S.; PÉPION, R.; CALLET, P. L. Suitable methodology in subjective video quality assessment: a resolution dependent paradigm. *Third International Workshop on Image Media Quality and its Applications (IMQA)*, Kyoto (Japão), p. 1–6, Set. 2008.

QMEE. *Online in 60 seconds [infographic] – a year later*. Disponível em: <<http://blog.qmee.com/online-in-60-seconds-infographic-a-year-later/>>. Acesso em: 5 Abr. 2015.

RAAKE, A.; EGGER, S. Quality and Quality of Experience. MÖLLER, S.; RAAKE, A. (Org.). *Quality of Experience: Advanced Concepts, Applications and Methods*. New York: Springer, 2014. p. 11–34.

RAYCHAUDHURI, D.; MANDAYAM, N. B. Frontiers of Wireless and Mobile Communications. *Proceedings of the IEEE*, v. 100, n. 4, p. 824–840, Abr. 2012.

ROHANI, M.; AVANAKI, A. N.; NADER-ESFAHANIA, S.; BASHIRPOUR, M. A Reduced Reference Video Quality Assessment method based on the human

motion perception. *Proceedings of the 5th International Symposium on Telecommunications (IST)*, Tehran (Irã), p. 831–835, Dez. 2010.

ROSÁRIO, D.; CERQUEIRA, E.; NETO, A.; RIKER, A.; IMMICH, R.; CURADO, M. A QoE handover architecture for converged heterogeneous wireless network. *Wireless Networks*, v. 19, n. 8, p. 2005–2020, Nov. 2013.

RUBINO, G.; VARELA, M.; BONNIN, J. Controlling Multimedia QoS in the Future Home Network Using the PSQA Metric. *Computer Journal*, v. 49, n. 2, p. 137–155, Mar. 2006.

SAAD, M. A.; BOVIK, A. C. Natural motion statistics for no-reference video quality assessment. *Proceedings of the International Workshop on Quality of Multimedia Experience (QoMEX)*, San Diego (Estados Unidos), p. 163–167, Jul. 2009.

SALOMON, D.; MOTTA, G. *Handbook of Data Compression*. 5 ed. New York: Springer, 2010. 1370 p.

SAMAAN, N.; KARMOUCH, A. Towards Autonomic Network Management: an Analysis of Current and Future Research Directions. *IEEE Communications Surveys Tutorials*, v. 11, n. 3, p. 22–36, Jul./Set. 2009.

SAMSUNG. *Samsung Smart TV*. Disponível em: <<http://www.samsung.com/br/consumer/tv-audio-video/tv/smart-tv>>. Acesso em: 5 Abr. 2015.

SAUTER, M. *Beyond 3G – Bringing Networks, Terminals and the Web Together: LTE, WiMAX, IMS, 4G Devices and the Mobile Web 2.0*. United Kingdom: Jowhn Wiley Sons Ltd, 2009. 362 p.

SCHOLLMEIER, R. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. *Proceedings of the First International Conference on Peer-to-Peer Computing*, Linköpings (Suécia), p. 101–102, Ago. 2001.

SCILAB. *Scilab – Free and open source software for numerical computation*. Disponível em: <<http://www.scilab.org/>>. Acesso em: 5 Abr. 2015.

SEELING, P.; REISSLEIN, M. Video Transport Evaluation With H.264 Video Traces. *IEEE Communications Surveys Tutorials*, v. 14, n. 4, p. 1142–1165, Out./Dez. 2012.

SEMPTOSHIBA. *Semp Toshiba Smart TV*. Disponível em: <<http://www.semptoshiba.com.br/ecommerce/home-produtos/tv/smart-tv.html>>. Acesso em: 5 Abr. 2015.

SESHADRINATHAN, K.; SOUNDARARAJAN, R.; BOVIK, A. C.; CORMACK, L. K. A Subjective Study to Evaluate Video Quality Assessment Algorithms. *SPIE Proceedings Human Vision and Electronic Imaging*, v. 7527, p. 1–10, Fev. 2010a.

SESHADRINATHAN, K.; SOUNDARARAJAN, R.; BOVIK, A. C.; CORMACK, L. K. Study of Subjective and Objective Quality Assessment of Video. *IEEE Transactions on Image Processing*, v. 19, n. 6, p. 1427–1441, Jun. 2010b.

SESHADRINATHAN, K.; BOVIK, A. Automatic prediction of perceptual quality of multimedia signals—a survey. *Multimedia Tools and Applications*, v. 51, n. 1, p. 163–186, Jan. 2011.

SHEIKH, H.; BOVIK, A. Image information and visual quality. *IEEE Transactions on Image Processing*, v. 15, n. 2, p. 430–444, Fev. 2006.

SHEIKH, H.; BOVIK, A.; DE VECLIANA, G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, v. 14, n. 12, p. 2117–2128, Dez. 2005.

SHMUELI, R.; HADAR, O.; HUBER, R.; MALTZ, M.; HUBER, M. Effects of an Encoding Scheme on Perceived Video Quality Transmitted Over Lossy Internet Protocol Networks. *IEEE Transactions on Broadcasting*, v. 54, n. 3, p. 628–640, Set. 2008.

SIMPSON, W. *Video Over IP: IPTV, Internet Video, H.264, P2P, Web TV, and Streaming: A Complete Guide to Understanding the Technology*. 2 ed. Burlington: Focal Press, 2008. 504 p.

SONG, L.; TANG, X.; ZHANG, W.; YANG, X.; XIA, P. The SJTU 4K Video Sequence Dataset. *Proceedings of the Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, Klagenfurt am Wörthersee (Austria), p. 34–35, Jul. 2013.

SONY. *Sony Smart TV*. Disponível em: <<http://www.sony.com.br/eletronicos/televisores/filtro/smart-tv/>>. Acesso em: 5 Abr. 2015.

STAELENS, N.; DEN BROECK, W. V.; PITREY, Y. Lessons Learned during Real-life QoE Assessment. *Proceedings of the 8th European Interactive TV Conference (EuroITV)*, Tampere (Finlândia), p. 305–306, Jun. 2010.

STANKIEWICZ, R.; JAJSZCZYK, A. A survey of QoE assurance in converged networks. *Computer Networks*, v. 55, n. 7, p. 1459–1473, Maio 2011.

STERNBERG, R. J. *Psicologia Cognitiva*. 5 ed. São Paulo: Cengage Learning, 2010. 598 p.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. *Introdução ao Data Mining*. Rio de Janeiro: Ciência Moderna, 2009. 908 p.

TAPOLCAI, J.; CHOLDA, P.; CINKLER, T.; WAJDA, K.; JAJSZCZYK, A.; AUTENRIETH, A.; BODAMER, S.; COLLE, D.; FERRARIS, G.; LONSETHAGEN, H.; SVINNSET, I.-E.; VERCHERE, D. Quality of resilience (QoR): NOBEL approach to the multi-service resilience characterization. *2nd International Conference on Broadband Networks*, Boston (Estados Unidos), p. 1328–1337, Out. 2005.

TIMMERER, C.; WALTL, M.; RAINER, B.; MURRAY, N. Sensory Experience: Quality of Experience Beyond Audio-Visual. MÖLLER, S.; RAAKE, A. (Org.). *Quality of Experience: Advanced Concepts, Applications and Methods*. New York: Springer, 2014. p. 351–365.

TRACE. *Video Trace Library (TRACE)-YUV Video Sequences*. Disponível em: <<http://trace.kom.aau.dk/yuv/index.html>>. Acesso em: 5 Abr. 2015.

TSI. *Ultra High Definition HEVC DASH Data Set*. Disponível em: <<http://download.tsi.telecom-paristech.fr/gpac/dataset/dash/uhd/>>. Acesso em: 5 Abr. 2015.

TUM. *TUM 1080p25 Data Set*. Disponível em: <<http://www.ldv.ei.tum.de/en/research/videolab/data-sets-downloads/tum-1080p25-data-set/>>. Acesso em: 5 Abr. 2015a.

TUM. *TUM 1080p50 Data Set*. Disponível em: <<http://www.ldv.ei.tum.de/en/research/videolab/data-sets-downloads/tum-1080p50-data-set/>>. Acesso em: 5 Abr. 2015b.

VENTURA, F. *Samsung anuncia, em média, um novo Galaxy a cada 10 dias*. Disponível em: <<http://gizmodo.uol.com.br/samsung-linha-galaxy/>>. Acesso em: 5 Abr. 2015.

VLC. *VideoLan VLC*. Disponível em: <<http://www.videolan.org/>>. Acesso em: 5 Abr. 2015.

WANG, H.-S.; MOAYERI, N. Finite-state Markov channel—a useful model for radio communication channels. *IEEE Transactions on Vehicular Technology*, v. 44, n. 1, p. 163–171, Fev. 1995.

WANG, Z.; BOVIK, A.; SHEIKH, H.; SIMONCELLI, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, v. 13, n. 4, p. 600–612, Abr. 2004.

WEBSTER, A. A.; JONES, C. T.; PINSON, M. H.; VORAN, S. D.; WOLF, S. An Objective video quality assessment system based on human perception. *Proc. SPIE Human Vision, Visual Processing, and Digital Display IV*, v. 1913, n. 1993, p. 15–26, Set. 1993.

WEKA. *Weka 3: Data Mining Software in Java*. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 5 Abr. 2015.

WIFI. *WiFi Alliance*. Disponível em: <<http://www.wi-fi.org/>>. Acesso em: 5 Abr. 2015.

WIMAX. *WiMAX Forum*. Disponível em: <<http://www.wimaxforum.org/>>. Acesso em: 5 Abr. 2015.

WINDOWS. *Adobe Media Server*. Disponível em: <<http://technet.microsoft.com/en-us/windowsserver/dd448620.aspx>>. Acesso em: 5 Abr. 2015.

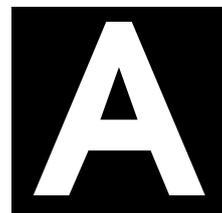
WINKLER, S.; MOHANDAS, P. The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics. *IEEE Transactions on Broadcasting*, v. 54, n. 3, p. 660–668, Set. 2008.

WIRELESSHD. *WirelessHD Specification Version 1.1 Overview*. Disponível em: <<http://www.wirelesshd.org/pdfs/WirelessHD-Specification-Overview-v1.1May2010.pdf>>. Acesso em: 5 Abr. 2015.

WU, H.; RAO, K. *Digital Video Image Quality and Perceptual Coding*. Boca Raton: CRC Press, 2006. 615 p.

XUE, J.; CHEN, C. W. A study on perception of mobile video with surrounding contextual influences. *Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, Melbourne (Austrália), p. 248–253, Jul. 2012.

- YAMAGISHI, K.; KAWANO, T.; HAYASHI, T. Hybrid Video-Quality-Estimation Model for IPTV Services. *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Honolulu (Estados Unidos), p. 1–5, Nov./Dez. 2009.
- YANG, Y.; LU, Z.; WEN, X.; ZHENG, W.; ZHANG, A. A No-Reference Video Quality Estimation Model over Wireless Networks. *IEEE Vehicular Technology Conference (VTC Fall)*, San Francisco (Estados Unidos), p. 1–5, Set. 2011.
- YAO, J.; ZHANG, Y.; XU, G.; JIN, M. No-Reference Objective Quality Assessment for Video Communication Services Based on Feature Extraction. *Proceedings of the 2nd International Congress on Image and Signal Processing (CIS)*, Tianjin (China), p. 1–6, Out. 2009.
- YASM. *The Yasm Modular Assembler Project*. Disponível em: <<http://yasm.tortall.net/>>. Acesso em: 5 Abr. 2015.
- YITONG, L.; YUN, S.; YINIAN, M.; JING, L.; QI, L.; DACHENG, Y. A study on Quality of Experience for adaptive streaming service. *Proceedings of 2013 IEEE International Conference on Communications Workshops (ICC)*, Budapest (Hungria), p. 682–686, Jun. 2013.
- YOU, J.; REITER, U.; HANNUKSELA, M. M.; GABBOUJ, M.; PERKIS, A. Perceptual-based quality assessment for audio-visual services: A survey. *Signal Processing: Image Communication*, v. 25, n. 7, p. 482–501, Ago. 2010.
- ZEPERNICK, H.-J.; ENGELKE, U. Quality of Experience of Multimedia Services: Past, Present, and Future. *Adjunct Proceedings of the 9th European Interactive TV Conference (EuroITV'09)*, Lisboa (Portugal), p. 115–119, Jun./Jul. 2011.



PUBLICAÇÕES

CARVALHO, P. D. O.; COUTO, J. L. O.; MAIA, O. B.; LEMOS NETO, A. C.; GUIMARÃES, F. O.; DE ERRICO, L. Um Sistema de Engenharia de Tráfego Auto-Gerenciável para Transmissão de Conteúdos Multimídia em Redes IP. *III Workshop em Sistemas Distribuídos Autônomo (WoSIDA)*, Brasília (Brasil), p. 1-4, Maio 2013.

MAIA, O. B.; YEHA, H. C.; DE ERRICO, L. A concise review of the quality of experience assessment for video streaming. *Computer Communications*, v. 57, p.1-12, 2015.

TERMO DE CONSENTIMENTO LIVRE E INFORMADO

Prezado(a) Senhor(a),

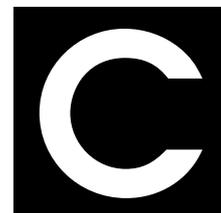
Gostaríamos de convidá-lo(a) a participar de nosso estudo sobre qualidade de vídeos transmitidos para dispositivos móveis. Trata-se de uma tese de doutorado, desenvolvido por Orlewilson Bentes Maia e orientado pelos professores Luciano de Errico (Laboratório de Redes, Tecnologias e Serviços de Comunicação a Distância – LabCOM), e Hani Camille Yehia (Centro de Estudos da Fala, Acústica, Linguagem e Música – CEFALA), do Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da Universidade Federal de Minas Gerais (UFMG).

Garantimos que a qualquer momento da realização desse estudo, qualquer participante e/ou estabelecimento envolvido poderá receber esclarecimentos adicionais que julgar necessários. Qualquer participante selecionado(a) poderá recusar-se a participar ou retirar-se da pesquisa em qualquer fase da mesma, sem nenhum tipo de penalidade, constrangimento ou prejuízo aos mesmos. Informamos que todas as informações prestadas por você são sigilosas e serão tratadas anonimamente quando utilizadas para fins desta pesquisa. As possíveis utilizações das informações são estritamente acadêmico: artigos científicos, relatórios de pesquisa e publicação de livros.

Agradecemos antecipadamente a sua colaboração e informamos que em caso de dúvidas, encaminhe um e-mail para Orlewilson Maia – orlewilson.maia@gmail.com.

Atenciosamente,

Orlewilson Bentes Maia



FORMULÁRIO SOBRE OS DADOS PESSOAIS DO(A) VOLUNTÁRIO(A)

1. Nome:

2. Idade (anos):

3. Gênero:

() Masculino () Feminino

4. E-mail para contato:

5. Escolaridade:

() Ensino Fundamental Incompleto. Qual o nome do curso?

() Ensino Fundamental Completo. Qual o nome do curso?

() Ensino Médio Incompleto. Qual o nome do curso?

() Ensino Médio Completo. Qual o nome do curso?

() Graduação Incompleta. Qual o nome do curso?

() Graduação Completa. Qual o nome do curso?

() Mestrado Incompleto. Qual o nome do curso?

() Mestrado Completo. Qual o nome do curso?

() Doutorado Incompleto. Qual o nome do curso?

() Doutorado Completo. Qual o nome do curso?

6. Qual a sua ocupação?

7. Em qual cidade/estado você mantém residência?

8. Você assiste a vídeos provenientes da Internet como, por exemplo, YouTube ou Netflix?

Sim Não

8.1 Caso assista, faz uso de quais dispositivos para assistir vídeos provenientes da Internet (você pode marcar mais de uma opção)?

– Televisão:

Não Sim. Por quanto tempo (minutos) ao dia, em média?

– Computador/laptop:

Não Sim. Por quanto tempo (minutos) ao dia, em média?

– Smartphone:

Não Sim. Por quanto tempo (minutos) ao dia, em média?

– Tablet:

Não Sim. Por quanto tempo (minutos) ao dia, em média?

9. Quando surge alguma falha durante a exibição do vídeo (por exemplo, indicação da imagem de carregando ou erros durante a exibição), por até quanto tempo você esperaria a normalidade da exibição do vídeo, nos seguintes casos:

– O serviço de transmissão de vídeos é gratuito, como o YouTube. (marque somente uma opção)

Até 5 segundos Até 10 segundos Até 30 segundos

Até 60 segundos Não esperaria

– O serviço de transmissão de vídeos é pago, como o Netflix. (marque somente uma opção)

- Até 5 segundos Até 10 segundos Até 30 segundos
 Até 60 segundos Não esperaria

Após responder o questionário, favor enviar para orlewilson.maia@gmail.com.



TESTE DE ACUIDADE VISUAL

O teste de acuidade visual teve como finalidade avaliar a limitação da visão do(a) candidato(a) com relação a detalhes espaciais (forma e contorno) de objetos. Neste trabalho, utilizou-se a tabela de Snellen de tamanho reduzido (folha A4), conforme ilustrada na Fig. 83.

O(a) candidato(a) ficou posicionado a 3 metros de distância da tabela de Snellen e depois foi explicado o procedimento de avaliação:

1. Permanecer usando o óculos de grau, caso tenha;
2. Ficar sentado em uma cadeira de frente à tabela;
3. Cobrir o olho esquerdo;
4. Ler em voz alta a letra indicada com uma régua até alcançar a última linha;
5. Repetir o mesmo procedimento com o olho direito coberto.

A pessoa estava apta para o teste de deficiência visual de cores quando ela conseguia ler todas as letras. Caso contrário, ela apresentava alguma limitação visual.

ACUVUE[®]
BRAND CONTACT LENSES

The Snellen Test for Visual Acuity

R O H K

20/100

N H O V C Z

20/70

C O R V Z D V H

20/50

R Z C V H

K S D N V

20/30

K D V H

V C R O

20/20

Figura 83: Tabela de Snellen utilizada nas avaliações de acuidade visual.

Fonte: [Acuvue \(2015\)](#).



TESTE DE DEFICIÊNCIA VISUAL DE CORES

O teste de deficiência visual de cores visa averiguar se uma pessoa possui alguma deficiência visual de origem congênita para cores, também conhecida como daltonismo. Para isso, utiliza-se o teste de Ishihara (ISHIHARA, 1972). Ele é constituído de placas onde estão expostos vários círculos com cores ligeiramente diferentes. No centro dessas placas, estão presentes alguns números ou algumas linhas.

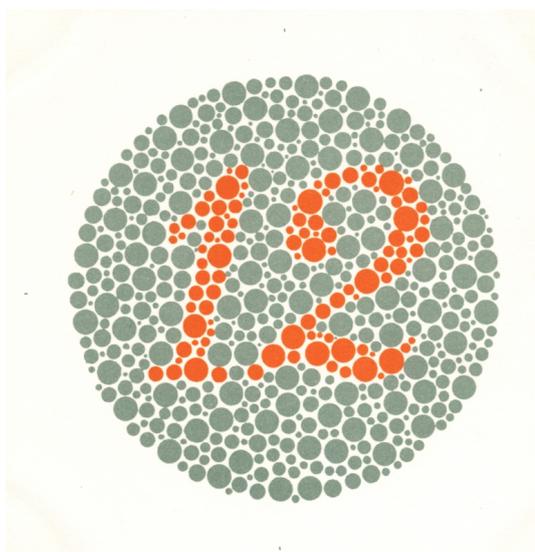
Neste trabalho, adotou-se o exame simplificado do teste de Ishihara, o qual consiste na leitura de seis placas escolhidas aleatoriamente de acordo com os grupos definidos na Tabela 37. Nela, são apresentadas as respostas de acordo com o grau de deficiência de cores.

Durante o exame, a visão da pessoa foi considerada normal para cores quando ela conseguia identificar os números e as linhas mostrados nas seis placas. Caso contrário, ela possuía alguma deficiência visual de cor. Nas Figs. 84 – 86 são mostradas as placas usadas neste trabalho.

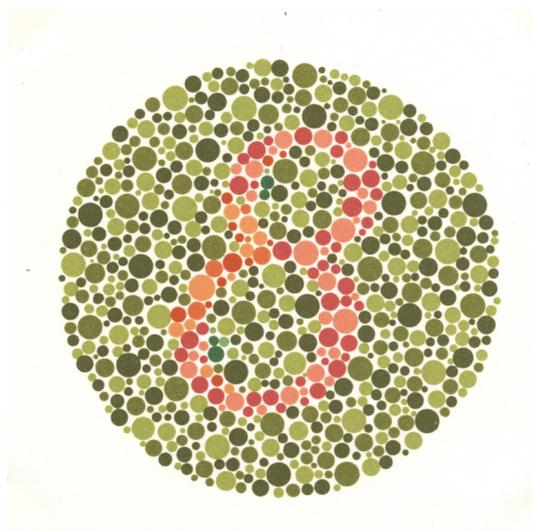
Tabela 37: Valores lidos de acordo com a deficiência visual de cores.

Grupo	Placa	Pessoa normal	Pessoa com daltonismo parcial (cor vermelha e verde)	Pessoa com daltonismo
1	1	12	12	12
2	2	8	3	X ²
	3	29	70	X ²
3	4	5	2	X ²
	5	3	5	X ²
	6	15	17	X ²
	7	74	21	X ²
4	8	6	X ²	X ²
	9	45	X ²	X ²
5	10	5	X ²	X ²
	11	7	X ²	X ²
	12	16	X ²	X ²
	13	73	X ²	X ²
6	14	X ²	5	X ²
	15	X ²	45	X ²

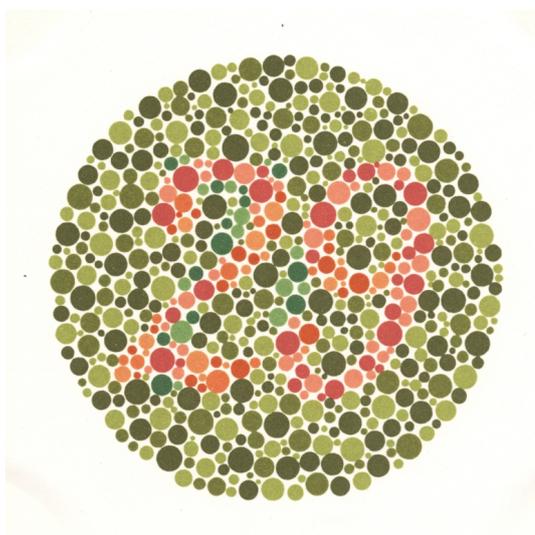
²A pessoa não consegue identificar o número ou a linha.



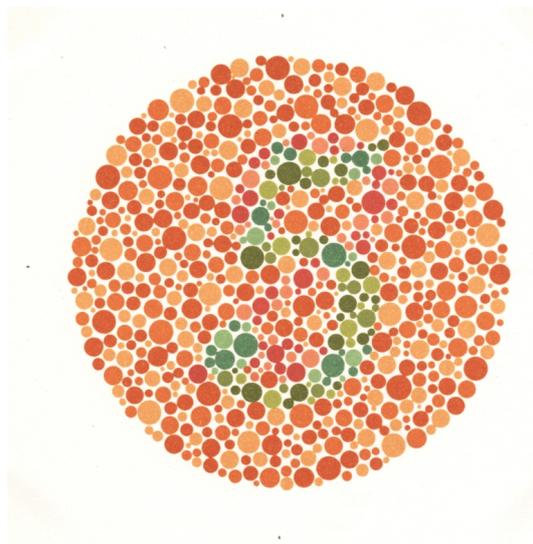
(a) Placa 1



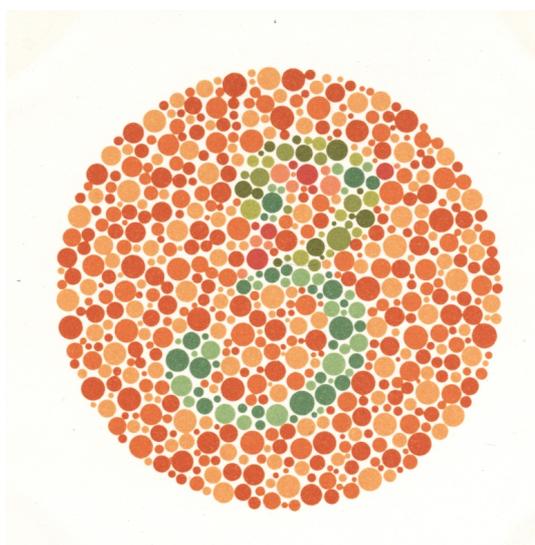
(b) Placa 2



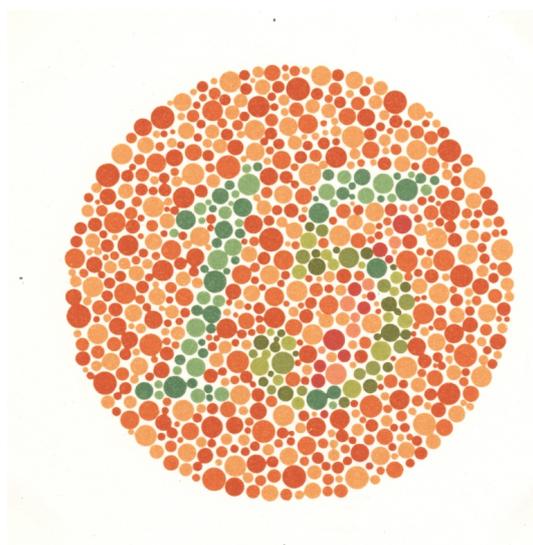
(c) Placa 3



(d) Placa 4

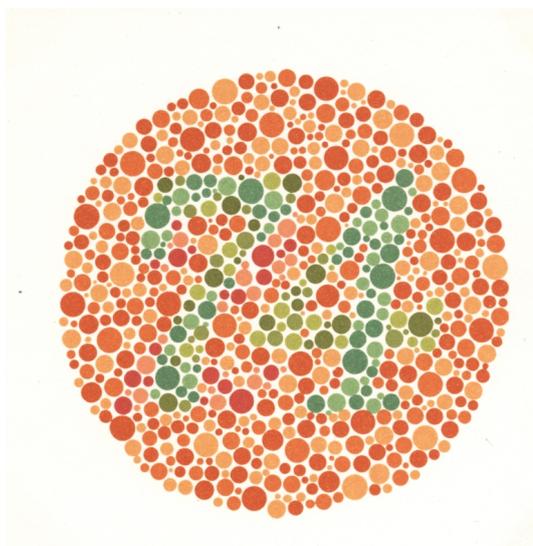


(e) Placa 5

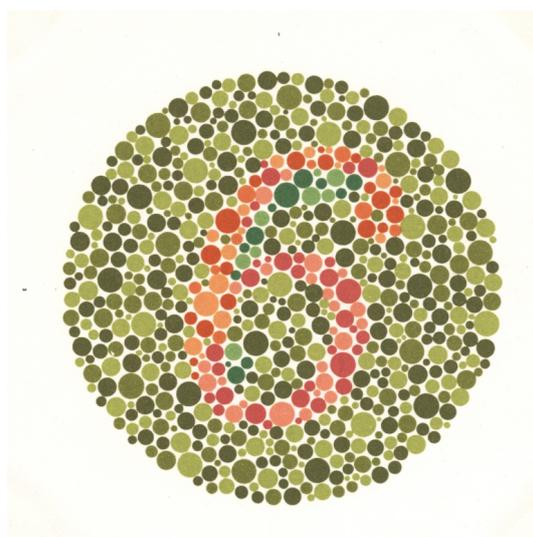


(f) Placa 6

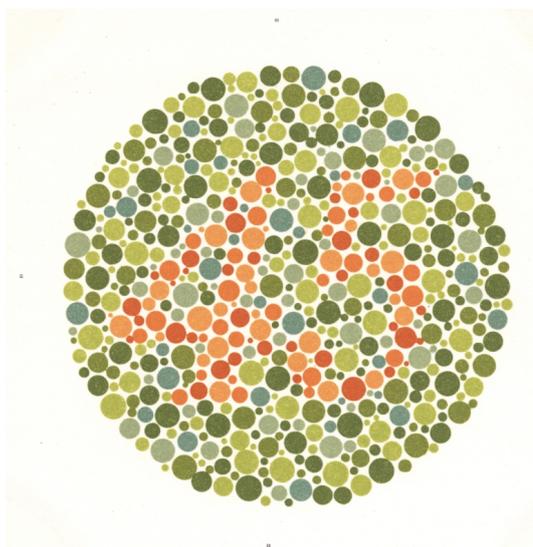
Figura 84: Placas utilizadas no exame simplificado do teste de Ishihara (1 de 3).



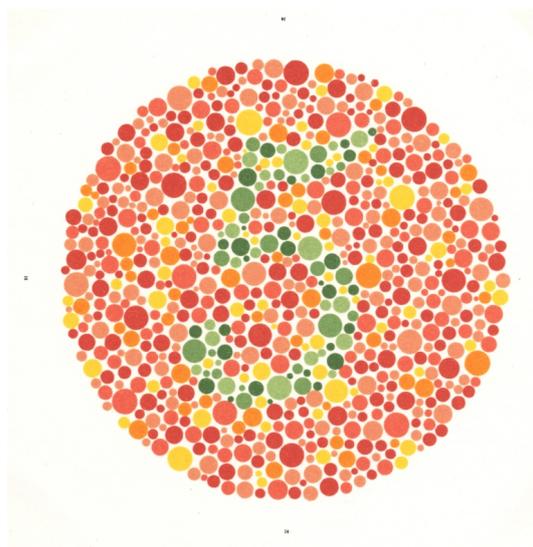
(a) Placa 7



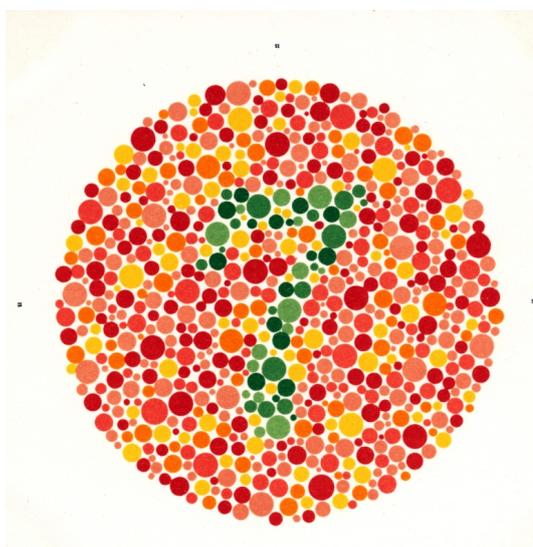
(b) Placa 8



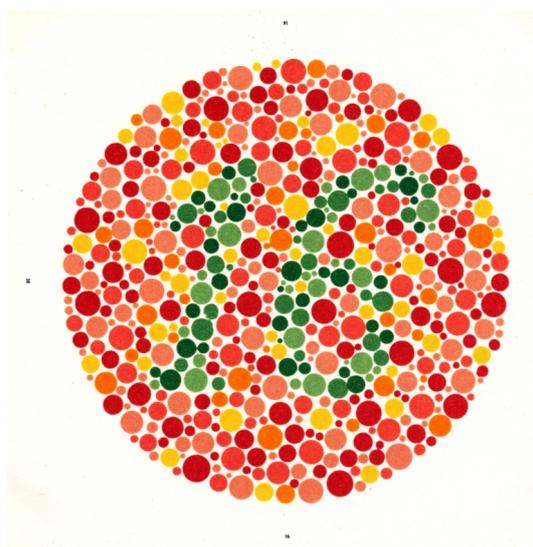
(c) Placa 9



(d) Placa 10



(e) Placa 11



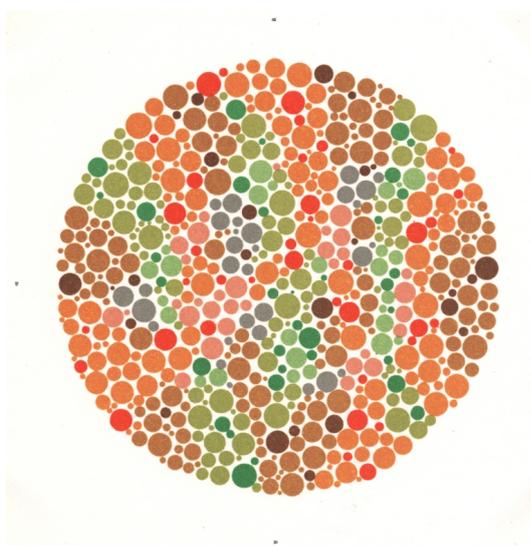
(f) Placa 12

Figura 85: Placas utilizadas no exame simplificado do teste de Ishihara (2 de 3).



(a) Placa 13

(b) Placa 14



(c) Placa 15

Figura 86: Placas utilizadas no exame simplificado do teste de Ishihara (3 de 3).