

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL - CPGEI

EMMERSON SANTA RITA DA SILVA

**ESPECIFICAÇÃO, PROJETO E DESENVOLVIMENTO DE
FERRAMENTAS DE AUXÍLIO À AVALIAÇÃO SUBJETIVA DE
VÍDEO**

DISSERTAÇÃO DE MESTRADO

**CURITIBA
2009**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

EMMERSON SANTA RITA DA SILVA

**ESPECIFICAÇÃO, PROJETO E DESENVOLVIMENTO DE
FERRAMENTAS DE AUXÍLIO À AVALIAÇÃO SUBJETIVA DE
VÍDEO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do grau de “Mestre em Ciências” – Área de Concentração: Telemática.

Orientador: Prof. Dr. Keiko Verônica Ono Fonseca

**CURITIBA
2009**

INCLUA NESSA PÁGINA A FICHA CATALOGRÁFICA

DEDICATÓRIA

Às memórias de Wendel Santa Rita, meu irmão amado, que mostrou com a sua curta vida como se ama e se vive plenamente nesse mundo.

Às minha vozinhas, Cândida Santa Rita e Rosa Pereira, pelo amor incondicional e pelos exemplos de vida e de amor.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a Deus, Senhor de todas as coisas, que guia as nossas vidas e guiou a minha também nesse caminho.

À minha família, pilar e sustentáculo, onde sempre encontro guarida e amparo nas horas difíceis, e que mesmo à distância não deixaram de acreditar em mim por nenhum momento.

Aos novos amigos que fiz aqui no LCD e LOOP, pois sem a contribuição de todos e, em especial, aos meus queridos amigos Rodolfo Patyk pela ajuda na montagem do dispositivo e Wyllian Bezerra da Silva que sempre apoiou e me auxiliou no caminho até aqui, sem esquecer-me do Diogo Pereira e de Lucieli Rossi pela ajuda na reta final.

Ao meu velho amigo Sergio Bezerra por acreditar em mim e contribuir para que eu chegasse ao final dessa etapa.

Um agradecimento especial a D. Wilce, mulher trabalhadora e mãe amorosa que me acolheu em sua casa como se acolhe a um filho, sem nunca ter me visto na vida, pois sem o seu apoio nada teria acontecido da forma que aconteceu.

Aos meus orientadores, prof. Dr. Keiko V. Ono e prof. Dr. Alexandre Pohl, por acreditarem em mim e terem me tomado nas mãos e, como um oleiro, transformarem-me de um estudante em um pesquisador. Sem a confiança e aposta que vocês fizeram, essa vitória que também é de vocês, não teria acontecido.

Ao UNINORTE, IFAM e CAPES pelo investimento financeiro, pois sem isso seria impossível chegar até aqui.

À UTFPR por me receber e me proporcionar esse avanço educacional.

Uma vitória dessas nunca é individual, ela é um feito coletivo e louvo a Deus por cada um que contribuiu para isso e serei agradecido pelo resto de minha vida.

Muito obrigado...

EPÍGRAFE

"... Ora, não percebeis que com os olhos alcançais toda a beleza do mundo? O olho é o senhor da astronomia e o autor da cosmografia; ele desvenda e corrige toda a arte da humanidade; conduz os homens às partes mais distantes do mundo; é o príncipe da matemática, e as ciências que o têm por fundamento são perfeitamente corretas.

O olho mede a distância e o tamanho das estrelas; encontra os elementos e suas localizações; ele... deu origem a arquitetura, a perspectiva, e a divina arte da pintura... Que povos, que línguas poderão descrever completamente sua função! O olho é a janela do corpo humano pela qual ele abre os caminhos e se deleita com a beleza do mundo."

LEONARDO DA VINCI (1452-1519)

RESUMO

SILVA, E.S.R. Especificação, Projeto e Desenvolvimento de Ferramentas de Auxílio à Avaliação Subjetiva de Vídeo, 2009. 118 f. Dissertação (Mestrado em Telemática) - Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná, Curitiba, 2009.

Este trabalho descreve a especificação, projeto e desenvolvimento de ferramentas de auxílio à avaliação subjetiva de vídeos. A especificação se baseou no estudo dos métodos propostos na norma ITU-R_BT.500-11 para avaliação de vídeo digital sem compressão degradado por artefatos gerados sinteticamente. O controle da geração de artefatos permitiu produzir diferentes níveis de degradação de vídeo e avaliar o impacto visual em relação àqueles encontrados em situações reais. Os resultados obtidos e observações realizadas determinaram a especificação de requisitos de um ambiente de avaliação de vídeo digital e a implementação de ferramentas de *hardware/software* para o auxílio nas avaliações subjetivas realizadas nesse ambiente

Palavras-Chave: Qualidade de vídeo. Vídeo digital. Avaliação subjetiva da qualidade de vídeo.

ABSTRACT

SILVA, E.S.R. Specification, Design and Development of a Support Tool for Subjective Evaluation of Video, 2009. 118 f. Dissertação (Mestrado em Telemática) - Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná, Curitiba, 2009.

This dissertation describes the specification, design and development of tools for subjective assessment of videos. The specification is based on a study of methods proposed in ITU-R_BT.500-11 for evaluation of uncompressed digital videos were which degraded by synthetic artifacts. The results and observations of a previous study led to the specification of the requirements for an environment designed for the assessment of digital video. Based on the achieved specification, tools (hardware / software) were developed to help subjective evaluations of digital videos. The new environment and its tools were tested and their results analyzed.

Keywords – *Digital Video, Quality Metrics, Video Artifacts, Evaluation Tools.*

LISTA DE FIGURAS

FIGURA 1 – REPRESENTAÇÃO ESQUEMÁTICA DO SISTEMA VISUAL HUMANO.....	20
FIGURA 2 – O SISTEMA ÓTICO DO OLHO HUMANO	21
FIGURA 3 – ESPECTRO ELETROMAGNÉTICO.....	22
FIGURA 4 – REPRESENTAÇÃO DE CONTRASTE SIMULTÂNEO.....	23
FIGURA 5 – CONTRASTE ENTRE DUAS REGIÕES	23
FIGURA 6 – EFEITO <i>MACH BAND</i>	24
FIGURA 7 - DIFERENTES COMPRIMENTOS DE ONDA ASSOCIADOS A DIFERENTES CORES	24
FIGURA 8 – CORES PRIMÁRIAS E SECUNDÁRIAS DA LUZ E DE PIGMENTOS	25
FIGURA 9 - AMOSTRAS DO FORMATO <i>YCbCr</i> : (A) 4:4:4, (B) 4:2:2 E (C) 4:2:0.....	26
FIGURA 10 – DEGRADAÇÕES INSERIDAS POR BLOCOS-DCT (B) E <i>WAVELET</i> (C).....	28
FIGURA 11 – EXEMPLO DO EFEITO DE BLOCAGEM.....	29
FIGURA 12 – EXEMPLO DE <i>STAIRCASE EFFECT</i> (EFEITO ESCADA).....	30
FIGURA 13 – EXEMPLO DE <i>RINGING EFFECT</i>	31
FIGURA 14 – EXEMPLO DE <i>MOSAIC PATTERN</i> (PADRÃO MOSAICO).....	32
FIGURA 15 – REPRESENTAÇÃO GRÁFICA DAS MÉTRICAS DE QUALIDADE DE VÍDEO	33
FIGURA 16 – ESCALA UTILIZADA NO MÉTODO DSIS	37
FIGURA 17 - ESCALA UTILIZADA NO MÉTODO DSCQS	38
FIGURA 18 – ESCALA UTILIZADA PELO SDSCE (ADAPTADO DE (ITU-R, 2002)).....	39
FIGURA 19 – APRESENTAÇÃO DOS ESTÍMULOS NO MÉTODO ACR	40
FIGURA 20 – ESCALA DE CATEGORIAS UTILIZADA PELO MÉTODO ACR.....	40
FIGURA 21 – APRESENTAÇÃO DOS ESTÍMULOS NO MÉTODO DCR	41
FIGURA 22 – APRESENTAÇÃO DOS ESTÍMULOS NO MÉTODO PC	42
FIGURA 23 – COMPARAÇÃO ENTRE UM VÍDEO ORIGINAL E DOIS VÍDEOS DEGRADADOS	45
FIGURA 24 – DIAGRAMA DO SISTEMA DE MEDIDA DA SIMILARIDADE ESTRUTURAL.	46
FIGURA 25 – PROPOSTA DE DESENVOLVIMENTO DA FERRAMENTA DE AVALIAÇÃO SUBJETIVA	59
FIGURA 26 – FLUXOGRAMA DE ATIVIDADES DESENVOLVIDAS NA AVALIAÇÃO SUBJETIVA PRELIMINAR.....	61
FIGURA 27 – AMBIENTE DE AVALIAÇÃO UTILIZADO	63
FIGURA 28- VÍDEOS ORIGINAIS UTILIZADOS NOS EXPERIMENTOS.....	63
FIGURA 29 – FORMA DE APRESENTAÇÃO DAS SEQUÊNCIAS DE VÍDEOS.....	66
FIGURA 30 – ESCALA UTILIZADA PELO SDSCE (ADAPTADO DE (ITU-R, 2002)).....	67
FIGURA 31 – TESTE DE ACUIDADE VISUAL (A) E EXEMPLO DE TESTE DE <i>ISHIHARA</i>	68
FIGURA 32 - DIAGRAMA EM BLOCO DO PROCESSO DE VOTAÇÃO	73
FIGURA 33 – VISÃO DO GERAL DO SISTEMA	73
FIGURA 34 – BRAÇO DESENVOLVIDO PARA A AUTOMATIZAÇÃO DAS MEDIDAS SUBJETIVAS	75
FIGURA 355 – CAMADAS DA APLICAÇÃO	77
FIGURA 366 – DIAGRAMA DE CLASSES DA ARQUITETURA.....	78
FIGURA 377 – TECNOLOGIAS ENVOLVIDAS NA IMPLEMENTAÇÃO DO SISTEMA.....	79
FIGURA 388 – DIAGRAMA DE CASOS DE USO DA APLICAÇÃO	81
FIGURA 39 – DIAGRAMA DE CLASSE DA APLICAÇÃO	82
FIGURA 400 – DIAGRAMAS DE SEQUÊNCIA.....	83
FIGURA 411 – DIAGRAMA DE ESTADOS DA APLICAÇÃO.....	83
FIGURA 422 – INTERFACE PRINCIPAL DA APLICAÇÃO.....	84
FIGURA 433 – MODELO ENTIDADE-RELACIONAMENTO DA APLICAÇÃO.....	86
FIGURA 444 – MODELO RELACIONAL DA APLICAÇÃO.....	88
FIGURA 455 – DIAGRAMA EM BLOCO DO <i>FIRMWARE</i>	89

FIGURA 466 – PROCESSO DE AVALIAÇÃO DOS VÍDEOS NO AMBIENTE PROPOSTO.....	91
FIGURA 477 – VISÃO GERAL DO AMBIENTE DE AVALIAÇÃO.....	92
FIGURA 488 – AVALIADOR FAZENDO UMA AVALIAÇÃO UTILIZANDO O MÉTODO SDSCE DE VÍDEOS NO FORMATO CIF.	93
FIGURA 49 – VISÃO GERAL DA AVALIAÇÃO NO AMBIENTE PROPOSTO	94
FIGURA 50 – SESSÃO DE AVALIAÇÃO SUBJETIVA UTILIZANDO AVALIAÇÃO MANUAL.....	96
FIGURA 51 - RELACIONAMENTO ENTRE O MOS E O PSNR COM BASE NAS DEGRADAÇÕES PARA O ARTEFATO DE BORRAMENTO	97
FIGURA 522 - RELACIONAMENTO ENTRE O MOS E O PSNR COM BASE NAS DEGRADAÇÕES PARA O ARTEFATO DE RÚIDO GAUSSIANO.....	98
FIGURA 53 - RELACIONAMENTO ENTRE O MOS E O PSNR COM BASE NOS VÍDEOS UTILIZADOS NO EXPERIMENTO (A, B, C E D).	100
FIGURA 54 – AVALIAÇÕES SUBJETIVAS UTILIZANDO O AMBIENTE DE AVALIAÇÃO E SUAS FERRAMENTAS.....	101
FIGURA 55 – GRÁFICOS DO MOS X MSSIM DOS VÍDEOS (A) E (B).....	103
FIGURA 56 - RELACIONAMENTO ENTRE O MOS E O MSSIM COM BASE NAS DEGRADAÇÕES PARA O ARTEFATOS DE RÚIDO GAUSSIANO (A) E BORRAMENTO (B).....	104

LISTA DE TABELAS

TABELA 1 – FORMATO DOS QUADROS DE VÍDEO.....	27
TABELA 2 – SELEÇÃO DE MÉTODOS DE TESTE (ADAPTADO DE (ITU-R, 2002))	36
TABELA 3 – CONFIGURAÇÕES DO MICROCOMPUTADOR UTILIZADO NOS TESTES	81
TABELA 4 – RESULTADO DOS TESTES DE USABILIDADE	105

LISTA DE QUADROS

QUADRO 1 – CONDIÇÕES GERAIS DE VISUALIZAÇÃO E APRESENTAÇÃO DOS TESTES	56
QUADRO 2 – PROCESSAMENTO DOS VÍDEOS UTILIZADOS NOS TESTES	64
QUADRO 3 – SITUAÇÕES DOS TESTES E DEGRADAÇÕES	65

LISTA DE ABREVIATURAS E SIGLAS

1080p	- Formato de vídeo 1920x1080 pixels a 60 <i>frames</i> completos por segundo.
API	- <i>Application Programming Interface</i> (Interface de Programação de Aplicativos)
AWT	- <i>Abstract Window Toolkit</i> (Caixa de ferramentas abstratas para Windows)
CCIR	- <i>Consultative Committee for International Radio</i> (Comitê Consultivo Internacional de Rádio)
CI	- Circuito Integrado
CIE	- <i>Commission Internationale de L'éclairage</i> (Comissão Internacional de Iluminação)
CIF	- <i>Common Intermediate Format</i> (Formato Intermediário Comum)
CRT	- <i>Cathode-Ray Tube</i> (Tubo de raios catódicos)
DCE	- <i>Data Communication Equipment</i> (Equipamento de Comunicação de Dados)
DCT	- <i>Discrete Cosine Transform</i> (Transformada Discreta do Cosseno)
DSCQS	- <i>Double Stimulus Continuous Quality Scale</i> (Escala Contínua de Qualidade com Duplo Estímulo)
DSIS	- <i>Double Stimulus Impairment Scale</i> (Escala de Imparidades com Duplo Estímulo)
DTE	- <i>Data Terminal Equipment</i> (Equipamento de Terminal de Dados)
DTV	- <i>Digital Television</i> (Televisão Digital)
DVB	- <i>Digital Video Broadcasting</i> (Radiodifusão de Vídeo Digital)
E/S	- Entrada e saída
EBU	- <i>European Broadcasting Union</i> (União de Radiodifusão Européia)
EPROM	- <i>Erasable Programmable Read-Only Memory</i> (Memória somente-leitura apagável e programável)
GPL	- <i>General Public License</i> (Licença Pública Geral)
HD	- <i>High Definition</i> (Alta Definição)
HDTV	- <i>High Definition Television</i> (Televisão de Alta Definição)
HVS	- <i>Human Visual System</i> (Sistema Visual Humano)
ITU	- <i>International Telecommunications Union</i> (União Internacional de Telecomunicações)
J2EE	- <i>Java Enterprise Edition</i> (Java Edição Empresarial)
JDBC	- <i>Java Data Base Connectivity</i> (Conectividade Java com Base de Dados)
JVM	- <i>Java Virtual Machine</i> (Máquina Virtual Java)
LCD	- <i>Liquid Crystal Display</i> (Monitor de Cristal Líquido)
LED	- <i>Light Emitting Diode</i> (Diodo Emissor de Luz)
MCUs	- <i>Microcontroller Unit</i> (Unidades de Microcontroladores)
MER	- Modelo Entidade-Relacionamento
MOS	- <i>Mean Opinion Score</i> (Média de Pontuação das Opiniões)
MPEG	- <i>Moving Picture Experts Group</i> (Grupo de Especialistas em Imagens com Movimento)
MPQM	- <i>Moving Pictures Quality Metric</i> (Métrica de Qualidade de Imagens em

	Movimento)
MSE	- <i>Mean Squared Error</i> (Erro Quadrático Médio)
MSSIM	- <i>Mean Structural SIMilarity Index</i> (Média do Índice de Similaridade Estrutural)
NTSC	- <i>National Television Systems Committee</i> (Comitê Nacional de Sitemas de Televisão)
ODBC	- <i>Open Data Base Connectivity</i> (Conectividade Aberta de Base de Dados)
OMG	- <i>Object Management Group</i> (Grupo de Gerenciamento de Objetos)
PAL	- <i>Phase Alternate Line</i> (Linha de Fase Alternada)
PC	- <i>Personal Computer</i> (Computador Pessoal)
Pixel	- <i>Picture Element</i> (Elemento de Imagem)
PSNR	- <i>Peak Signal to Noise Ratio</i> (Razão Sinal-ruído de Pico)
QCIF	- <i>Quarter Common Intermediate Format</i>
RAM	- <i>Random Access Memory</i> (Memória de Acesso Rândomico)
RGB	- Espaço de cores RGB (onde R = <i>Red</i> , G = <i>Green</i> , B = <i>Blue</i>)
RMSE	- <i>Root Mean Squared Error</i> (Erro Quadrático Médio Principal)
ROM	- <i>Read Only Memory</i> (Memória de Somente de Leitura)
RS - 232	- Padrão adotado pela indústria para troca de dados binários na forma serial assíncrona
SAASVD	- Sistema de Auxilio à Avaliação Subjetiva de Vídeo Digital
SDSCE	- <i>Simultaneous Double Stimulus for Continuous Evaluation</i> (Avaliação Contínua com Duplo Estímulo Simultâneo)
SECAM	- <i>Séquentiel Couleur à Mémoire</i> (Cor Sequencial com Memória)
SGBD	- Sistema de Gerenciamento de Banco de Dados
SQL	- <i>Structured Query Language</i> (Linguagem de Consulta Estruturada)
SSCQE	- <i>Single Stimulus Continuous Quality Evaluation</i> (Avaliação Contínua da Qualidade com Estímulo Simples)
SSIM	- <i>Structural SIMilarity Index</i> (Índice de Similaridade Estrutural_
UCS	- Escala de Cromaticidade Uniforme
UHF	- <i>Ultra High Frequency</i> (Frequência Ultra Alta)
UML	- <i>Unified Modeling Language</i> (Linguagem de Modelagem Unificada)
VQEG	- <i>Video Quality Experts Group</i> (Grupo de Especialistas em Qualidade de Vídeo)

SUMÁRIO

1. INTRODUÇÃO.....	15
1.1. MOTIVAÇÕES	15
1.2. OBJETIVOS	16
1.2.1 OBJETIVO GERAL	16
1.2.2 OBJETIVOS ESPECÍFICOS	16
1.3. ESTRUTURA DA DISSERTAÇÃO	17
1.4. PUBLICAÇÕES	18
2 FUNDAMENTAÇÃO TEÓRICA	19
2.1 SISTEMA VISUAL HUMANO	20
2.2 ARTEFATOS	27
2.3 DESCRIÇÃO DAS MÉTRICAS SUBJETIVAS E OBJETIVAS UTILIZADAS.....	32
2.3.1 MÉTRICAS SUBJETIVAS.....	35
2.3.2 MÉTRICAS OBJETIVAS DE QUALIDADE DE IMAGEM/VÍDEO.....	43
2.4 TRABALHOS RELACIONADOS.....	47
2.5 RESUMO E CONCLUSÃO DO CAPÍTULO	53
3 ESTUDO DE CASO PRELIMINAR DE AVALIAÇÃO SUBJETIVA	54
3.1 FUNDAMENTOS METODOLÓGICOS DA AVALIAÇÃO.....	54
3.1.1 CONDIÇÕES DE OBSERVAÇÃO	55
3.1.2 INFLUÊNCIA DA DISTÂNCIA E DO TAMANHO DA IMAGEM NA OBSERVAÇÃO.....	56
3.1.3 ANÁLISE DE VARIABILIDADE ENTRE AVALIADORES	56
3.1.4 MATERIAL DE TESTE.....	57
3.1.5 FORMA DE APRESENTAÇÃO DAS SEQUÊNCIAS DE VÍDEO	57
3.1.6 ESCALAS.....	58
3.1.7 ESTIMAÇÃO DA QUALIDADE SUBJETIVA	59
3.2 EXPERIMENTO DE AVALIAÇÃO SUBJETIVA PRELIMINAR	60
3.2.1 LIMITAÇÕES E CONSIDERAÇÕES.....	62
3.2.2 MATERIAL DE TESTE.....	63
3.2.3 PROCESSAMENTO.....	64
3.2.4 SITUAÇÕES DE TESTE.....	65
3.2.5 PREPARAÇÃO DO MATERIAL.....	65
3.2.6 FORMA DE APRESENTAÇÃO	66
3.2.7 DETALHES DO TESTE.....	66
3.3 RESUMO E CONCLUSÃO DO CAPÍTULO	68
4 MODELAGEM E PROJETO DO AMBIENTE E SUAS FERRAMENTAS.....	70
4.1 DESENVOLVIMENTO.....	70
4.1.1 ANÁLISE DE REQUISITOS	70
4.2 ESPECIFICAÇÃO TÉCNICA	72
4.3 ASPECTOS FUNCIONAIS	74
4.4 ARQUITETURA DO SISTEMA PROPOSTO	76
4.5 PROJETO	80
4.6 PROJETO DO BANCO DE DADOS.....	84

4.6.1	REQUISITOS DO SISTEMA DE BANCO DE DADOS DO SAASVD	84
4.6.2	MODELO ENTIDADE-RELACIONAMENTO E MODELO RELACIONAL	85
4.7	PROJETO DE FIRMWARE.....	89
4.8	ESPECIFICAÇÃO DE TESTES DE VALIDAÇÃO DO PROJETO	90
4.9	DIAGRAMA DE BLOCOS FINAL DO SAASVD	90
4.10	VISÃO GERAL DO AMBIENTE DE AVALIAÇÃO SUBJETIVA.....	92
4.11	RESUMO E CONCLUSÃO DO CAPÍTULO	94
5	VALIDAÇÃO DO AMBIENTE E DE SUAS FERRAMENTAS.....	96
5.1	RESULTADOS PRELIMINARES COM AVALIAÇÃO MANUAL	96
5.2	ANÁLISE DOS RESULTADOS OBTIDOS COM O NOVO AMBIENTE.....	101
5.3	RESUMO E CONCLUSÃO DO CAPÍTULO	105
6	CONCLUSÕES E TRABALHOS FUTUROS.....	106
6.1	CONCLUSÕES	106
6.2	TRABALHOS FUTUROS	108
	REFERÊNCIAS	109

1. INTRODUÇÃO

1.1. MOTIVAÇÕES

Vídeos e imagens digitais estão sujeitas a uma grande variedade de distorções durante a aquisição, processamento, compressão, armazenamento, transmissão e a reprodução. Qualquer um destes processos pode resultar em degradações da qualidade visual do vídeo ou de imagens digitais (WANG et al, 2004), as quais são geralmente caracterizadas por artefatos visíveis no dispositivo de monitoração.

A avaliação da qualidade percebida pelo usuário é fundamental para orientar projetos de sistemas de imagem (PAPPAS e SAFRANEK, 2000). A visualização dos artefatos na imagem/vídeo, sua quantidade, frequência, local de aparecimento, tipo, contexto, herança cultural e experiência pessoal do espectador, áudio associado, entre outros, determinam a qualidade da imagem percebida pelo espectador humano. A avaliação da qualidade da imagem (ou vídeo) requer o estabelecimento ou adoção de métricas para capturar informações de interesse específico, por exemplo, em projeto de sistemas de monitoramento, de transmissão, de codificação/decodificação, de produção do vídeo etc. Neste sentido, métricas de avaliação de qualidade de imagem/vídeo são importantes para uma larga gama de aplicações, como, por exemplo, monitorar e ajustar a qualidade da imagem de um sistema de provisão de vídeo digital para a Internet.

As métricas de qualidade subjetiva, obtidas a partir de certo número de observadores humanos, têm sido consideradas por muitos como a mais confiável forma de medição de qualidade (KOUMARAS, PLIAKAS e KOURTIS, 2007). Uma análise estatística dos dados coletados dos observadores humanos resulta em uma nota MOS (*Mean Opinion Score – Média de Pontuação das Opiniões*) (ITU-R, 2002), a qual é utilizada para a avaliação da qualidade percebida. No entanto, esta abordagem é lenta e cara para a maioria das aplicações (ESKICIOGLU e FISHER, 1995). Todavia, existe um grande esforço no sentido de desenvolvimento de medidas quantitativas (WANDELL, 1995; LAMBRECHT, 1998; VQEG, 2003; VQEG, 2008) ou de ferramentas que minimizem custo e tempo de avaliação.

Por sua vez, métricas objetivas (WINKLER, 2005) têm por objetivo prever quanto da distorção da imagem/vídeo será observada por usuários. Exemplos dessas métricas são o MSE

(*Mean Squared Error – Erro Quadrático Médio*), PSNR (*Peak Signal to Noise Ratio – Razão Sinal-ruído de Pico*) e SSIM (*Structural SIMilarity Index – Índice de Similaridade Estrutural*) (WANG et al, 2004;WU e RAO, 2006) que exigem a imagem/vídeo original como referência para avaliar a imagem/vídeo degradado. O SSIM foi desenvolvido sobre as características do Sistema Visual Humano (SVH) e mantém melhor correlação com medidas subjetivas (WU et al, 2006), enquanto o PSNR e o MSE apresentam baixa correlação com resultados obtidos através da avaliação subjetiva (WANG et al, 2006).

Este trabalho apresenta um conjunto de ferramentas de auxílio à avaliação subjetiva que compõem um ambiente de análise de qualidade para experimentos em laboratórios. Através do seu uso pode-se avaliar/validar resultados de avaliações obtidos com métricas objetivas. O ambiente de avaliação é composto por: ferramenta de *hardware*, ambiente físico, gerador de artefatos artificiais e do *software* de compilação e análise dos dados. Os resultados obtidos em (ALBINI, 2009) de medidas objetivas realizadas em vídeos corrompidos por artefatos gerados artificialmente e avaliados pela SDSCE (*Simultaneous Double Stimulus for Continuous Evaluation – Duplo Estímulo Simultâneo para Avaliação Contínua*) conduziram à especificação, projeto e implementação das ferramentas supracitadas.

1.2. OBJETIVOS

1.2.1 Objetivo Geral

Desenvolver ferramentas de auxílio à avaliação subjetiva para teste ou validação de métricas subjetivas e objetivas de qualidade de vídeo digital.

1.2.2 Objetivos Específicos

- Especificar um ambiente para avaliação subjetiva de vídeo digital.
- Projetar um sistema para automatizar e minimizar o tempo necessário para a realização dos testes subjetivos.
- Desenvolver *hardware* para coleta de dados dos avaliadores.
- Especificar e implementar um sistema de representação dos resultados das avaliações subjetivas.

- Gerar gráficos das avaliações subjetivas e comparações com métricas objetivas.
- Analisar e discutir os resultados experimentais

1.3. ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está organizada em seis capítulos descritos a seguir.

No Capítulo 2 faz-se uma revisão da literatura sobre o Sistema Visual Humano e suas características de forma a descrever, sucintamente, aspectos fisiológicos e funcionais. Aspectos referentes aos vídeos digitais são abordados também e, para finalizar esse capítulo, descrevem-se as métricas objetivas, subjetivas e artefatos de vídeos digitais, pois os mesmos serão objetos de estudos neste trabalho. Ainda neste capítulo, fazemos uma breve discussão dos trabalhos relacionados com o objetivo de colocar o leitor na linha do tempo das pesquisas na área de avaliação da qualidade de vídeo digital e também no que diz respeito à importância da avaliação subjetiva.

O capítulo 3 descreve o estudo de caso que foi realizado em função da necessidade de coleta de requisitos para o desenvolvimento do ambiente e de suas ferramentas. Neste capítulo fundamenta-se metodologicamente a avaliação subjetiva e descrevem-se detalhadamente os testes subjetivos realizados baseados nas normas vigentes e literatura atual.

O Capítulo 4 apresenta a modelagem e o projeto das ferramentas, bem como apresenta o ambiente (composto de ambiente físico e ferramentas de *hardware/software*) fruto da pesquisa e dos experimentos de avaliação subjetiva. Aqui, todos os artefatos de desenvolvimento de *hardware* e *software* são apresentados, tais como diagramas UML, MER, Modelo Relacional, Interface do sistema, projeto de *firmware*. Este capítulo serve como guia de desenvolvimento e de melhorias para o sistema desenvolvido, uma vez que apresenta o projeto dos sistemas de *hardware/software* implementados.

Na sequência (Capítulo 5) apresentam-se os resultados obtidos nos dois experimentos realizados para este trabalho, tanto no que diz respeito à avaliação subjetiva preliminar quanto na avaliação automatizada proposta. Este capítulo mostra os resultados gerados pelo sistema e dos testes de usabilidade realizados para a avaliação das ferramentas apresentadas no ambiente de avaliação subjetiva. No capítulo 6, apresentam-se as conclusões e propostas de trabalhos futuros.

1.4. PUBLICAÇÕES

SILVA, Emmerson S. R. da, POHL, Alexandre de A. P.,ALBINI, F.L.P., FONSECA, Keiko V. O., **Projeto e Construção de um Ambiente de Auxílio à Avaliação Subjetiva de Vídeo.** SET2009, Sociedade Brasileira de Engenharia de Televisão, agosto de 2009.

SILVA, Emmerson S. R. da, FONSECA, Keiko V. O., POHL, Alexandre de A. P.,BEZERRA, Sérgio A. C., SILVA, Wyllian B. da., **Sistema de Auxílio à Avaliação Subjetiva de Vídeo Digital.** WebMedia2009, Simpósio Brasileiro de Sistemas Multimídia e Web, outubro de 2009.

2 FUNDAMENTAÇÃO TEÓRICA

A avaliação subjetiva aplica métodos para determinar o desempenho dos sistemas de vídeo utilizando medições que podem prever diretamente as reações dos usuários que utilizarão os sistemas testados. Em geral, classificam-se as avaliações subjetivas em dois tipos: as que servem para estabelecer avaliações do desempenho dos sistemas em condições ideais (chamadas avaliações de qualidade) e as necessárias para criar sistemas para manter qualidade sob condições ótimas que não dizem respeito às redes de transporte ou emissão (chamadas de avaliações de imparidade).

Especificamente em sistemas de difusão de sinais de televisão existe uma grande variedade de métodos de teste básicos para avaliações. Um resumo dos problemas de avaliação típicos e destes métodos de teste é apresentado na norma ITU (ITU-R, 2002). Por exemplo, a norma especifica que a avaliação deve durar até 30 minutos e contar com pelo menos 15 observadores, que devem ser instruídos sobre os detalhes da avaliação (quando avaliar, como os vídeos são apresentados, como funciona a escala e tipos de falha no vídeo) antes da mesma iniciar (ITU-R, 2002).

As avaliações verificam o impacto do aparecimento de artefatos sobre a percepção do usuário em relação a qualidade da imagem/vídeo. Exemplos de artefatos em vídeo analógico são o “chuveiro” e o “fantasma”, causados principalmente por ruído branco gaussiano e multi-caminho, respectivamente (SCHWARTZ e SHAW, 1975). Nos vídeos digitais são exemplos de artefatos a blocagem (*blocking*) e o borramento (*blurring*), decorrentes principalmente pela compressão da imagem (ZAMPOULO, GOMES e SEARA, 2008) utilizando-se a Transformada Discreta do Cosseno (DCT – *Discrete Cosine Transform*), a perda de informações da imagem (LIU e BOVIK, 2002), (ZAMPOULO, GOMES e SEARA, 2008), ou até mesmo devido à interpolação ou efeito zoom (AUTRUSSEAU e CALLET, 2005).

Nas seções que seguem apresentam-se os conceitos necessários ao entendimento do trabalho realizado.

2.1 SISTEMA VISUAL HUMANO

Os sistemas de codificação de vídeo digital buscam o compromisso entre a qualidade do que é transmitido e o que o ser humano considera agradável e aceitável, uma vez que o ser humano é o consumidor final. Portanto, a resposta do Sistema Visual Humano (SVH) é de vital importância para as medidas de qualidade de vídeo.

O SVH é um sistema complexo que capta, interpreta e responde aos estímulos visuais (WU e RAO, 2006). Pode-se representar o SVH como um conjunto de blocos, conforme mostrado na Figura 1, onde há elementos ópticos responsáveis pela formação da imagem anatômica e fisiologicamente responsáveis pelo controle de exposição, detectores responsáveis pela captura da luz, sua transformação em impulsos elétricos e respectivo processamento (WANG e BOVIK, 2006).

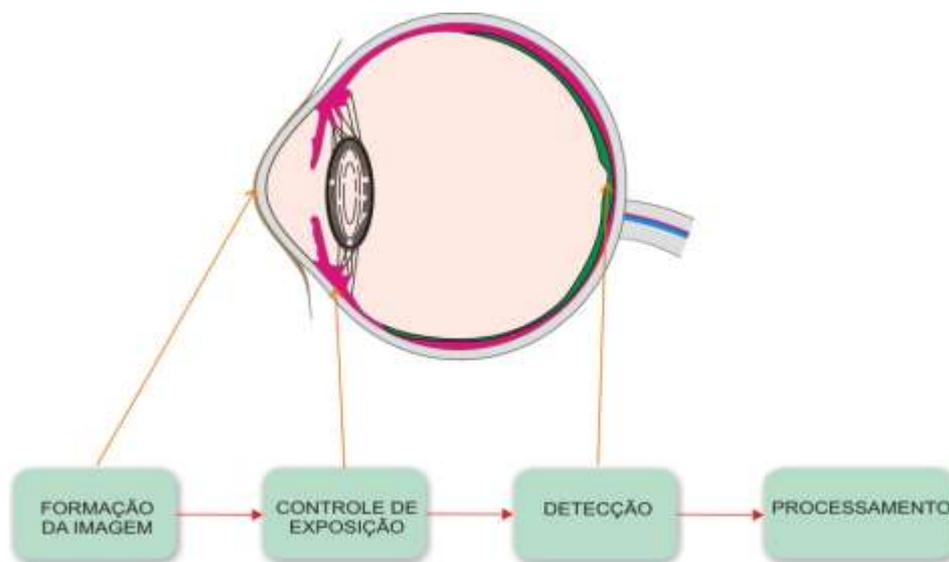


Figura 1 – Representação esquemática do sistema visual humano

Fonte: adaptado de (WANG e BOVIK, 2006).

No primeiro bloco do sistema a imagem de um objeto ou cena de vídeo é captada pelo sistema ocular e é projetada sobre a retina, que é uma membrana foto-sensível. Obtém-se como resultado uma imagem chamada de imagem de retina. A Figura 2 apresenta os elementos da óptica ocular e nessa figura pode-se destacar a córnea, a pupila e o cristalino (WANG e BOVIK, 2006).

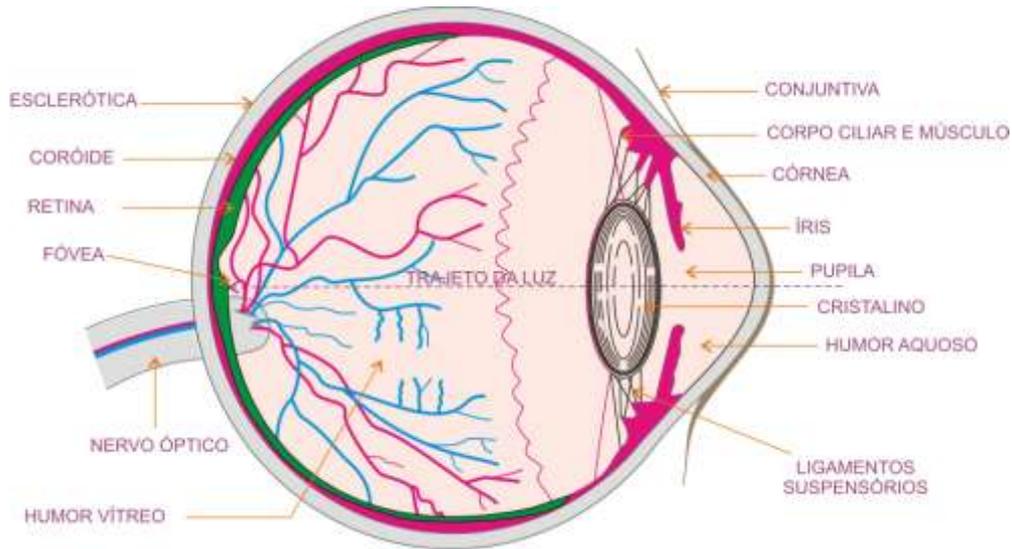


Figura 2 – O sistema ótico do olho humano

Fonte: adaptado de (NEWELL, 1992)

A retina tem papel fundamental na visão, pois a mesma é composta por células responsáveis pela identificação dos objetos, chamadas de células fotorreceptoras. Existem dois tipos de fotorreceptores: os cones e os bastonetes. Os cones são responsáveis pela visão em cores sob condições normais de iluminação (divididos nos tipos L (vermelho), M (verde) e S (azul)) e os bastonetes são responsáveis pela visão monocromática em condições de baixa luminosidade. Os sinais que chegam aos fotorreceptores passam por vários neurônios interconectados antes de serem transmitidos para as células que são responsáveis pela transmissão das informações visuais, chamadas de células ganglionares (WANG e BOVIK, 2006). A quantidade de bastonetes é 20 vezes maior que a de cones no olho humano (NEWELL, 1992), dessa forma, pode-se afirmar que o olho humano é mais sensível a luminância que a crominância. Além disso, devido a diferenças em torno de como as células fotorreceptoras (cones e bastonetes) captam as diversas frequências do espectro eletromagnético, o olho humano pode detectar ondas num espectro mais estreito desse espectro, aproximadamente de 380 a 760 nm conforme apresentado na Figura 3.

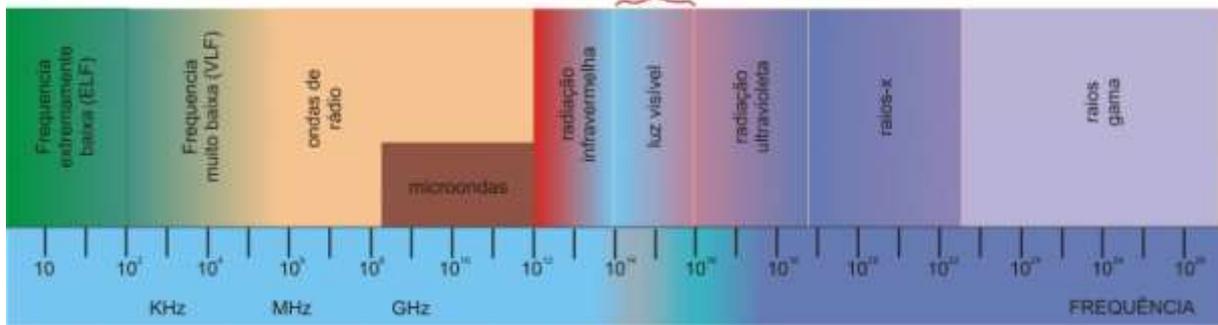


Figura 3 – Espectro eletromagnético

Fonte: adaptado de (ALENCAR, 2007)

O entendimento do funcionamento do SVH e a incorporação desse conhecimento ao processamento digital de imagens podem garantir melhores sistemas e índices de qualidade do ponto de vista do observador humano, que possui papel fundamental na caracterização da qualidade das imagens/vídeos. Além disso, segundo (WU e RAO, 2006) a sensibilidade do olho humano para distinguir diferenças de intensidade de luminosidade não depende somente dessa diferença mais também do nível de intensidade, uma vez que a resposta do olho a essas mudanças de intensidade luminosa não é linear.

O SVH é mais sensível a áreas com baixo nível de cinza e é indiferente a áreas de alto nível de cinza, por influência das condições de luminosidade do ambiente. O brilho de um monitor de vídeo afeta a sensibilidade a mudanças, logo, a sensibilidade do SVH a mudanças é diretamente proporcional ao brilho do monitor e os erros de quantização não são percebidos em faixas de baixa luminância, mas são percebidos em uma faixa de média luminância (WU e RAO, 2006). A principal razão é que nossa percepção é sensível ao contraste de luminância e não ao seu valor absoluto. Contraste é o termo geralmente usado para descrever diferenças em valores de luminância em duas áreas vizinhas.

A Figura 4 demonstra o contraste simultâneo, onde um círculo de uma dada luminosidade é colocado em três fundos diferentes. Quanto mais escuro for o fundo mais brilhante parece o círculo. A luminância percebida de um objeto, chamada brilho, depende da luminância das áreas circundantes, ou seja, do contraste entre o objeto e o fundo.

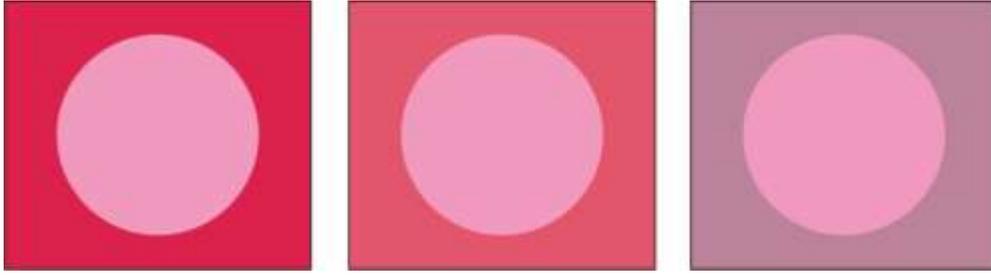


Figura 4 – Representação de contraste simultâneo

Fonte: adaptado de (PEDRINI e SCHWARTS, 2008)

A figura 5 mostra um círculo de luminância $I + \Delta I$ cercado por um retângulo mais largo com luminância I . Se ΔI for zero as duas regiões não são distinguíveis. Aumentando-se ligeiramente ΔI , para um determinado valor, há o contraste entre as duas regiões.

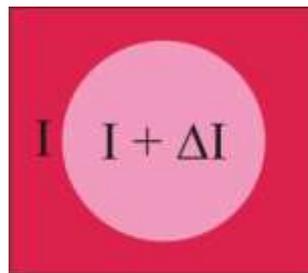


Figura 5 – Contraste entre duas regiões

Fonte: adaptado de (PEDRINI e SCHWARTS, 2008)

Bordas salientes em uma imagem confundem o SVH, como pode ser visto na Figura 6, ou seja, o SVH tem grande sensibilidade a pequenas diferenças na intensidade da energia luminosa. Em cada barra é mantida constante a intensidade dos tons de cinza, no entanto ao se observar a fronteira de uma barra com a outra o SVH percebe que essa parte (direita da fronteira) é mais escura do que o meio da barra e o contrário é observado na parte mais a esquerda, pois o SVH percebe essa área como sendo mais clara que o meio da barra. Este efeito visível nas arestas designa-se por *Mach Band* (PEDRINI e SCHWARTS, 2008). A transição de uma faixa para outra é enfatizada pela sensibilidade ao contraste do SVH (WU, RAO, 2006).



Figura 6 – Efeito *Mach Band*

Fonte: adaptado de (PEDRINI e SCHWARTS, 2008)

A utilização das cores no processamento digital de imagens se destaca por possibilitar a identificação de objetos e a sua extração de uma cena, bem como na análise realizada por seres humanos, uma vez que o olho humano é capaz de discernir milhares de tons e intensidades de cores (GONZALEZ e WOODS, 2000). Nesse trabalho, o foco é na análise realizada pelos seres humanos, uma vez que se trata do desenvolvimento de ferramentas de auxílio à avaliação subjetiva.

O espectro de cores pode ser dividido em seis regiões, como mostra a Figura 7: violeta, azul, verde, amarelo, laranja e vermelho. Na realidade, na decomposição das cores não se tem uma definição do final de uma cor, mas sim um degradê, que é uma mistura suave de uma cor com a outra (GONZALEZ e WOODS, 2000).

	Violeta	380 a 440 nm
	Azul	440 - 490 nm
	Verde	490 - 565 nm
	Amarelo	565 - 590 nm
	Laranja	590 - 630 nm
	Vermelho	630 - 780 nm

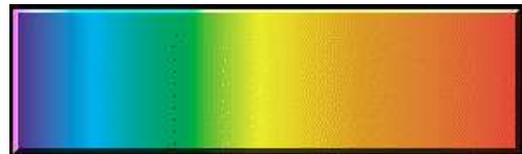


Figura 7 - Diferentes comprimentos de onda associados a diferentes cores

Fonte: adaptado de (GONZALEZ e WOODS, 2000).

Uma fonte de luz cromática possui três características: radiância (energia total que flui de uma fonte de luz), luminância (quantidade de energia percebida por um observador) e o brilho (um descritor subjetivo e essencial para a descrição das sensações de cores).

Devido à disposição das células fotoreceptoras presentes no SVH, todas as cores são percebidas como combinações das cores primárias (R - *Red* = vermelho, G - *Green* = verde e B - *Blue* = azul). Com a necessidade de padronização o CIE (Comissão Internacional de Iluminação – “*Comission Internationale de l’Eclairage*”) (CIE, 2009) designou em 1931 os seguintes valores de comprimentos de onda para as cores primárias: azul = 435,8 nm, verde = 546,1 nm e vermelho = 700 nm. As cores primárias podem ser adicionadas para gerar novas cores, chamadas de cores secundárias. As TVs em cores utilizam-se da natureza aditiva das cores da luz para gerar as imagens produzidas na recepção. A Figura 8 apresenta a mistura das cores primárias para geração das cores secundárias e as cores primárias de pigmentos (GONZALEZ e WOODS, 2000).

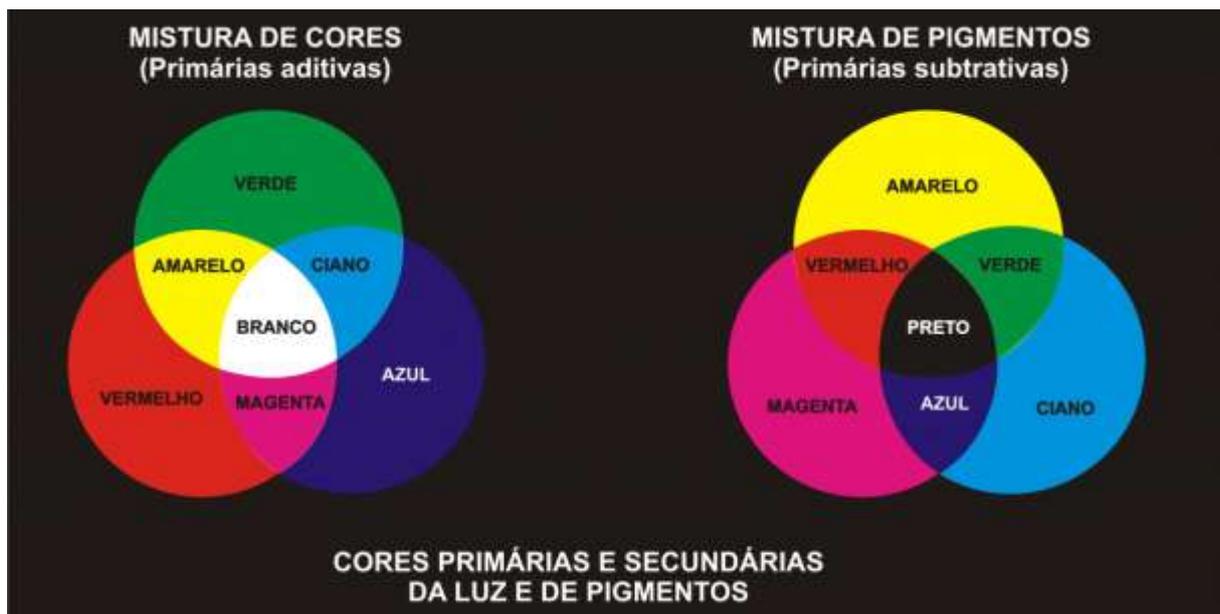


Figura 8 – Cores primárias e secundárias da luz e de pigmentos

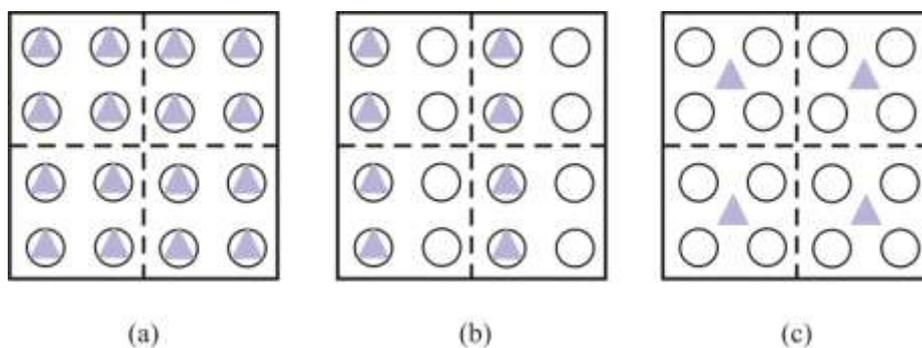
Fonte: adaptado de (GONZALEZ e WOODS, 2000).

Esta seção não tem a pretensão de esgotar ou de aprofundar a discussão sobre o SVH, uma vez que este assunto é multidisciplinar e extrapola os objetivos deste trabalho. O que se pretende é destacar as características mais importantes do ponto de vista do observador humano e sua relação com a percepção dos artefatos de vídeo.

FORMATOS DE VÍDEO

Os pixels podem ser amostrados de diversas maneiras, para exemplificar pode-se citar a representação que utiliza as componentes de cores primárias, vermelho (R – *Red*), verde (G – *Green*) e azul (B – *Blue*), formando o padrão RGB, e o outro é o formado pelas componentes Y, Cb e Cr, onde Y representa a luminância e Cb e Cr a crominância (cor).

Na tentativa de se obter vídeos que fossem de qualidade com a menor quantidade de *bits*, verificou-se a partir do estudo das características do SVH e da teoria das cores que necessariamente não era necessário amostrar as três faixas de cores como no padrão RGB e sim um percentual aceitável desse espectro para as cores, uma vez que se verificou que o SVH é mais sensível a luminância que a crominância. Os sinais de crominância toleram menos resolução na codificação sem perda significativa de informação para o SVH. Pode-se, portanto, diminuir a taxa de bits na codificação de um sinal de vídeo ao se utilizar componente de luminância e crominância ao invés de RGB. A Figura 9 (a), (b) e (c) apresentam, respectivamente, os vídeos amostrados nos formatos 4:4:4, 4:2:2 e 4:2:0. O formato 4:4:4 representa as componentes Y, Cb e Cr amostrados com taxas iguais e para cada *pixel* existe uma amostra de cada componente. No formato 4:2:2, são amostrados com a metade da resolução horizontal e com a mesma resolução vertical. Ou seja, para cada 4 componentes de luminância na horizontal existem duas amostras Cb e duas Cr. Já o formato 4:2:0 não segue a mesma lógica de formação dos formatos citados. Neste formato as amostras Cb e Cr possuem metade da resolução horizontal e vertical da luminância (RICHARDSON, 2003).



Legenda:
 ○ Luminância (Y)
 ▲ Crominância (Cb = U e Cr = V)

Figura 9 - Amostras do formato YCbCr: (a) 4:4:4, (b) 4:2:2 e (c) 4:2:0

Fonte: Adaptado de (RICHARDSON, 2003)

No caso do projeto desenvolvido, o formato dos vídeos adotados foi o *Common Intermediate Format* (CIF). O CIF foi definido para uso em sistemas de videoconferência. O formato é derivado de SIF (*Source Intermediate Format*) e também base de um conjunto de formatos de vídeo conforme apresentados na Tabela 2.

Tabela 1 – Formato dos quadros de vídeo

Formatos de vídeo	Resolução de luminância (horizontal × vertical)	Bits por quadro (4:2:0, oito bits por amostra)	Aplicação
Sub-QCIF	128x96	147456	Aplicações multimídia móveis
Quarter CIF (QCIF)	176x144	304128	Videoconferência e aplicações Multimídia móveis
CIF	352x288	1216512	Videoconferência
4CIF	704x576	4866048	Padrão de televisão e DVB

Fonte: adaptado de (RICHARDSON, 2003)

2.2 ARTEFATOS

Nesta seção, apresentam-se os artefatos e suas manifestações em vídeos digitais. O propósito é descrever os principais tipos de artefatos, suas características e como os mesmos são formados.

Quantificar os artefatos em termos do seu impacto visual é uma tarefa muito difícil. Devido à complexidade do SVH (conforme seção 2.1), que ainda não foi satisfatoriamente modelado, a distorção percebida não é diretamente proporcional ao erro de quantização absoluto, mas também está sujeita às características espaço-temporais locais e globais das sequências de vídeo (WU e RAO, 2005).

Os artefatos se diferem para cada sistema, tanto para o digital quanto para o analógico. Os artefatos podem ser introduzidos através da filtragem, conversão, transformação, quantização e/ou transmissão (WU e RAO, 2005; WANG e BOVIK, 2006). Blocagem, Borramento, erro de quantização são algumas degradações introduzidas no processo de compressão. Essas

degradações são influenciadas pelo algoritmo de compressão utilizado, conteúdo das imagens/vídeos, origem da fonte e das condições do canal de transmissão. A Figura 31 apresenta artefatos gerados por compressão DCT e *Wavelet*.



(a) Original (b) bloco-DCT (c) *Wavelet*

Figura 10 – Degradações inseridas por blocos-DCT (b) e *Wavelet* (c)

Fonte: (WINKLER, 2005)

Na Figura 10 o efeito de blocagem é visível na parte inferior de (b); o efeito *staircase* (escada) pode ser visto ao redor da borda branca inclinada do farol em (b). O borramento é evidente em ambas as imagens comprimidas; O efeito *ringing* pode ser observado em torno de contornos e bordas.

Observa-se que embora os vídeos comprimidos sejam geralmente degradados em relação ao original, alguns processos de filtragem podem reduzir o número de degradações existentes no vídeo fonte.

Efeito de Blocagem (Blocking Effect)

O efeito de blocagem é a distorção mais extensamente estudada (SHANIKA et al,1995; MONTEZ e BECKER, 2004; WU e RAO, 2005) decorre da codificação em blocos de imagens.

Visualmente, o efeito é visto como uma descontinuidade entre blocos adjacentes em uma imagem. As bordas retas induzidas pelo limite de blocos descontinuados e o espaçamento regular dos blocos resulta no efeito de blocagem. A Figura 11 apresenta um exemplo típico de blocagem.



Figura 11 – Exemplo do efeito de Blocagem

Fonte: (WU e RAO, 2006)

O efeito de blocagem é definido aqui como as descontinuidades encontradas nos limites dos blocos adjacentes em uma imagem reconstruída (BOVIK, 2000; PAPPAS e SAFRANEK, 2000; WU e RAO, 2006), uma vez que é uma prática comum medir e reduzir o efeito de blocagem levando-se em conta os *pixels* nos limites do bloco. Porém, o termo pode ser generalizado para significar uma diferença global entre blocos adjacentes, mas caracterizam-se estas diferenças como artefatos individuais.

Este tipo de degradação se faz presente em cenas complexas ou quando elevadas taxas de compressão são utilizadas. Para dada granularidade de quantização a blocagem é mais perceptível nas áreas mais homogêneas de cada quadro do vídeo (FARIA E SILVA, 1998).

Efeito de Borramento (Blurring Effect)

O borramento se manifesta como uma perda de detalhes espaciais em regiões de moderada a alta atividade espaciais nas regiões da imagem, como em áreas com texturas ou ao redor das bordas de objeto da cena (WU e RAO, 2005). Um exemplo de borramento pode ser observado na Figura 10.

Para codificação *intraframe* de macroblocos, o borramento está diretamente relacionado à supressão dos coeficientes de ordem mais alta da DCT por meio da quantização.

O borramento é um artefato caracterizado pela perda das componentes de alta frequência (quantização pouco refinada). As componentes de alta frequência têm menos energia e são menos perceptíveis ao SVH. Na prática, para uma dada taxa, há um compromisso entre alta resolução e baixo ruído, sendo esse ajuste subjetivo e dependente da aplicação.

O borramento ou erros de crominância em vídeos digitais comprimidos podem ser mais perceptíveis que as demais degradações, uma vez que os sistemas empregam uma subamostragem na crominância. Na presença de objetos coloridos na cena em movimento rápido, a subamostragem da crominância pode levar a um descasamento entre luminância e crominância, causando, com isso, perda de resolução e tornando os artefatos mais visíveis (SILVA, 1998).

Efeito Escada (Staircase Effect)

O efeito escada está relacionado a ambos os efeitos de padrão de blocagem e mosaico em termos da manifestação de descontinuidade entre blocos adjacentes. Quando uma extremidade diagonal é representada dentro de uma série de blocos sucessivos, a consequência da quantização grosseira é a reconstrução da borda diagonal como vários passos horizontais ou verticais. Estes passos individuais não se fundem suavemente no limite do bloco com a continuação da extremidade em blocos adjacentes. A Figura 12 apresenta um exemplo do *staircase effect* (efeito escada).

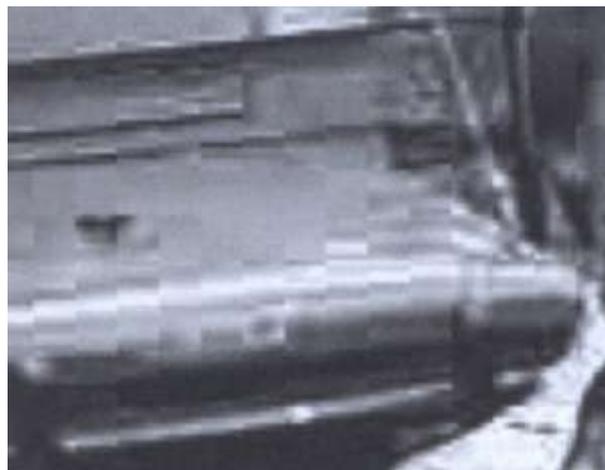


Figura 12 – Exemplo de *Staircase Effect* (Efeito Escada)

Fonte: (WU e RAO, 2006)

Observa-se na Figura 12 que as discontinuidades acontecem nos limites dos blocos. O efeito escada apresenta significância, especialmente, em torno das bordas e perto do topo da imagem (WU e RAO, 2006).

Efeito Vibrante (Ringing Effect)

O efeito *ringing* é mais evidente junto aos contrastes altos da borda em áreas de textura geralmente suave na reconstrução, e aparece como uma ondulação fora da borda até os limites das cercanias do bloco. A Figura 13 apresenta um exemplo de *ringing effect*.



Figura 13 – Exemplo de *Ringing Effect*

Fonte: (WU e RAO, 2006)

Pode se notar na Figura 13 que o efeito *ringing* é mais evidente ao longo da borda das mesas de tênis e na parte inferior do braço do jogador. Em geral, a textura suave em torno de blocos resulta em uma maior visibilidade do efeito *ringing*, onde de outra forma um efeito de mascaramento seria introduzido (WU e RAO, 2006).

Padrão Mosaico (Mosaic Pattern)

A natureza geral do efeito de padrão de mosaico é a aparente divergência entre todo, ou parte, do conteúdo de blocos adjacentes. Este tem um efeito semelhante a usar visualmente má montagem de azulejos em um mosaico. Isto pode significar um bloco com certo contorno ou textura dissimilar aos blocos circundantes, ou um bloco usado na representação de um objeto que não mistura satisfatoriamente com os outros blocos constituintes.

O padrão de mosaico, tipicamente, coincide com o efeito de blocagem. Porém, a existência do efeito de blocagem entre dois blocos não necessariamente implica a manifestação do padrão de mosaico entre os mesmos dois blocos. Por exemplo, uma área suave com textura é altamente suscetível ao efeito de blocagem, mas as características suaves dos blocos adjacentes não induziriam um padrão de mosaico. A Figura 14 apresenta um exemplo do padrão mosaico.



Figura 14 – Exemplo de *Mosaic Pattern* (Padrão Mosaico)

Fonte: (WU e RAO, 2006)

Na Figura 14 o padrão mosaico é mais evidente em torno do rosto do personagem e adjacente às bordas horizontais perto da janela do automóvel.

2.3 DESCRIÇÃO DAS MÉTRICAS SUBJETIVAS E OBJETIVAS UTILIZADAS

Esta seção descreve o processo utilizado na avaliação da qualidade visual em conformidade com as recomendações ITU (ITU-R, 2002). Tal processo desenvolve-se nas seguintes fases: (i) formação de uma base de dados obtida experimentalmente em que avaliadores são solicitados a emitir opiniões pessoais sobre a qualidade dos vídeos exibidos; (ii) Comparação com outras métricas objetivas propostas na literatura. Além disso, descreveremos as métricas objetivas e subjetivas utilizadas no desenvolvimento do trabalho.

Recomendações internacionais para avaliação subjetiva da qualidade de vídeos digitais (ITU-R, 2002) incluem especificações para apresentar muitos tipos diferentes de testes

subjetivos. Cada qual com suas peculiaridades e diferenças. A Figura 15 apresenta uma representação gráfica das métricas de qualidade de vídeo.

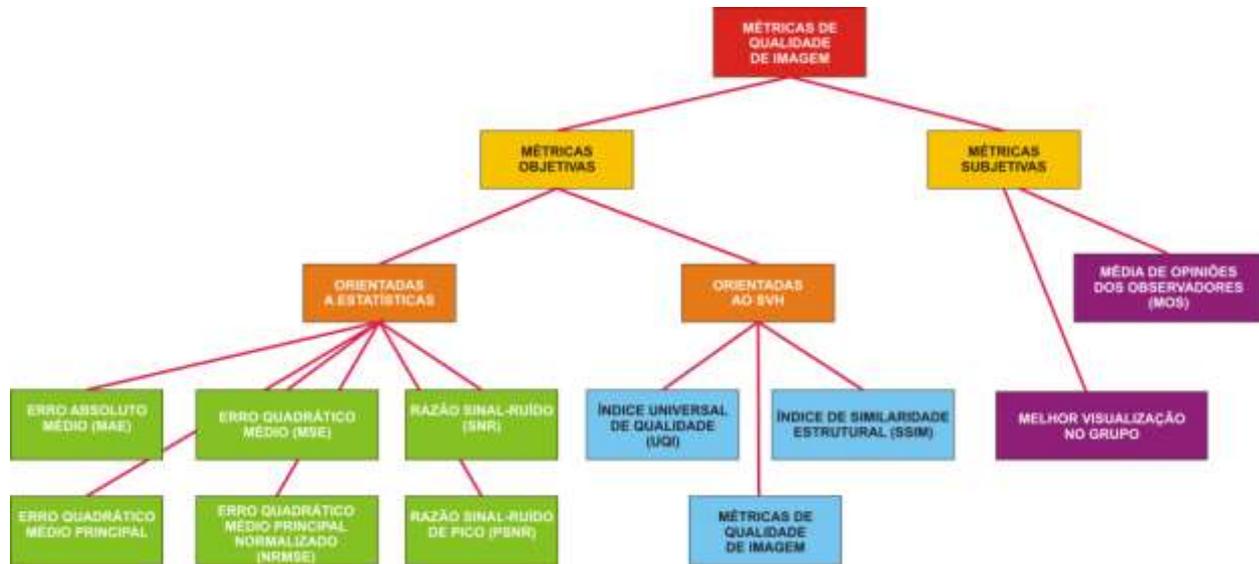


Figura 15 – Representação gráfica das métricas de qualidade de vídeo

Fonte: (KONNIK, 2009)

No diagrama apresentado na Figura 15 dividem-se as métricas em objetivas e subjetivas. As métricas objetivas dividem-se em Orientadas a Estatísticas e Orientadas ao SVH, e as avaliações subjetivas são baseadas nas opiniões de usuários.

As medidas de qualidade de imagem e vídeo desempenham um papel importante em uma variedade de aplicações de processamento de imagem e vídeo. Muito frequentemente a qualidade de um vídeo precisa ser quantificada. Isto pode ser feito por sessões de prova subjetivas, ou por métricas objetivas – métricas computacionais (WINKLER, 2005).

Em sistemas de processamento de imagem/vídeo destinados ao usuário final, deseja-se que os parâmetros utilizados na medição da qualidade sejam otimizados para aumentar a qualidade visual percebida pelos mesmos. Levando-se isso em consideração, a avaliação subjetiva é a melhor forma de avaliação da qualidade, uma vez que o ser humano será o consumidor final das imagens/vídeos (WANG e BOVIK, 2004). Ainda com relação à avaliação subjetiva, (KOUMARAS, PLIAKAS e KOURTIS, 2007) asseguram que o MOS, que é calculado a partir das notas dos avaliadores, é o método mais preciso de medida de qualidade

visual percebida. Contudo, tal estratégia, se por um lado representa a qualidade percebida, por outro exige tempo considerável e demandam um elevado custo.

Uma alternativa à avaliação subjetiva é a avaliação objetiva em que a avaliação humana é substituída por uma função matemática, produzindo uma avaliação automática da qualidade visual, reduzindo o tempo e o custo em relação à avaliação subjetiva. Esses modelos matemáticos, também conhecidos como métricas podem ser perceptuais e não perceptuais. As perceptuais apresentam baixo custo computacional, mas apresentam baixa correlação com a qualidade visual percebida pelos usuários, enquanto as não perceptuais apresentam maior correlação, porém agregado a isso, maior custo computacional.

As métricas perceptuais e não perceptuais podem ser classificadas segundo WANG e BOVIK (2006); WU e RAO (2006); KOUMARAS, PLIAKAS, COURTIS (2007), em:

- a) Referência Completa (*Full-Reference*) – quando uma imagem de referência é utilizada para fins de comparação com uma imagem/vídeo cuja qualidade se quer avaliar.
- b) Referência Reduzida (*Reduced-Reference*) – quando se dispõe de alguns parâmetros da imagem de referência, as quais são comparadas com o mesmo conjunto de parâmetros da imagem/vídeo cuja qualidade se quer avaliar.
- c) Sem Referência (*No-Reference*) – quando a avaliação da qualidade da imagem/vídeo é realizada apenas usando a imagem/vídeo que se quer avaliar a qualidade.

Através da avaliação subjetiva de vídeo é possível captar e medir características percebidas da imagem/vídeo diretamente, enquanto que esse componente subjetivo falta às avaliações objetivas, uma vez que as mesmas são somente capazes de medir parâmetros físicos da imagem/vídeo (SILVA, 1998).

Devido à necessidade de se melhorar as métricas objetivas e de se caracterizar e introduzir parâmetros referentes ao SVH, a utilização de uma abordagem não consiste na não adoção de outra, uma vez que o próprio VQEG (2008) concluiu em seu último relatório que para se validar métricas objetivas era necessário a validação utilizando métricas subjetivas. O problema na utilização das métricas subjetivas ainda se foca nos custos de tempo e a necessidade de avaliadores humanos.

Cuidados devem ser tomados com as condições de teste, cultura e experiência dos observadores humanos e idade pois esses fatores podem afetar a avaliação subjetiva, tendo em vista suas características. Além desses fatores o projeto do teste também pode influenciar os resultados dos testes subjetivos.

A seguir, descrevem-se as métricas subjetivas de acordo com a (ITU-R, 2002) e as métricas objetivas MSE, PSNR, SSIM e MSSIM.

2.3.1 Métricas Subjetivas

A percepção humana da qualidade de um vídeo é afetada pela fidelidade espaço-temporal. A fidelidade espacial refere-se às partes do cenário que podem ser vistas nitidamente e a temporal refere-se ao fato do movimento aparecer natural e suave.

Em adição a isso, a opinião dos observadores sobre a qualidade é afetada por outros fatores tais como o meio visual, estado mental do observador, experiência em avaliação de vídeos e o nível de interação do observador com o cenário visual. Experiências mostram que, se o ambiente visual é confortável, a opinião de um observador sobre a qualidade visual de um vídeo tende a ser mais alta do que se o ambiente é desconfortável. Além desses fatores a atenção visual e o efeito recente (memória recente) influenciam na percepção da qualidade visual. Em função de todos esses aspectos, a medição da qualidade de vídeo digital de forma precisa em termos quantitativos e qualitativos torna-se bastante difícil.

Avaliação subjetiva são métodos usados para determinar o desempenho dos sistemas de transmissão/recepção utilizando medições que podem antecipar diretamente as reações das pessoas que possam ver os sistemas testados. Neste contexto, entende-se que através dessa avaliação pode não ser possível caracterizar plenamente o desempenho do sistema por meios objetivos e, por conseguinte, é necessário completar as medições objetivas com as medições subjetivas.

Em geral, existem duas classes de avaliações subjetivas. Em primeiro lugar, existem as que servem para estabelecer avaliações do desempenho dos sistemas em condições ideais, estes são tipicamente chamados avaliações de qualidade. Em segundo lugar, existem avaliações em que a capacidade de criar sistemas para manter a qualidade sob condições ótimas que não dizem respeito às redes de transporte ou de emissão, estes, na maioria das vezes, são chamadas de avaliações de imparidade (ITU-R, 2002).

Para conduzir adequadamente avaliações subjetivas, é necessário primeiro selecionar entre as opções disponíveis aquelas que melhor se adequam aos objetivos e circunstâncias do problema na avaliação em foco.

Testes subjetivos são parte importante na etapa de projeto de codificadores. Mesmo em situações em que medidas subjetivas que apresentam boa relação com a qualidade percebida da

imagem/vídeo estão disponíveis, os testes subjetivos são usados para complementar, verificar ou calibrar tais medidas objetivas.

O objetivo dos testes subjetivos é estabelecer, através de métodos empíricos, bases para os julgamentos. É importante que o método de medida usado conduza a resultados válidos e representativos das opiniões.

Uma grande variedade de métodos de testes básicos foi usada em avaliações de televisão. Na prática, porém, métodos particulares deviam ser usados para tratar problemas de avaliação particulares. Ilustra-se na Tabela 3 os métodos de avaliação subjetiva típicos referenciados em (ITU-R, 2002).

Tabela 2 – Seleção de métodos de teste (adaptado de (ITU-R, 2002))

Métodos	Problemas de avaliação
<i>DSIS – Double Stimulus Impairment Scale</i> (Escala de Imparidades com Duplo Estímulo)	Medidas de robustez do sistema (isto é, características de falha)
<i>DSCQS – Double Stimulus Continuous Quality Scale</i> (Escala Contínua de Qualidade com Duplo Estímulo)	Medidas de qualidade de sistemas relativos a uma referência, medição da qualidade de codificação de imagens estereoscópicas
<i>SSCQE – Single Stimulus Continuous Quality Evaluation</i> (Avaliação Contínua da Qualidade com Estímulo Simples)	Medidas de qualidade de sistemas sem referência
<i>SDSCE – Simultaneous Double Stimulus for Continuous Evaluation</i> (Avaliação Contínua com Duplo Estímulo Simultâneo)	Medir a fidelidade entre duas sequências de vídeos prejudicados, comparar diferentes ferramentas de recuperação de erros.

Fonte: (ITU-R, 2002)

Em seguida apresenta-se uma breve descrição baseada em (ITU-R, 2002) e em (ITU-T, 2008) das principais métricas subjetivas.

2.3.1.1 Métodos apresentados pela norma ITU-R BT.500.11 (ITU-R, 2002)

Os métodos apresentados na norma ITU-R BT.500-11 (ITU-R, 2002) que serviram de referência para esse trabalho são apresentados a seguir.

DSIS – Double Stimulus Impairment Scale (Escala de Imparidades com Duplo Estímulo)

Nesse método, uma sequência de vídeo é apresentada ao avaliador, sendo mostrada sempre aos pares: a primeira como referência (sem falhas) e a segunda contendo algum defeito, devendo o observador avaliar o vídeo imperfeito.

Assim, o avaliador assistirá a um vídeo sem defeitos, e em seguida verá o mesmo vídeo com defeitos, para então avaliar a qualidade do segundo tendo com referência o primeiro. O vídeo padrão e o vídeo testado devem ficar à mostra durante 10 segundos cada.

Essa avaliação se dá dentro da escala apresentada na Figura 16:

5 imperceptível
4 perceptível, mas não irritante
3 levemente irritante
2 irritante
1 bastante irritante

Figura 16 – Escala utilizada no método DSIS

Deste modo, o observador deve votar em uma das cinco opções de acordo com a opinião sobre o vídeo defeituoso apresentado. Os vídeos devem ser apresentados numa ordem aleatória e de preferência, nunca serem repetidos de uma avaliação para outra.

Finalmente, as falhas devem variar em intensidade de modo que, ao final das avaliações, a média se aproxime do conceito 3 (levemente irritante).

DSCQS – *Double Stimulus Continuous Quality Scale* (Escala Contínua de Qualidade com Duplo Estímulo)

Nesse método, bem como no DSIS, os vídeos são apresentados aos pares, sendo uma proveniente diretamente da fonte e a outra contendo os defeitos.

Entretanto, enquanto no método DSIS o observador sabia qual era o vídeo de referência e deveria medir somente o vídeo defeituoso, nesse caso o par de vídeos vem em uma ordem aleatória e o observador deve avaliar ambas as imagens.

O observador deve assistir aos dois filmes, um após o outro, em uma ordem sempre aleatória. Esse processo se repete um número de vezes de acordo com o tipo de vídeo: caso seja mais estático, são 5 repetições, onde cada vídeo é mostrada entre 3 e 4 segundos, e a votação se dá nas duas últimas; caso contrário, ou seja, o vídeo envolva movimento, devem ocorrer somente 2 repetições, sendo cada um dos vídeos mostrados durante 10 segundos, e a votação deve acontecer na segunda repetição.

A votação se dá através de uma marca dentro de uma escala contínua apresentada na Figura 17:

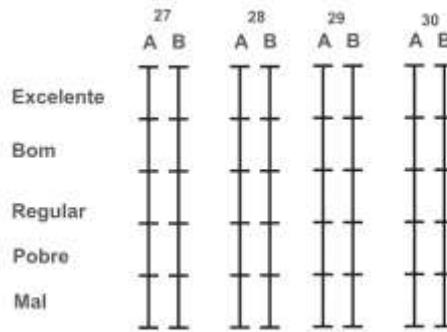


Figura 17 - Escala utilizada no método DSCQS

Como visto, para cada par de vídeos, tem-se duas barras divididas em cinco partes com os conceitos escritos ao lado, apenas para dar uma melhor noção ao avaliador.

SSCQE – Single Stimulus Continuous Quality Evaluation (Avaliação Contínua da Qualidade com Estímulo Simples)

Nesse método a avaliação é contínua ao longo do tempo, e não possui uma referência, diferente dos métodos anteriores. Assim, durante a apresentação do vídeo, que ocorre somente uma vez, os observadores devem avaliar a qualidade de imagem.

A escala deve ser nos mesmos moldes da escala utilizada pelo método DSCQS, conforme a Figura 17, ou seja, uma escala contínua, dividida em cinco partes, onde o topo correspondente a uma qualidade de imagem excelente e a base, a uma qualidade ruim.

SDSCE – Simultaneous Double Stimulus for Continuous Evaluation (Avaliação Contínua com Duplo Estímulo Simultâneo)

Esse método, bem como o SSCQE, tem como diferencial a avaliação contínua durante o tempo, porém no SDSCE existe a presença de uma referência, assim o observador compara os dois vídeos e avalia a fidelidade entre os mesmos.

Os dois vídeos são apresentados ao mesmo tempo e normalmente em dois monitores distintos, mas caso o tamanho do vídeo seja suficientemente pequeno pode-se apresentar ambos em um mesmo monitor.

O mecanismo de escala deve apresentar a mesma característica do mecanismo utilizado no SSCQE, conforme apresentado na Figura 18, visto que ambos os métodos são de avaliação contínua ao longo do tempo. Assim sendo, e sabendo que o critério a ser avaliado é a fidelidade

da qualidade entre o vídeo padrão e o vídeo defeituoso, o mecanismo deve estar no topo de escala quando não houver diferenças, e na base quando a fidelidade for zero.

Qualidade		Imparidade	
1	Excelente	1	Imperceptível
2	Bom	2	Perceptível, mas não irritante
3	Regular	3	Ligeiramente irritante
4	Pobre	4	Irritante
5	Ruim	5	Muito irritante

Figura 18 – Escala utilizada pelo SDSCE (Adaptado de (ITU-R, 2002))

Para todos os métodos apresentados a avaliação deve durar até 30 minutos, e contar com pelo menos 15 observadores, que devem ser instruídos sobre os detalhes da avaliação (quando avaliar, como os vídeos são apresentados, como funciona a escala e tipos de falha no vídeo) antes da mesma iniciar (ITU-R, 2002).

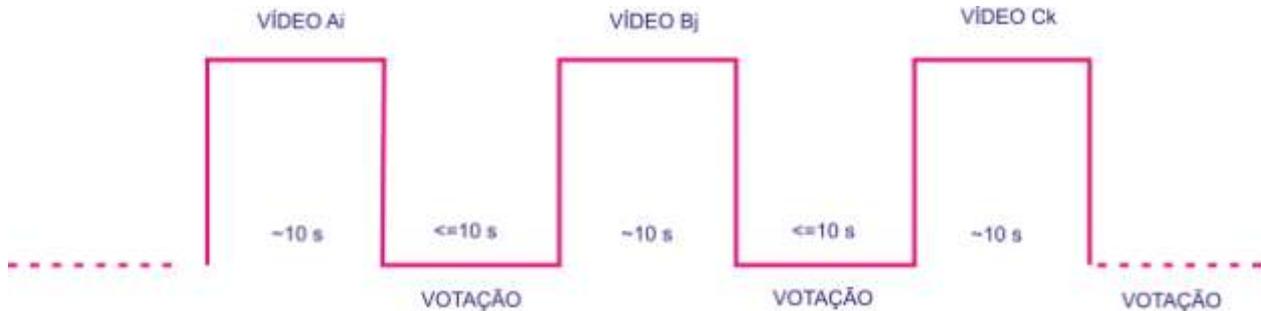
2.3.1.2 Métodos apresentados pela norma ITU-T P.910.11 (ITU-R, 2008)

Além da recomendação ITU-R BT.500.11 (ITU-R, 2002) que se aplica a televisão digital de resolução convencional, existe um conjunto de normas do ITU-R que provêm instruções e métodos disponíveis para o acesso à qualidade da imagem/vídeo em um ambiente de laboratório controlado. A Recomendação ITU-R BT.710 (ITU-R, 1998) trata especificamente de televisão de alta definição e a Recomendação ITU-T P.910 (ITU-T, 2008) faz referência à avaliação subjetiva em aplicações multimídia. O ITU-T Rec. P.910 (ITU-T, 2008) define os seguintes procedimentos de prova:

Absolute Category Rating (ACR – Categoria de Avaliação Absoluta)

O ACR é um método de estímulo único. Nesta categoria as sequências de teste são apresentadas uma de cada vez e são avaliados de forma independente em uma escala de categorias. O método especifica que depois de cada apresentação os avaliadores são solicitados a avaliar a qualidade da sequência mostrada. O padrão de tempo para a apresentação do estímulo é ilustrado pela Figura 19. Se um período de votação constante é utilizada (por exemplo, vários espectadores executar simultaneamente a partir de um mesmo vídeo), então o tempo de votação

deve ser inferior ou igual a 10 s. O tempo de apresentação poderá ser reduzido ou aumentado de acordo com o conteúdo do material de ensaio.



onde:

- Ai Sequência A sob condição de teste i
- Bj Sequência B sob condição de teste j
- Ck Sequência C sob condição de teste k

Figura 19 – Apresentação dos estímulos no método ACR

Fonte: adaptado de (ITU-T, 2008)

O fato de que a referência não é mostrada com cada clipe nos testes, faz do ACR um método muito eficiente em comparação com DSIS ou DSCQS, que têm quase 2 ou 4 vezes mais tempo de duração, respectivamente (WIKLER, 2005).

A seguinte escala de cinco níveis de classificação de qualidade global deve ser utilizada conforme a Figura 20:

5	Excelente
4	Bom
3	Razoável
2	Pobre
1	Ruim

Figura 20 – Escala de categorias utilizada pelo método ACR

Fonte: Adaptado de (ITU-T, 2008)

A utilizada é nos mesmos moldes da escala utilizada pelo método DSCQS, ou seja, uma escala contínua, dividida em cinco partes, onde o topo correspondente a uma qualidade de imagem excelente e a base, a uma qualidade ruim.

Absolute Category Rating with Hidden Reference (ACR-HR – Categoria Absoluta de Avaliação com Referência Oculta).

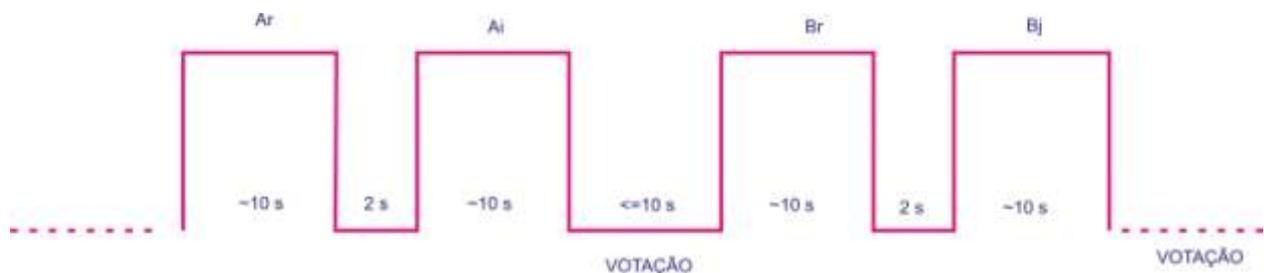
O ACR-HR é uma categoria de julgamento onde as sequências de teste são apresentadas uma de cada vez e são avaliadas de forma independente em uma escala de categoria. O procedimento de ensaio deve incluir uma versão de referência de cada sequência de teste mostrada como qualquer outro teste de estímulo. Isto é denominado com uma condição de referência escondida. Durante a análise dos dados um *score* de qualidade diferencial (DMOS) é calculado entre cada sequência de teste e seus correspondentes vídeos de referência (escondido). Este procedimento é conhecido como "referência escondida". O método especifica que depois de cada apresentação os avaliadores são solicitados a avaliar a qualidade da sequência mostrada.

O padrão de tempo para a apresentação do estímulo é o mesmo apresentado para o ACR (Figura 19). Assim como, a escala de categorias é a mesma utilizada pelo ACR (Figura 20).

Degradation Category Rating (DCR – Categoria de Avaliação de Degradação).

O DCR implica que as sequências de teste são apresentadas em pares: o primeiro estímulo apresentado em cada par é sempre a referência de origem, enquanto que o segundo estímulo é a mesma fonte, apresentados através de um dos sistemas em teste (semelhante ao DSIS). A utilização de formatos de imagem/vídeo reduzidos (por exemplo, CIF, QCIF, SIF), permite mostrar a referência e a sequência de testes simultaneamente no mesmo monitor.

O padrão de tempo para a apresentação do estímulo é ilustrado pela Figura 21.



onde:

Ai Sequência A sob condição de teste i

Ar, Br Sequências A e B, respectivamente, no formato de fonte de referência

Bj Sequência B sob condição de teste j

Figura 21 – Apresentação dos estímulos no método DCR

Fonte: adaptado de (ITU-T, 2008)

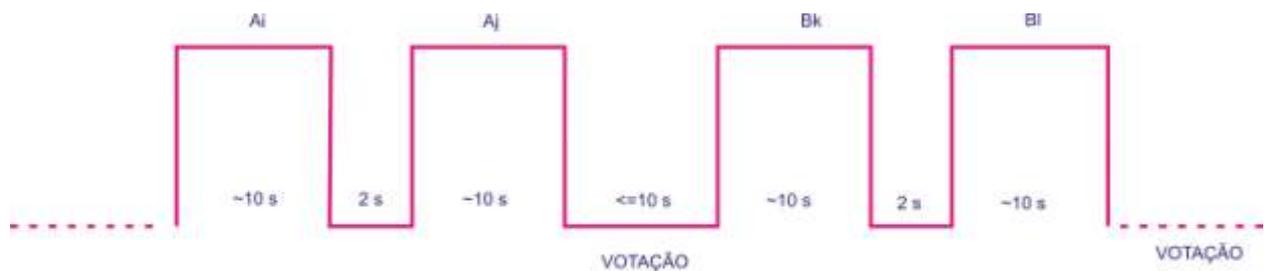
A escala de cinco níveis de avaliação de degradações é semelhante à utilizada pelo método SDSCE (Figura 18) no que diz respeito, nessa escala, às medidas de imparidade.

Pair Comparison Method (PC – Método de Comparação de Pares).

O método PC implica que as sequências de teste são apresentadas em pares, consistindo na mesma sequência que está sendo apresentado por intermédio de um sistema em teste e em seguida, através de outro sistema.

Os sistemas em testes (A, B, C, etc.) são geralmente combinadas em todos os possíveis $n(n-1)$ combinações AB, BA, CA, etc., Assim, todos os pares de sequências devem ser exibidos em ambas as ordens possíveis (por exemplo, AB, BA). Depois que cada par é avaliado no qual um elemento do par é o escolhido no contexto do cenário de teste.

O padrão de tempo para a apresentação do estímulo é ilustrado pela Figura 22.



onde:

Ai e Aj Sequência A sob condições de teste i e j respectivamente
 Bk e Bl Sequência B sob condição de teste k e l respectivamente

Figura 22 – Apresentação dos estímulos no método PC

Fonte: (ITU-T, 2008)

Para o método PC, o número de repetições, não precisa ser geralmente considerado, porque o próprio método implica a repetida apresentação das mesmas condições, embora em diferentes pares.

2.3.1.3 Comparação entre as Normas

Para todos os métodos apresentados na referência P.910 (ITU-T, 2008) a quantidade mínima de observadores é a mesma (15 avaliadores é o recomendado) da norma BT.500-11

(ITU-R, 2002). Também são calculados os valores do MOS que representam a qualidade subjetiva de um determinado vídeo, com exceção do método DSCQS que utiliza o DMOS (MOS Diferencial).

Os procedimentos de testes mencionados acima geralmente têm diferentes aplicações. Os métodos de estímulo simples (DSCQS, DSIS, ACR, DCR, PC) compartilham os mesmos inconvenientes, entretanto: mudanças na complexidade da cena, multiplexação estatística ou transmissões de erros possibilitam variações substanciais na qualidade que não são uniformemente distribuídos ao longo do tempo (WINKLER, 2005).

Segundo WINKLER(2005 *apud* ALDRIDGE,1995) os métodos de estímulo simples não são adequados para avaliação de sequências longas por causa do efeito recente, um desequilíbrio nas avaliações em relação aos vídeos avaliados no final de 10–20 segundos devido a limitações de memória de trabalho humana. Além disso, foi observado que a apresentação de uma referência ou a repetição das sequencias no método DSCQS põe os temas em uma situação de remoção devido aos avaliadores familiarizarem-se com o material sob investigação WINKLER (2005 *apud* LODGE, 1996). A avaliações utilizando-se o SSCQE são mais difíceis de lidar na análise por causa das diferenças potenciais em tempos de reação de espectador e a inerente auto-correlação das séries temporais de dados (WINKLER, 2005; ITU-R, 2002; ITU-T, 2008).

A seguir descrevem-se as métricas objetivas PSNR, MSE, SSIM e MSSIM, pois as mesmas foram utilizadas para realizar a relação com o MOS e validação da avaliação subjetiva.

2.3.2 Métricas Objetivas de Qualidade de Imagem/Vídeo

Os métodos de avaliação objetiva mais simples e mais difundidos são definidos estatisticamente como MSE e PSNR. Estes métodos são baseados em *pixel*, isto é, o vídeo distorcido e a vídeo de referência são comparados *pixel a pixel*. O MSE é computado de acordo com (WINKLER, 2005).

O SSIM foi concebido para ter uma maior aproximação com o SVH, pois leva em consideração a luminância, estrutura e o contraste. O MSSIM é uma média dos valores do SSIM calculados a partir do vídeo avaliado (WANG et al, 2004).

2.3.2.1 MSE

O MSE (*Mean Squared Error*) ou erro quadrático médio é média das diferenças ao quadrado entre os valores de níveis de cinza dos pixels de duas imagens ou sequências I_1 e I_2 segundo a Equação 1:

$$MSE = \frac{1}{TXY} \sum_t \sum_x \sum_y [I_1(t, x, y) - I_2(t, x, y)]^2 \quad (1)$$

para imagens de tamanho $X \times Y$ e T frames na sequência (WINKLER, 2005).

2.3.2.2 PSNR

O PSNR (*Peak Signal to Noise Ratio – Razão Sinal-Ruído de Pico*) é medido em uma escala logarítmica e depende do erro quadrático médio (MSE) entre duas imagens, uma original e uma processada, relativo a $(2^n - 1)^2$ (o quadrado do maior valor possível de sinal na imagem, onde n é o número de *bits* de cada imagem) (WINKLER, 2005), segundo a Equação 2:

$$PSNR_{dB} = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \quad (2)$$

Tanto o MSE quanto o PSNR são definidos somente para informações de luminância. Logo, quando a crominância entra no circuito não se pode garantir a fidelidade do cálculo. Tecnicamente, o MSE mede diferenças entre imagens, enquanto o PSNR mede a fidelidade entre imagens (WINKLER, 2005).

O PSNR pode ser calculado rapidamente e é por isso uma métrica de qualidade bastante popular, largamente utilizada para comparar a “qualidade” de imagens de vídeo comprimidas e descomprimidas (RICHARDSON, 2003). A Figura 23 apresenta um comparativo de 3 imagens: A primeira (a) é a original, as (b) e (c) são versões degradadas (borradas) da imagem original. A imagem (b) tem um PSNR medido de 33,21 dB enquanto a imagem (c) tem um PSNR de 20,93 dB. A imagem c, do ponto de vista da avaliação subjetiva é de melhor qualidade quando comparada à imagem b, porém o valor do PSNR não confirma o que é visto de forma subjetiva.

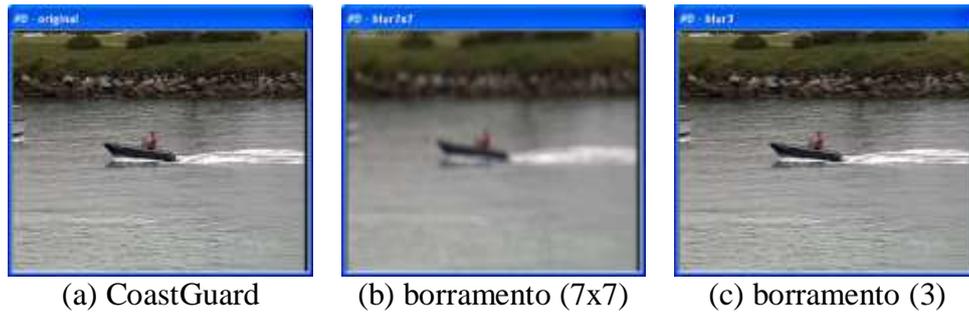


Figura 23 – Comparação entre um vídeo original e dois vídeos degradados

A medida do PSNR possui várias limitações (HUYNH-THU, 2008; SHI, 2009). Ele necessita da imagem/vídeo original no mesmo instante de tempo para comparação, porém, isso pode não estar disponível em todos os momentos, dificultando a verificação de fidelidade em relação ao original. O PSNR não possui uma correlação forte com as medidas subjetivas de qualidade de vídeo, em relação àquelas definidas na norma ITU-R (ITU-R, 2002). Para uma imagem dada ou uma sequência de imagens, um PSNR alto geralmente indica uma alta fidelidade e um PSNR baixo indica baixa qualidade. Porém, um valor de PSNR particular não necessariamente se compara com um valor “absoluto” da qualidade subjetiva (XIONG e ZHU, 2008).

Por causa das limitações de métricas como o PSNR, muitos trabalhos foram realizados nos últimos anos para tentar desenvolver um teste objetivo mais sofisticado que se aproxime dos resultados subjetivos (WATSON, MCGOWAN e MULLIGA, 1999; WANG et al, 2004). Muitas métricas foram propostas, mas nenhuma delas emergiu como uma alternativa clara aos testes subjetivos (VQEG, 2003).

2.3.2.3 SSIM

O índice SSIM é uma métrica *full-reference* (referência completa). O SSIM é aplicado sobre a luminância, contraste e a estrutura da imagem (LU et al, 2002; WANG e BOVIK, 2002; WANG, SIMONCELLI e BOVIK, 2003; WANG, LU e BOVIK, 2004; SESHADRINATHAN e BOVIK, 2004; HUYNH-THU, 2008) e se destina a melhorar os métodos tradicionais como o PSNR e o MSE que provaram ser incompatíveis com a percepção humana.

O SSIM entre duas imagens x e y de tamanho $N \times N$ é dada pela Equação 3:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

onde:

- μ_x a média de x ;
- μ_y a média de y ;
- σ_x^2 a variância de x ;
- σ_y^2 a variância de y ;
- $2\sigma_{xy}$ a covariância de xy ;
- $C_1 = (k_1L)^2$, $C_2 = (k_2L)^2$ duas variáveis para estabilizar a divisão;
- L a faixa dinâmica dos valores dos *pixels* (normalmente é $2^{\text{bits por pixel}} - 1$);
- $k_1 = 0,01$ e $k_2 = 0,03$. Estes valores são um tanto arbitrários, mas nos experimentos realizados pelos autores, mostrou-se que o desempenho do algoritmo do índice SSIM é bastante sensível a variações destes valores.

A fim de avaliar a qualidade de imagem, esta fórmula é aplicada apenas sobre a luminância. A resultante do índice SSIM é um valor entre 0 e 1, e o valor 1 só é alcançável no caso de dois conjuntos idênticos de dados (imagens). Normalmente, é calculado sobre a janela de dimensão 8x8. A janela pode ser deslocada *pixel* por *pixel* sobre a imagem, mas segundo Wang (2004) para reduzir a complexidade computacional, propõe-se a utilização apenas de um subconjunto das possíveis janelas. A Figura 24 ilustra o diagrama do sistema de medida do índice SSIM.

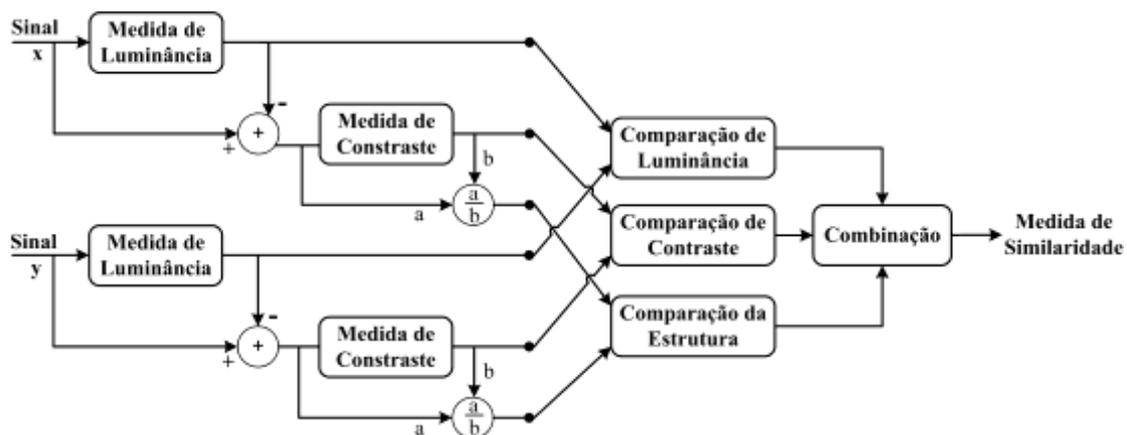


Figura 24 – Diagrama do Sistema de Medida da Similaridade Estrutural.

Fonte: adaptado de (WANG et al. 2004).

2.3.2.4 MSSIM

Normalmente necessita-se de uma única medida de qualidade global de toda a imagem/vídeo. Usa-se, então, uma média do SSIM (MSSIM) para avaliar o índice global de qualidade de imagem representada pela Equação 4:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(X_j, Y_j) \quad (4)$$

onde X e Y são a imagem de referência e as imagens distorcidas, respectivamente, X_j e Y_j são os conteúdos das imagens na janela local j , e M é o número de janelas do local da imagem (WANG et al, 2004).

Ainda não existem padrões para um sistema de medida da qualidade objetiva (“automática”) que seja aplicável em vídeos digitalmente codificados. Em reconhecimento a isso, o Grupo VQEG (*Video Quality Experts Group*) do ITU-T anseia desenvolver padrões para avaliação objetiva da qualidade de vídeo. Em março de 2000, o VQEG reportou a primeira bateria de testes onde dez sistemas competindo foram testados sobre condições idênticas. Infelizmente, nenhuma das dez propostas foi considerada aplicável para padronização e o VQEG completou uma segunda bateria de avaliações em 2003. O VQEG continua fazendo testes e apresentou mais recentemente (2008) mais um conjunto de testes onde várias instituições de pesquisa participaram e ainda assim, não se chegou a um consenso sobre um padrão definitivo de qualidade objetiva de vídeo.

2.4 TRABALHOS RELACIONADOS

Vários trabalhos têm sido publicados sobre avaliação da qualidade de vídeo digitais, discutindo os mais diversos problemas envolvidos (criação, manipulação, transmissão e recuperação de vídeo digital) e a grande maioria utiliza a avaliação subjetiva para validar as suas propostas de métricas objetivas. Nesta seção, discorre-se sobre os trabalhos relacionados que serviram de base de recursos teórico-práticos para a implementação do sistema aqui proposto. A seguir descrevem-se trabalhos relacionados com métricas objetivas, com mais conotação no PSNR.

Em (BASSO et al, 1996), o PSNR é utilizado para medições de qualidade e comparado ao MPQM. O estudo mostra que a métrica MPQM (*Moving Pictures Quality Metric* - Métrica de

Qualidade de Imagens em Movimento), por exemplo, apresenta maior correlação com as métricas subjetivas.

Watson e colaboradores (WATSON et al, 1999) apresentaram uma métrica de qualidade de vídeo digital chamada de DVQ (*Digital Video Quality* - Qualidade de Vídeo Digital). A métrica proposta foi aplicada para sequências de vídeo digital corrompido por vários artefatos de compressão típicos, e compararam os resultados das avaliações de qualidade feitas por observadores humanos foram comparados.

Zhenghua Yu e H.R Wu em seu trabalho (YU e WU, 2000) objetivam proporcionar uma melhoria do conhecimento da modelagem do SVH, como as características do SVH foram incorporadas na medida de qualidade, e as questões de implementação das métricas, bem como as direções das investigações futuras.

YuanFei (YUAN e HUANG, 2007) apresentam um estudo sobre qualidade de vídeo baseado em uma versão modificada do PSNR que leva em consideração características do SVH, conseguindo com isso, melhor correlação com as medidas subjetivas. O PSNR também é utilizado em (TULU, 2008) para avaliação de qualidade em vídeos usados em telemedicina. Os resultados indicam que o *jitter* (variação do atraso) e *delay* (atraso) têm efeitos significativos sobre os valores de qualidade objetiva.

Sukumar e colaboradores (SUKUMAR et al, 2007) concluem que a medição da qualidade visual é de fundamental importância para projetar algoritmos que, automaticamente, possam avaliar a qualidade das imagens ou vídeos de uma maneira perceptivelmente coerente. Neste trabalho os autores fizeram um estudo extensivo subjetivo utilizando diferentes limiares de modulação em uma tentativa de capturar características fisiológicas e psicovisuais do sistema visual humano (SVH).

De Angelis e colaboradores apresentam em seu trabalho (DE ANGELIS et al, 2007) o estado da arte no campo Avaliação da Qualidade de imagem (IQA), fornecendo uma classificação de alguns dos mais importantes métodos IQA objetivos e subjetivos. Mais ainda, alguns aspectos do campo são analisados do ponto de vista das medidas, também através da comparação com a área de medição de qualidade de software. Em particular, uma abordagem estatística para a avaliação da incerteza para métodos objetivos IQA é apresentada e um exemplo é fornecido. O tema da modelagem para os métodos de medição subjetiva IQA também é analisada. Finalmente, um estudo de caso de imagens corrompidas por ruído impulsivo é discutido

Em (HUYNH-THU e GHANBARI, 2008) os autores discorrem sobre a utilização do PSNR como a métrica objetiva mais utilizada na avaliação de imagens/vídeos digitais, bem como apresenta as desvantagens de utilizá-la com vídeos diferentes e a perda da correlação com medidas subjetivas. O artigo apresenta resultados de imprecisão do PSNR como preditor de qualidade subjetiva.

Em EDEN (2008) é introduzida uma abordagem para medir a qualidade de imagem com relação à percepção humana. A abordagem é baseada no PSNR estendido com um indicador de atividade espacial. Seus benefícios são provados por testes subjetivos.

Em todos os trabalhos citados acima foram utilizados o PSNR e variações do mesmo para medir a qualidade de vídeo digital e as correspondentes medidas subjetivas para validar as propostas feitas em cada trabalho. Observa-se que a medida subjetiva acompanha todas as propostas como ferramenta de validação dos mesmos, uma vez que a finalidade das experiências é conseguir métricas objetivas que possam se aproximar dos valores de qualidade dados por avaliadores humanos. Outros autores como (WANG et al, 2004) propõem uma nova métrica para medir qualidade de vídeo em detrimento da utilização do PSNR, uma vez que o mesmo apresenta uma correlação pobre com medidas subjetivas. A seguir apresentam-se trabalhos relacionados a novas métricas propostas por vários pesquisadores e os métodos de avaliação subjetivos utilizados para validação das mesmas são destacados.

Zhou Wang e Alan C. Bovik (WANG e BOVIK, 2002) fornecem algumas indicações do porquê a avaliação de qualidade de imagem é tão difícil, assinalando as debilidades do *framework* baseado em sensibilidade de erro, que foi usado pela maioria de avaliações de qualidade de imagem abordada na literatura. Além disso, foi proposta uma nova filosofia em métricas de qualidade de imagem.

Stefan Winkler e Ruth Campos (WINKLER e CAMPOS, 2003) utilizaram o SSCQE e o DSIS para realizar vários testes subjetivos. Deste modo comparam estes métodos de prova e apresentaram uma análise dos resultados experimentais devido à apresentação por *codec*. Finalmente, os autores usaram os dados subjetivos para confirmar a precisão de predição em tempo real métrica de qualidade sem-referência.

Wang, Bovik e outros autores em seu trabalho (WANG et al, 2004) partem do pressuposto de que a percepção visual humana é altamente adaptada para extrair informações estruturais de uma cena, e apresentam um quadro alternativo complementar para avaliação da qualidade com base na degradação das informações estruturais. Como um exemplo concreto desse conceito, os autores desenvolveram um índice da similaridade estrutural e demonstraram a

sua garantia através de um conjunto de exemplos intuitivos, bem como a comparação com ambas as avaliações subjetivas e o estado da arte dos métodos objetivos em um banco de imagens comprimidas nos formatos JPEG e JPEG2000.

Hands em (HANDS et al, 2005) analisa se as avaliações de peritos, normalmente evitada por laboratórios de testes subjetivos em favor de não-especialistas, fornecem mais informações e se utiliza de votos SSCQE com o objetivo de desenvolver modelos de qualidade de vídeo. Os resultados mostram que, para treinamento e teste objetivo de modelos de qualidade de vídeo, é preferível usar os dados obtidos a partir de especialistas do que os dados obtidos a partir de indivíduos não-especialistas.

Pinson e Wolf em (PINSON e WOLF, 2003) apresentam em seu artigo uma comparação entre métricas subjetivas destacando-se nesse estudo o DSCQS e SSCQE. Neste estudo os autores fazem comparações com seis experimentos subjetivos e comparam os resultados utilizando as métricas supracitadas bem como fornecem evidências de que a memória humana para efeitos de estimativa de qualidade variando no tempo parece ser limitada a cerca de 20 segundos.

Jumisko e colegas em (JUMISKO, ILVONEN e VAANANEN-VAINIO-MATILLA, 2005) apresentaram um estudo que revela que os critérios de avaliação subjetivos mostraram que os avaliadores prestam atenção ao conteúdo e deteriorações de qualidade especialmente em regiões de interesse, de acordo com o conteúdo e o tipo de vídeo exibido.

SHEIKH, SABIR e BOVIK (2006) apresentam resultados de um estudo extensivo de avaliação subjetiva de qualidade em que um total de 779 imagens distorcidas foram avaliados por cerca de duas dezenas de seres humanos. A qualidade dos dados obtidos a partir de cerca de 25.000 decisões individuais da qualidade subjetiva é usada para avaliar o desempenho de métricas (FR – *Full Reference*). Ao melhorar nosso entendimento, além de estudos de qualidade de vídeo realizado pela *Video Quality Experts Group* (VQEG, 2000), o estudo apresentado neste documento é o maior estudo de qualidade subjetiva de imagem na literatura em termos de número de imagens, tipos de distorção, e número de julgamentos humanos por imagem.

Monteiro e Nunes em (MONTEIRO e NUNES, 2007) apresentam uma ferramenta de quantificação para a avaliação de sistemas de comunicação de vídeo escalável e não escalável. O mecanismo proposto estima a Qualidade Subjetiva de Experiência (QoE) de um espectador humano de acordo com a resolução temporal, a resolução espacial do RMSE (*Root Mean Squared Error* - Erro Quadrático Médio Principal) entre a imagem original e a codificada. A

ferramenta de quantificação proposta foi obtida por testes subjetivos usando um painel de avaliadores e uma nova metodologia que mostrou bons fatores de correlação entre medidas dos dados e funções de estimação.

Vladimir Petrović em sua pesquisa (PETROVIC, 2007) enfoca a metodologia de avaliação da percepção da imagem de fusão através de testes comparativos e validação de métricas de avaliação objetiva de fusão. São descritos testes subjetivos para validação dos testes objetivos.

Zampoulo e colaboradores em (ZAMPOULO, GOMES e SEARA, 2008) propõem uma metodologia de avaliação de desempenho como também compara algumas métricas de qualidade visual para a caracterização de limiar de percepção em imagens degradadas por ruído aditivo gaussiano. Neste trabalho é realizada uma avaliação subjetiva com o intuito de estimar um valor de limiar de detecção para cada uma das métricas avaliadas bem como a qualidade obtida da correspondente estimativa. Neste trabalho, é citada a utilização de uma ferramenta de armazenamento de dados dos resultados das avaliações subjetivas.

ROUSE e HEMAMI (2008) investigam como se obter uma avaliação de informações objetivas visuais, predizendo o limite de reconhecimento de representações de imagem diferentes. Para isso utilizam-se do MSSIM para reconhecimento das estruturas básicas nas imagens.

Shengnan Ye e outros (YE, SU e XIAO, 2008) apresentaram uma nova métrica para avaliação de qualidade de vídeo que seleciona regiões de amostragem local de acordo com regiões sem borda e usa o SSIM como medida de distorção local.

Kwangjin Choi e colegas em (CHOI et al, 2008) apresentam um estudo e comparação de métricas de qualidade de vídeo streaming. Destacando-se as medidas tanto objetivas quanto subjetivas de medição de qualidade.

De forma a entender e possibilitar a utilização das novas tecnologias de *displays* de vídeo destaca-se a seguir trabalhos relacionados a estudos comparativos de visores e LCD e CRT.

Margaret H. Pinson e Stephen Wolf em (PINSON e WOLF, 2004) apresentam uma comparação dos resultados de testes de qualidade de vídeo subjetiva profissional a partir de comparação entre um monitor de tubo de raios catódicos (CRT monitor de televisão) com a de um display de cristal líquido (monitor LCD).

No relatório de G. Alan Brown (BROWN, 2007) é apresentado um estudo sobre as condições de exibição de vídeos, testes de monitores LCD e suas características para manter uma qualidade aceitável pelos consumidores de vídeo.

Hisatake e colegas (HISATAKE et al, 2007) analisam a correlação entre três tipos de visualizações de desfoque de movimento (borramento de cor, luminância do borramento e nitidez dos objetos da imagem) e propriedades físicas. Foi realizada também avaliação subjetiva para especificar o nível ótimo ou permitido de desfoque de movimento com várias imagens em um monitor CRT e quatro tipos de LCDs, com características diferentes.

SOUK e LEE (2007) fizeram um estudo da qualidade da imagem e do aprimoramento de tecnologias aplicadas ao SVH. Um mecanismo de busca foi aplicado à melhoria da qualidade de imagem em movimento. Questões de luminância relacionadas com display de cristal líquido (LCD) foram revistas, tais como fadiga ocular de alta luminância e relação de contraste perceptível máxima do display.

Em seu estudo sobre a utilização de LCDs na avaliação de vídeo KUO, LIN e HWANG (2008) desenvolveram um *framework* para obter a qualidade perceptiva em painéis de LCD na exibição de TV (LCD-TV) do ponto de vista do usuário. Os resultados indicaram que a luminância, contraste e o tempo de resposta são fatores importantes em selecionar um LCD-TV. Nessa linha de análise de LCDs e de imagens em HDTV, Hans Hoffmann em (Hoffmann et al, 2008) propõem um novo método para medir a qualidade de imagens chamado TSCES (*Triple Stimulus Continuous Evaluation Scale* – Escala de Avaliação Contínua de Estímulo Triplo), para permitir uma comparação direta para diferentes formatos de HDTV. O método utiliza três exibições em um arranjo vertical particular. O novo método forneceu resultados estáveis e consistentes.

HUYNH-THU e GHANBARI (2006) investigam o impacto perceptivo dos fotogramas repetidos e suas perdas na qualidade percebida. Os testes subjetivos foram efetivados utilizando-se um display de LCD.

T.C.M de Koning et al em (KONING et al, 2007) exploram a relação entre medidas subjetivas e objetivas de qualidade de vídeo.

O VQEG no relatório, “*Final Report from the Video Quality Experts Group on the Validation of Objective Models of Multimedia Quality Assessment*” (VQEG, 2008) apresenta os resultados do (VQEG) sobre testes de validação de modelos objetivos de qualidade de vídeo para celular/PDAs e Serviços de Comunicações de Internet de banda larga. Este documento fornece entrada para os organismos de normalização pertinentes responsáveis pela produção de recomendações internacionais. O teste multimídia contém duas avaliações paralelas de material de vídeo de teste. Uma avaliação é realizada por grupos de observadores humanos (ou seja, testes subjetivos). A outra é por modelos objetivos computacionais de qualidade de vídeo (ou seja, os

modelos proponentes). Os modelos objetivos são destinados a prever os julgamentos subjetivos. O plano de testes subjetivos oficial está disponível em (PLANO DE TESTES SUBJETIVOS, 2009).

Brunnström em seu artigo *VQEG Validation and ITU Standardization of Objective Perceptual Video Quality Metrics*, (BRUNNSTRÖM et al, 2009), revê o trabalho realizado pelo VQEG, em especial aos ensaios de validação dos testes propostos pelo VQEG. Para a medição da qualidade objetiva o VQEG afirma que existem três aspectos do desempenho: a precisão da previsão (ou seja, a previsão exata do MOS subjetivo de cada sequência), os requisitos computacionais e tempo de execução. Esse trabalho justifica a implementação de um ambiente e suas ferramentas uma vez que o VQEG, como grupo referência em qualidade de imagens e vídeos, sugere que, para medir a qualidade objetiva, um dos aspectos que devem ser levados em consideração é a avaliação subjetiva, bem como outros autores citados nesta seção utilizam a avaliação subjetiva como parte dos experimentos de teste e a validação de novas métricas objetivas.

2.5 RESUMO E CONCLUSÃO DO CAPÍTULO

Neste capítulo apresentaram-se os principais conceitos norteadores para o entendimento do trabalho desenvolvido, bem como caracterizam-se a necessidade de estudo e de desenvolvimento de ferramentas para auxiliar no estudo dos vídeos digitais e suas peculiaridades.

Destaca-se aqui o SVH, pois o que se busca é conseguir entender e produzir vídeos que sejam agradáveis a esse sistema tão complexo. Detalhes sobre representação das cores e dos sistemas de representação de cores, bem como os conceitos de vídeos digitais. As métricas objetivas utilizadas para compor os testes realizados foram explanadas, bem como os detalhes referentes aos principais métodos de avaliação subjetiva constantes das normas, além dos artefatos de vídeo digital.

Os trabalhos relacionados apresentaram vários trabalhos que destacam a avaliação subjetiva como parte fundamental, mesmo quando o estudo é focado em métricas de qualidade objetiva de vídeo digital. Além disso, observou-se que a utilização de ferramentas é uma necessidade para minimizar o tempo de resposta das avaliações subjetivas.

A seguir descreve-se a metodologia utilizada para a modelagem do ambiente de avaliação de vídeo digital proposto neste trabalho.

3 ESTUDO DE CASO PRELIMINAR DE AVALIAÇÃO SUBJETIVA

Nos capítulos anteriores, apresentaram-se os conceitos referentes ao SVH, sobre os sistemas de cores e vídeos digitais e, além desses, as métricas que foram utilizadas para o desenvolvimento do trabalho de modo a informar ao leitor sobre as características de cada métrica e os artefatos de vídeo digital.

Neste capítulo, apresenta-se a metodologia utilizada para o desenvolvimento do ambiente de avaliação subjetiva de vídeos digitais. O propósito é descrever como foi feita a extração dos requisitos por meio da realização de um experimento de avaliação subjetiva que possibilitou a verificação de todos os aspectos necessários para a implementação do ambiente e do sistema proposto.

3.1 FUNDAMENTOS METODOLÓGICOS DA AVALIAÇÃO

O método aplicado a este estudo de caso é baseado no escalonamento de opiniões, pois nestes métodos diferentes valores de uma degradação são aplicados a uma imagem/vídeo em ordem aleatória. O próximo passo é a classificação realizada pelos observadores das imagens/vídeos a partir de uma escala de categorias (SILVA, 1998).

As normas de avaliação subjetiva tratam dos aspectos relativos às condições de observação, seleção e preparação do material, número e tipos dos avaliadores enfatizam, principalmente, a seleção da metodologia adequada ao problema, com relação à seleção da forma de apresentação e ao tipo de escalas de notas. A quantidade de dados que precisa ser coletada depende de fatores intercorrelacionados tais como: intervalo de confiança, desvio padrão e amplitude relativa do efeito que se deseja medir (ITU-R HANDBOOK, 1996).

A variação perceptível na qualidade da imagem por determinado tipo de degradação depende fortemente da natureza da imagem, evidenciando a necessidade de escolha da imagem em função do tipo de degradação que se deseja avaliar.

Na análise dos resultados de um experimento, o objetivo principal é estimar da melhor forma possível a distribuições de opiniões para uma determinada condição da imagem/vídeo. A exatidão não pode ser obtida dada a incerteza inerente ao processo de amostragem.

As características das classes de avaliadores desempenham papel importante na análise dos resultados, pois um modelo aplicável a um grupo de especialistas pode não ser aplicável a

um grupo de não especialistas. Mesmo dentro de um grupo pode haver diferenças dependendo do nível social e instrução.

3.1.1 Condições de observação

As condições de observação para métodos subjetivos de avaliação da qualidade de imagem/vídeo digitais variam de acordo com as normas e, especificamente, no que diz respeito à televisão convencional e para as de alta resolução. Dessa forma, os parâmetros devem ser mantidos o mais próximo possível dos descritos nas padronizações para permitir comparações com resultados de outros laboratórios, como por exemplo prático, o VQEG (2008) e seus membros em vários países realizam testes baseados nas mesmas condições e padronizações para conseguir um maior amplitude das comparações dos testes realizados.

Os parâmetros representados são um resumo dos valores estabelecidos pela ITU-R BT 500.11 (ITU-R, 2002), ITU-R BT. 710 (ITU-R, 1998), norma NBR 5413 (NBR, 1992) e relatório VQEG (VQEG, 2008):

(continua)

Condições gerais	
Razão entre a distância de visualização e a altura da tela:	
TV padrão	4H – 6H
HDTV	3H
Ajustes do monitor	
Luminância de pico	70 cd/m ²
Razão entre a luminância da tela do tubo inativo e a luminância de pico	≤ 0,02
Razão entre a luminância da tela, quanto mostrando somente o nível preto em uma sala completamente escura, e a correspondente ao nível de pico de branco	0,01
Condições da sala	
Razão entre a luminância do fundo atrás do monitor e a luminância de pico da imagem	0,15
Outra iluminação da sala	Baixa

(conclusão)

Condições especiais	
Número de avaliadores por monitor	3
Monitor	Alta qualidade com 54 cm para TV convencional e 140 cm para HDTV
Brilho e contraste do display	Acertado via sinal PLUGE ¹
Natureza da sala de visualização	Sala com três paredes na cor “gelo” e uma parede na cor cinza (70 a 90% negro) atrás do monitor

onde H = altura da imagem/vídeo

Quadro 1 – Condições gerais de visualização e apresentação dos testes

Fonte: adaptado de (ITU-R, 2002; ITU-R, 1998; NBR, 1992; VQEG, 2008):

3.1.2 Influência da distância e do tamanho da imagem na observação

A sensibilidade do olho humano é influenciada por componentes da imagem, tais como: frequência, cor, movimento, contraste e brilho (GONZALEZ e WOODS, 2000). Além das características espaciais e temporais da imagem, diversos fatores relacionados às condições de visualização também influenciam de forma determinante a qualidade de uma imagem (WU e RAO, 2006).

A distância do observador para uma cena pode influenciar a qualidade do que se está assistindo, pois o aumento da distância do observador à cena diminui o ângulo de visão e com isso pode variar a opinião do observador sobre a cena observada. Segundo WESTERINK (1989) a qualidade subjetiva da imagem na distância ótima é independente da largura da imagem e depende linearmente com o logaritmo da resolução (relacionada à banda da imagem).

3.1.3 Análise de variabilidade entre avaliadores

Como regra geral (ITU-R, 2002), devem ser utilizados, pelo menos, 15 avaliadores. A escolha de avaliadores experientes ou não também determinam variabilidade nos resultados. Pode-se utilizar avaliadores especialistas e não especialistas, porém a utilização de especialistas influenciará diretamente na caracterização das diferenças entre as cenas. Ou seja, avaliadores

¹ <http://www.mediacollege.com/video/test-patterns/pluge.html>

especialistas observarão as diferenças entre as cenas devido a sua experiência em avaliações semelhantes.

Como o objetivo dos testes subjetivos era o de avaliar artefatos e ter critérios para a criação de ferramentas de auxílio na avaliação subjetiva optou-se por avaliadores inexperientes em avaliação de vídeos digitais.

Após serem expostos às degradações, com o tempo, os avaliadores tendem a se acostumar ou se tornarem mais avessos às degradações. Também, em função dessas características, a escolha entre especialistas e não-especialistas em avaliações subjetivas deve ser cuidadosamente considerada.

3.1.4 Material de teste

A escolha de materiais de teste é essencial para produzir melhores resultados para as avaliações subjetivas. A escolha das cenas é fundamental e deve ser considerado o que se deseja medir. Além da escolha baseada no que pretende avaliar, a quantidade das cenas e os níveis de degradação são importantes devido ao tempo máximo de duração do teste que é baseada na norma do teste subjetivo e pela necessidade de redundância.

3.1.5 Forma de apresentação das sequências de vídeo

Em relação à forma de apresentação dos vídeos, dois aspectos são fundamentais: a utilização ou não de vídeos de referência e a repetição do estímulo. A apresentação simultânea de dois vídeos, seja um deles de referência ou não, é a forma mais efetiva de avaliar um sistema, promovendo uma comparação mais eficiente entre os estímulos. Esta técnica implica, entretanto, na disponibilidade de um sistema de gerência que garanta a correta sincronização entre os vídeos e sem o qual a apresentação simultânea tende a deteriorar os resultados da avaliação (SILVA, 1998).

A utilização de vídeo de referência é preferível devido aos efeitos de adaptação do observador. A adaptação consiste na variação de intensidade das notas dos avaliadores em função dos demais vídeos em avaliação. A confiabilidade do método subjetivo é extremamente dependente do número de observadores e da diversidade de sequências utilizadas (ITU-R, 2002).

No caso da duração dos trechos de avaliação, estabeleceu-se a utilização de trecho de vídeo de 10 segundos de duração. Conforme a Recomendação ITU-R(2002) apenas uma nota é

dada para cada estímulo, mas, especificamente no caso de trechos de vídeo com movimento, faz-se necessária alguma forma de ponderar esta impressão de qualidade ao longo do tempo. Foram também utilizados vídeos com duração variando entre 3 minutos e 10 minutos, para avaliar as degradações ao longo do tempo em escala contínua.

3.1.6 Escalas

Em função dos elevados custos das avaliações subjetivas, grande empenho é dedicado ao desenvolvimento de metodologias que permitam a obtenção de resultados de máxima confiabilidade e mínimo esforço. A escala não deve causar dúvida ao avaliador e deve ter sensibilidade suficiente para se ajustar a todos os critérios utilizados pelo avaliador.

Com relação às características as escalas podem ser (SILVA, 1998):

- (i) Escala de categorias. O termo escala de categorias é usado para definir escalas onde a graduação dos níveis é definida por termos da língua portuguesa. São de dois tipos: relacionadas à qualidade da imagem/vídeo e definição do nível de degradação da imagem/vídeo.
- (ii) Escala numérica linear. As escalas numéricas exploram a habilidade do avaliador de emitir julgamentos em uma escala linear. Uma vez que os diversos níveis da escala não estão limitados pelos valores dos adjetivos, a escala pode ser utilizada em diferentes faixas de degradação.
- (iii) Escala racional. Este tipo de escala fornece informações que outras não são capazes, como por exemplo, a distância entre variáveis, além da vantagem de ser limitada, pois é característica da avaliação humana que os avaliadores evitam os extremos da escala, reduzindo a informação que pode ser obtida.
- (iv) Escala gráfica. As escalas gráficas, por si só, não identificam uma classe de métodos de avaliação subjetiva, mas caracterizam a forma de utilização das escalas.
- (v) Escala multidimensional. Nesta escala, deseja-se explorar a similaridade entre imagens com diferentes níveis e tipos de degradação.

3.1.7 Estimação da qualidade subjetiva

A necessidade de medir a qualidade de vídeos é consequência natural do desenvolvimento de técnicas, métodos e equipamentos, no que diz respeito à geração e distribuição de vídeo. As duas categorias principais para a avaliação de vídeo são a avaliação subjetiva, que utiliza observadores humanos, e a avaliação objetiva que se utiliza de medidas físicas da imagem.

Conforme citado na seção 2.4.1 a avaliação subjetiva é considerada a melhor avaliação da qualidade, uma vez que é realizada por observadores humanos. No entanto, é necessário o desenvolvimento de medidas objetivas de qualidade de vídeo. Medidas são necessárias para as organizações de padronização, usuários finais e provedores de serviços. Medidas objetivas imparciais, confiáveis, repetíveis e de baixo custo, tanto de vídeo quanto de desempenho dos sistemas de transmissão, aumentam a competição e avaliação de novos sistemas. A possibilidade de realização de medidas objetivas em tempo real abre possibilidades de monitoração contínua da qualidade de vídeo.

O estudo desenvolvido nesse trabalho, em primeiro plano, deseja estimar a qualidade subjetiva por comparação entre vídeos originais e degradados, automatizando a geração dos resultados por meio da especificação do ambiente de avaliação e suas ferramentas. O esquema utilizado está descrito na Figura 25.



Figura 25 – Proposta de desenvolvimento da ferramenta de avaliação subjetiva

No que tange ao modelo de descrição das degradações, a Recomendação ITU (ITU-R, 2002) define funções sigmóides não simétricas para a descrição da relação entre medidas objetivas e subjetivas feitas sobre uma imagem/vídeo, conforme ilustrado na seção referente a

resultados deste trabalho (Seção 5.2). Estas funções procuram modelar os efeitos de saturação e assimetria da percepção próprios do SVH.

As médias de distorção devem levar em consideração a faixa de degradações da imagem/vídeo e a fidelidade das observações humanas. É comum assumir que a qualidade do vídeo percebida pelos observadores humanos é proporcional àquela estimada por medidas objetivas. Esta relação, entretanto, não é verdadeira. A forma encontrada para descrever este efeito é através de uma função com comportamento linear na região central, mas com características de saturação nas bordas.

Outra característica importante é a assimetria da percepção. Essa assimetria significa que se tomando um vídeo e a ele associando a um ganho e uma perda na qualidade equivalente do ponto de vista objetivo, a perda de qualidade é, em termos subjetivos, mais significativa. Esta qualidade do vídeo resulta em uma habilidade assimétrica por parte dos avaliadores em acompanhar variações de qualidade em testes subjetivos, e faz com que o comportamento na região central se afaste, ainda mais, de uma reta.

3.2 EXPERIMENTO DE AVALIAÇÃO SUBJETIVA PRELIMINAR

O processo de captação dos requisitos se iniciou a partir da realização do trabalho de mestrado apresentado em (ALBINI, 2009). O referido trabalho trata da implementação de vários algoritmos que simulam artefatos em vídeos digitais. Estes algoritmos simularam os artefatos de blocagem, borramento (ver seção 2.5.3), ruído branco gaussiano, ruído sal e pimenta e travamento.

Para visualizar os resultados produzidos pelos algoritmos implementados uma metodologia de avaliação subjetiva foi aplicada para avaliá-los. O experimento de avaliação foi a base para estabelecer os requisitos do ambiente de avaliação. Este experimento envolveu procedimentos normalizados para avaliação subjetiva, a aplicação de questionários e entrevistas com os avaliadores sobre os procedimentos utilizados e suas sugestões de como agilizar a captação dos resultados.

A Figura 26 apresenta o fluxograma das atividades desenvolvidas no procedimento adotado por (SILVA, 2008; ALBINI, 2009). O objetivo dessa ilustração é dar uma visão macro dos procedimentos adotados no experimento preliminar e, nas seções posteriores, comparar com os resultados obtidos após a implementação do ambiente de avaliação.

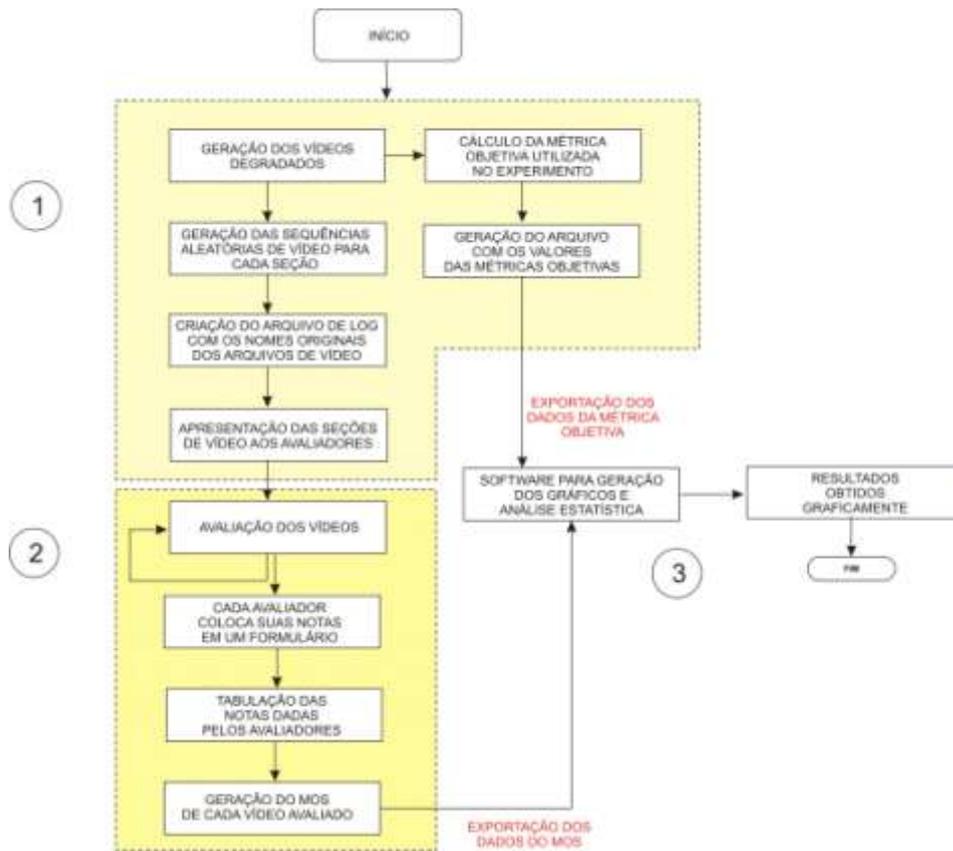


Figura 26 – Fluxograma de atividades desenvolvidas na avaliação subjetiva preliminar

No bloco identificado por 1, estão alguns procedimentos do sistema que foram implementados por (SILVA, 2008; ALBINI, 2009), descritos a seguir:

- Geração dos vídeos degradados: algoritmos utilizados para gerar as degradações nos vídeos de teste. Foram gerados cinco conjuntos de arquivos degradados separados por tipo de vídeo e degradação.
- Cálculo da métrica objetiva utilizada no experimento: é realizado um algoritmo para fazer os cálculos do MSE e do PSNR de cada um dos vídeos gerados.
- Geração do arquivo com os valores da métrica objetiva: geração dos arquivos texto com os nomes dos vídeos e seus respectivos valores de MSE e de PSNR, para posterior recuperação dos mesmos. Nesta etapa utilizaram-se somente essas duas métricas.
- Geração das sequências aleatórias de vídeo para cada sessão: entende-se por “sessão” o processo de apresentação das sequências de vídeo aos avaliadores. Tendo sido gerados os vídeos e realizados os cálculos das métricas objetivas, foi necessário, segundo a norma (ITU-R, 2002), produzir sequências diferentes para cada sessão.

- Criação do arquivo de *log* com os nomes originais dos arquivos de vídeo: foi criado um arquivo com o nome original e o nome que foi dado ao arquivo em determinada sessão, este arquivo será utilizado para posterior recuperação dessa informação para a tabulação dos dados.
- Apresentação das sessões de vídeo aos avaliadores: nesta etapa foi realizada a apresentação das sequências de vídeo usando o método subjetivo SDSCE descrito em (ITU-R, 2002).

No bloco 2, representa-se a seção de avaliação de vídeos, conforme descrito a seguir:

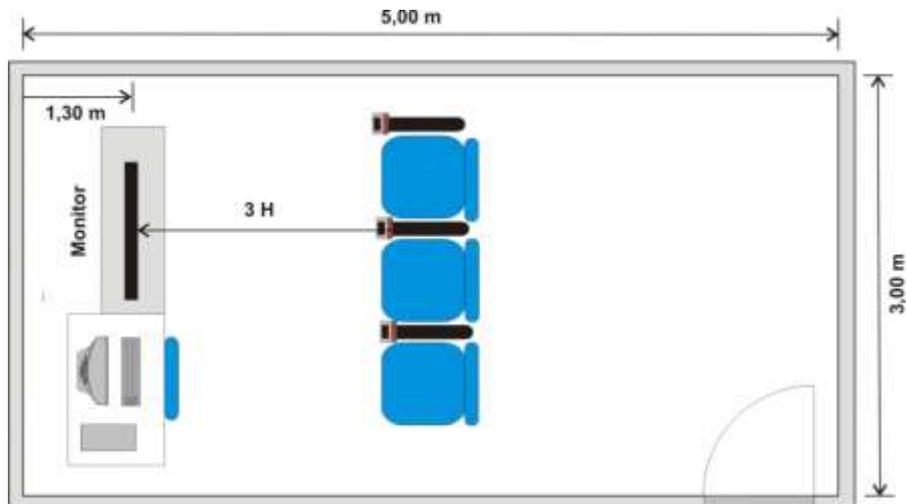
- Avaliação dos vídeos: realização das avaliações dos vídeos pelos avaliadores, com grupos de no máximo de três avaliadores. Os avaliadores foram instruídos sobre a avaliação (treinamento, teste de acuidade) e como preencher o formulário de avaliação distribuído no início da sessão. Esse formulário serviu para captar as notas dadas pelos usuários, bem para obter dados acerca da avaliação para a geração dos requisitos do sistema.
- Tabulação das notas dadas pelos avaliadores: ao final de cada sessão, os formulários foram recolhidos e os dados transferidos para uma planilha para posterior tabulação dos dados.
- Geração do MOS de cada vídeo avaliado: tendo os dados registrados na planilha, foi calculado o MOS dos vídeos de acordo com os dados recuperados das seções e do arquivo de *log*, o qual foi criado a partir da geração das sequências de vídeo. Após a tabulação e cálculo do MOS, os dados foram exportados para um *software* de geração de gráficos (Origin) e análise estatística dos resultados.

3.2.1 Limitações e considerações

As etapas descritas anteriormente e a abstração do fluxograma apresentado na Figura 26 apresentaram a forma como foi conduzido o experimento e feita a captação dos requisitos necessários à construção das ferramentas de auxílio à avaliação subjetiva. O processo inteiro levou muito tempo, em torno de 4 horas para a tabulação dos dados e inserção dos mesmos no gerador de gráficos, e muito trabalho manual foi dispendido para chegar aos resultados, mesmo tendo uma boa parte do processo automatizado.

Este experimento teve como objetivo captar os requisitos necessários para automatização do processo de geração dos resultados, bem como de fazer a sintonia fina no que diz respeito aos parâmetros necessários à especificação do ambiente de avaliação subjetiva e suas ferramentas. Os experimentos realizados permitiram a análise de requisitos que levaram à construção de um sistema de automatização dos resultados estatísticos das avaliações subjetivas.

Nesta avaliação subjetiva preliminar foi utilizado o método SDSCE (Seção 2.3.1.1), definido na Recomendação ITU (ITU-R, 2002). Todos os aspectos referentes às condições de visualização foram cuidadosamente analisados e controlados. A Figura 27 ilustra o ambiente de avaliação utilizado.



onde: H = altura da imagem/vídeo

Figura 27 – Ambiente de avaliação utilizado

3.2.2 Material de teste

Esta etapa da avaliação subjetiva compreendeu a reprodução e organização do material de forma criteriosa e sistemática.

As cenas utilizadas no experimento fazem parte da base IVC (AUTRUSSEAU, 2005): *Mobile*, *Stefan*, *Akiyo (News)* e *CoastGuard*, conforme apresentado na Figura 28.

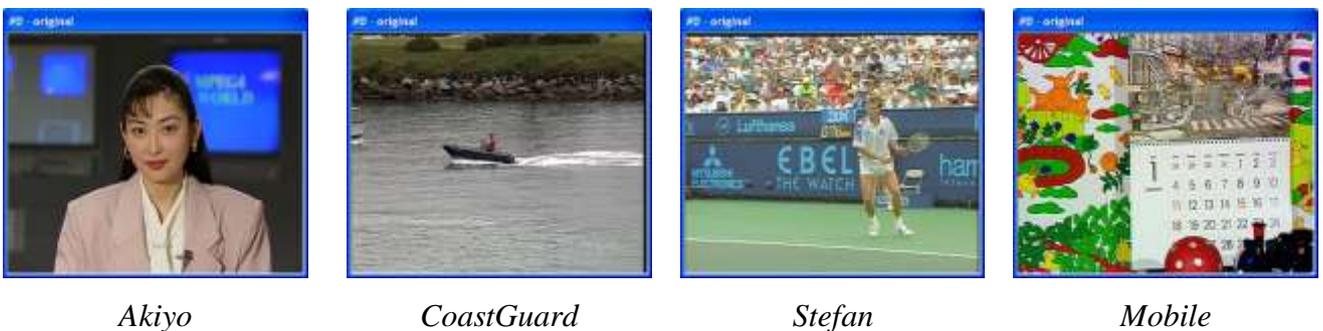


Figura 28- Vídeos originais utilizados nos experimentos.

As fontes de vídeo utilizadas foram sem compressão (formato YUV), com resolução CIF (352x288 *pixels*), amostrados na taxa 4:2:0. Essas fontes foram selecionadas por possuírem um formato onde os *pixels* são armazenados diretamente e não sofreram nenhuma interferência externa. Foram utilizadas com muita frequência em outros experimentos citados na literatura referente à avaliação da qualidade de vídeo digital e por possuírem características diferentes uns dos outros, como forma de analisar situações de visualização de cenas diferentes.

Os traços de vídeo foram eleitos visando abranger a maior diversidade de características, tais como: movimentação, textura e coloração. Os vídeos originais pertencem ao rol dos vídeos comumente utilizados em testes de ferramentas que envolvem este tipo de informação e características (TRAÇOS DE VÍDEO, 2008). Todos os traços possuem duração de 10 segundos, a uma taxa de 30 quadros por segundo. A seguir, são descritos sucintamente os procedimentos adotados.

3.2.3 Processamento

O Quadro 2 apresentado a seguir é uma descrição dos vídeos que foram analisados no experimento:

Descrição	Processamento	Observação
Vídeo original não degradado	nenhum	Os vídeos originais, ou seja, que não sofreram qualquer tipo de processamento foram tomadas como referência nas avaliações e foram incluídos entre as sequências avaliadas. A ausência de defeitos nestas cenas foi fundamental para a obtenção de resultados estáveis.
Vídeos degradados artificialmente	Via <i>software</i>	O processamento consistiu em introduzir degradações via software no material original. A simulação dos artefatos tem como referência os algoritmos desenvolvidos por (ALBINI, 2009).

Quadro 2 – Processamento dos vídeos utilizados nos testes

As condições de degradação foram utilizadas de forma a tornar o teste equilibrado e a garantir que todos os níveis da escala fossem utilizados.

3.2.4 Situações de Teste

As degradações introduzidas estão descritas no Quadro 3:

Processamento	Efeito
Varição de degradação na faixa de 2% a 5% dos <i>pixels</i> degradados	<i>Ruído Gaussiano</i>
Varição das máscaras: 3 <i>pixels</i> , 5 <i>pixels</i> , 7 <i>pixels</i> , 3x3 <i>pixels</i> , 5x5 <i>pixels</i> , 7x7 <i>pixels</i>	<i>Blurring</i>
Varição de 20% de blocos degradados	<i>Blocking</i>
1000 <i>pixels</i> degradados no quadro	<i>Sal e pimenta</i>
Duração de 10 segundos	<i>Flicker</i>

Quadro 3 – Situações dos testes e degradações

3.2.5 Preparação do material

A preparação do material é essencial para a obtenção dos resultados na avaliação subjetiva de vídeo digital.

3.2.5.1 Arranjo das sequências

O material preparado atendeu às condições de máxima aleatoriedade explorando todas as condições de teste e observando a limitação de 30 minutos, necessária para que se consiga manter a atenção dos avaliadores.

Os tipos de processamento apresentados totalizam 84 condições que foram exploradas. Essa amostra se mostrou suficiente para medir o ruído associado e a variabilidade das decisões subjetivas.

3.2.5.2 Ordenação

Foram preparadas dez sessões de apresentações que diferem entre si na escolha das sequências que foram utilizadas em cada uma das situações de teste descritas e na ordem em que estes itens são apresentados. A ordenação dos testes é diferente para cada sessão a fim de

compensar efeitos de adaptação ou de cansaço que pudessem vir a interferir no julgamento dos avaliadores.

As sequências foram arranjadas de forma pseudo-aleatória e cuidados foram tomados na ordenação para evitar a exibição sequencial de imagens perceptualmente “semelhantes”.

3.2.6 Forma de apresentação

Para a realização das avaliações foi utilizada uma TV LG *Full*-HD (1080p) de 42 polegadas². As condições de iluminação adotadas foram adaptadas a partir das referências descritas na norma (ITU-R, 2002). A distância média de visualização do avaliador foi de 1,20 m (3 x H, sendo H a altura da imagem/vídeo observada). A avaliação consistiu na apresentação aos avaliadores de 96 vídeos (24 para cada vídeo citado) apresentados de forma aleatória. O avaliador indica a sua nota conforme a escala utilizada na avaliação SDSCE. A Figura 29 apresenta a forma de apresentação das sequências de vídeos avaliados.

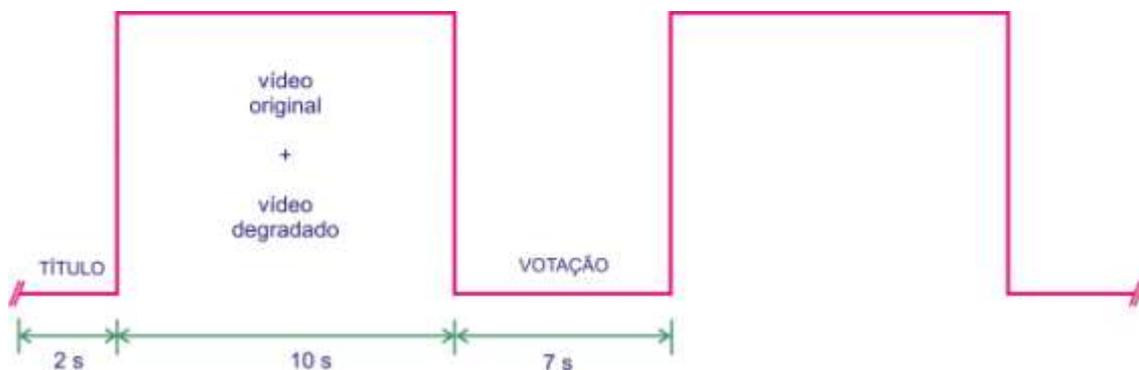


Figura 29 – Forma de apresentação das sequências de vídeos

3.2.7 Detalhes do teste

3.2.7.1 Avaliadores

Foram convidados 33 avaliadores para participarem do experimento inicial, dentre os quais, indivíduos de ambos os sexos, alunos e professores da UTFPR, com faixa etária entre 18 a 50 anos. Vale salientar que a norma (ITU-R, 2002) exige um mínimo de 15 participantes, no

² LCD TV LG Modelo 42LGGOFR

entanto neste trabalho foi utilizado mais que o dobro do necessário para validar a avaliação. Todos atestaram não ter experiência alguma em avaliação de vídeos digitais e processamento de imagem e vídeo.

As sessões de avaliação produziram um total de 3168 notas (96 notas para cada avaliador) para serem compiladas.

3.2.7.2 Votação

A duração de uma sessão de avaliação é um compromisso entre a necessidade de se ter redundância explorando diferentes parâmetros e o tempo em que é possível manter concentrada a atenção do avaliador. O tempo de duração seguiu as recomendações das normas contidas em (ITU-R, 2002), e foi de aproximadamente 30 minutos.

Para cada sessão, os avaliadores foram apresentados ao método de avaliação e às degradações que aparecem no teste, através das sequências de treinamento. As notas dadas as sequências de treinamento foram descartadas dos resultados finais. As instruções incluíram informações sobre o tipo de avaliação, escala utilizada, temporização das sequências e duração do intervalo de votação.

A votação das sequências foi escolhida através de uma escala de degradação por categorias com 5 pontos, formada pelos cinco níveis de degradação estabelecidos pelo ITU-R (ITU-R, 2002). Notas de 1 a 5, igualmente espaçadas, foram relacionados à escala adotada como mostrado na Figura 30.

Nota	Categoria	Posição na escala
1	Imperceptível	<input type="checkbox"/>
2	Perceptível, mas não irritante	<input type="checkbox"/>
3	Ligeiramente irritante	<input type="checkbox"/>
4	Irritante	<input type="checkbox"/>
5	Muito irritante	<input type="checkbox"/>

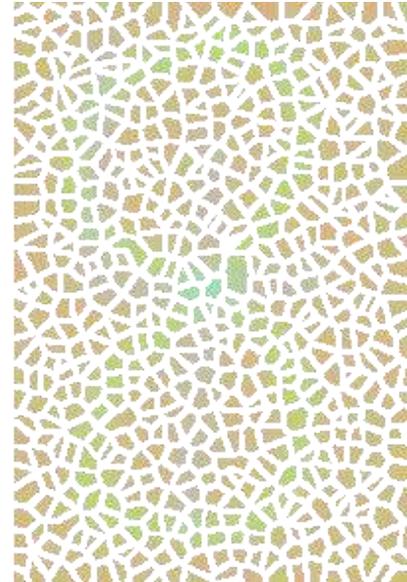
Figura 30 – Escala utilizada pelo SDSCE (Adaptado de (ITU-R, 2002))

Em resumo, o procedimento englobou as seguintes etapas:

- (i) Teste de acuidade visual (TESTE DE ACUIDADE VISUAL, 2008) e de daltonismo utilizando o teste de *ishihara* (TESTE DE *ISHIHARA*, 2008); A Figura 31 (a) e (b) apresenta exemplos de ambos os testes aplicados.



(a) Acuidade Visual (Tabela de Snellen) Fonte: (TESTE DE ACUIDADE VISUAL, 2008)



(b) Exemplo Teste de *Ishihara* Fonte: (TESTE DE *ISHIHARA*, 2008)

Figura 31 – Teste de Acuidade visual (a) e Exemplo de teste de *Ishihara*

- (ii) Treinamento e apresentação das possíveis imparidades encontradas (Blocagem, Borrramento, Ruído Gaussiano, Ruído sal e pimenta e travamento); e
 (iii) Apresentação da sequência de vídeo.

Nesta etapa o procedimento de coleta de dados se deu de forma manual utilizando-se um formulário padronizado que continha todos dados (informações pessoais, escala utilizada, área de observações) que foram posteriormente tabulados.

3.3 RESUMO E CONCLUSÃO DO CAPÍTULO

Neste capítulo destacaram-se os fundamentos metodológicos que nortearam a avaliação subjetiva preliminar que serviu de base para especificação, projeto e desenvolvimento de um ambiente de auxílio à avaliação subjetiva, objetivo principal deste trabalho.

Aqui também foi descrito o experimento de avaliação subjetiva preliminar que possibilitou a geração dos requisitos necessários ao desenvolvimento das ferramentas que

compõem o ambiente de avaliação subjetiva. O fluxo de dados e os procedimentos de teste realizados geraram resultados que foram utilizados como parâmetros de desenvolvimento.

A seguir, apresentam-se o processo de desenvolvimento do ambiente proposto e de suas ferramentas.

4 MODELAGEM E PROJETO DO AMBIENTE E SUAS FERRAMENTAS

Neste capítulo apresentam-se a modelagem e o projeto das ferramentas de auxílio à avaliação subjetiva propostas neste trabalho. O mesmo concentra-se nos aspectos referentes à análise, modelagem e projeto orientado a objetos para dar forma aos requisitos colhidos ao longo da experimentação preliminar apresentada na Seção 3.2

Esta etapa é fundamental, pois define as estruturas e o desenho da aplicação que foi desenvolvida de forma a facilitar o entendimento do sistema implementado.

4.1 DESENVOLVIMENTO

Nesta seção, descrevem-se, de forma sucinta, alguns conceitos, ferramentas e modelos que foram utilizados para o desenvolvimento do sistema proposto.

Para o desenvolvimento do ambiente foram empregadas técnicas de engenharia e ferramentas de *hardware* e *software*.

A partir dos dados e das observações anotadas, partiu-se para a especificação, projeto e desenvolvimento de um ambiente de auxílio à avaliação subjetiva de vídeo. Tendo em vista a possibilidade de ter escalas de avaliações contínuas e discretas, concebeu-se o ambiente e os dispositivos que auxiliarão aos pesquisadores na simulação de artefatos, coleta e análise dos dados de avaliações subjetivas.

A seguir são descritos alguns conceitos importantes ao processo de desenvolvimento das ferramentas de *software*.

4.1.1 Análise de Requisitos

A análise de requisitos é o primeiro passo na modelagem do processo. Nesta fase a preocupação é no que se deve fazer e não com a forma como será implementado o sistema. É o processo de aquisição, refinamento e verificação das necessidades do sistema. O objetivo é sistematizar o processo de definição dos requisitos, obtendo uma especificação correta e completa do mesmo para elaboração do Documento de Requisitos (SILVA, 2009c).

Os requisitos estão divididos em: Requisitos funcionais, requisitos não funcionais e requisitos de interface (PRESSMAN, 2006). Os requisitos funcionais são funções de como o sistema deve reagir a entradas de dados e como deve ser seu comportamento em determinadas

situações. Os requisitos não funcionais tratam de outros requisitos como usabilidade, segurança, integridade, corretude, etc.

Para se obter a lista de requisitos do sistema podem-se utilizar as seguintes técnicas: entrevista, questionário, observação direta e sessões de *brainstorming*. Para o sistema aqui proposto foram utilizadas as técnicas de entrevistas, questionários e observação direta a partir do experimento de avaliação subjetiva preliminar descrito na seção 3.2.

Levando-se em conta o experimento apresentado, foi possível realizar a especificação dos requisitos baseados nas normas contidas em (ITU-R, 2002), observações dos avaliadores participantes do experimento e experimentos subjetivos realizados em (ALBINI, 2009) chegou-se a lista dos requisitos funcionais e não funcionais, como segue:

Requisitos Funcionais:

- Calcular os valores das métricas objetivas de acordo com a necessidade da pesquisa em foco.
- Montar as sequências aleatórias dos vídeos a serem exibidos para cada seção, bem como armazenar os nomes originais dos arquivos para recuperação posterior dos mesmos.
- Armazenar as notas dadas pelos avaliadores para posterior geração dos resultados.
- Calcular o MOS das notas dadas pelos avaliadores para compor a análise no final do procedimento de testes.
- Representar todas as escalas de avaliação subjetiva contidas na norma (ITU-R, 2002). Esta representação será feita por meio de dispositivo eletrônico que representará os tipos de avaliação subjetiva por meio de botões e de um potenciômetro.
- Gerar os gráficos dos resultados da comparação entre as métricas objetivas estudadas e o seu respectivo teste subjetivo, a fim de estabelecer a correlação entre as mesmas.
- Enviar mensagens informativas para o avaliador tanto no monitor utilizado na avaliação, quanto no visor de LCD acoplado ao dispositivo eletrônico de avaliação subjetiva.

Requisitos não-funcionais:

- O ambiente deve prover iluminação adequada e regulável para as avaliações de vídeo.
- A distância dos avaliadores ao monitor de vídeo deve ser medida de acordo com as normas contidas em (ITU-R, 2002).

- As escalas utilizadas vão variar de acordo com o tipo de método de avaliação utilizado, podendo ser discreto ou contínuo.
- O monitor utilizado deverá estar de acordo com as normas estabelecidas (ITU-R, 2002, ITU-T, 1999; VQEG, 2008; EBU, 2003) para a utilização de monitor de LCD.
- O sistema deverá possibilitar respostas esclarecedoras e que auxiliem o avaliador, tanto no visor de LCD quanto no monitor de apresentação dos vídeos.
- O sistema poderá ser expandido acrescentando módulos externos.
- Os usuários dos sistemas serão os pesquisadores do grupo de qualidade de vídeo, avaliadores experientes ou não experientes em avaliações de vídeos digitais.
- O tempo de resposta do sistema deverá ser adequado aos experimentos de forma a não estressar nem os pesquisadores nem os avaliadores.
- A documentação do sistema poderá ser acessada *online* ou em manual adequado.
- O acesso ao sistema será protegido por senha de acesso e as informações serão controladas.

O processo de extração de requisitos é parte fundamental em qualquer processo de desenvolvimento de *software/hardware* não importando a metodologia utilizada para isso. No caso do sistema aqui proposto onde se utilizou a modelagem orientada a objeto, requisitos mal formulados levariam a um desastre no que diz respeito à qualidade do sistema, bem como a um possível fracasso no projeto do sistema, elevando os custos de manutenção.

A seção seguinte apresenta a especificação técnica do sistema proposto.

4.2 ESPECIFICAÇÃO TÉCNICA

Na especificação técnica é apresentada uma visão geral do sistema configurável de avaliação subjetiva e uma descrição prévia de cada módulo do qual o projeto é composto.

O Diagrama de blocos do módulo de avaliação é apresentado na Figura 32.



Figura 32 - Diagrama em bloco do processo de votação

O Diagrama mostra a representação esquemática da votação via microcontrolador utilizados pelos avaliadores na sala de avaliação.

A Figura 33 apresenta um esquema geral do projeto. O mesmo é um esquema composto de *hardware/software*, gerando, ao final da avaliação, o valor das notas dadas durante o período de avaliação.



Figura 33 – Visão do geral do sistema

O sistema será composto de: um sistema de *hardware* (microcontrolador, LCD e dispositivos de entrada) conectados a uma cadeira para avaliação, que receberá as notas dos avaliadores dadas aos vídeos e visualizados via LCD no kit do microcontrolador. Este dispositivo será conectado ao microcomputador via interface serial que se liga a um sistema de banco de dados. Mais detalhes do funcionamento são detalhados nas seções posteriores e em (SILVA, 2009c; SILVA, 2009d).

4.3 ASPECTOS FUNCIONAIS

O Sistema de Automatização de Avaliações Subjetivas de Vídeos Digitais – SAASVD trata da contabilização das notas dadas pelos avaliadores através do módulo de avaliação direta pelo usuário. Após receber a mensagem, o sistema de Banco de Dados armazena temporariamente os dados para posterior manipulação. O projeto foi desenvolvido de forma a garantir a utilização dos mais diversos tipos de avaliações subjetivas tratados em (ITU-R, 2002). Além disso, permitirá a inclusão de módulos de cálculo de métricas objetivas trazendo com isso, maior flexibilidade para os pesquisadores.

O *software* é responsável por realizar o acesso ao banco de dados e foi construído através de técnicas de Engenharia de Software. O *software* permitirá a recepção dos dados via porta serial da memória do kit do microcontrolador e os avaliadores serão orientados sobre o tipo de avaliação e qual dispositivo utilizar (teclado ou potenciômetro). Uma vez escolhido o tipo de avaliação e o dispositivo de interfaceamento, os vídeos serão apresentados de acordo com a norma e o tipo escolhido para a pesquisa. Após lançamento das notas, o sistema enviará as mensagens contendo as notas para o microcontrolador e as mesmas aparecerão no *display* de LCD localizado na cadeira, e após essa etapa, o sistema envia a mensagem do microcontrolador ao programa, que finalmente enviará ao banco de dados.

O sistema via interface direta com o *kit* do microcontrolador possui uma memória EPROM para armazenar o *firmware* que terá a função de comunicar os dados que saem dos dispositivos de votação (teclado ou potenciômetro linear) através da porta serial e chegam ao computador que possui um *front-end* em Java que enviará para o banco de dados (MySQL) em tempo real. A Figura 34 apresenta o braço desenvolvido para avaliação.

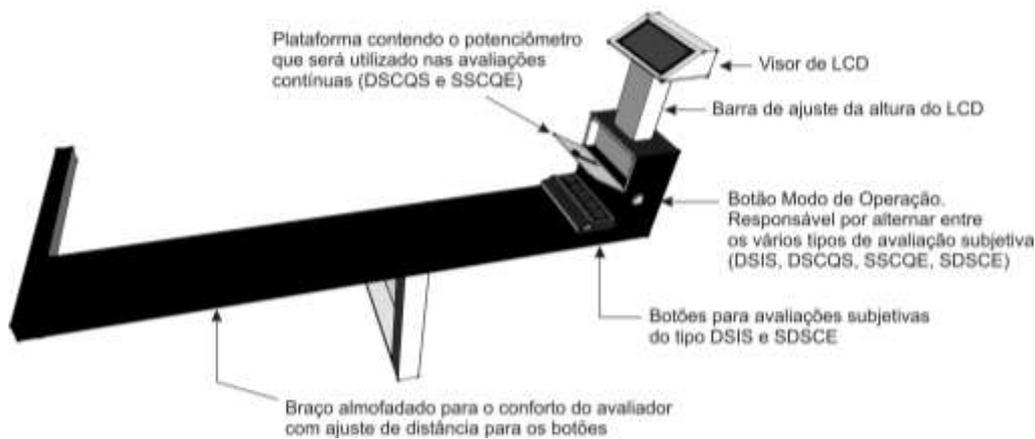


Figura 34 – Braço desenvolvido para a automatização das medidas subjetivas

4.3.1 Especificação do *Hardware*

Nesta seção apresenta-se a descrição do *hardware* utilizado como parte do desenvolvimento da cadeira de avaliação que faz parte do SAASVD.

4.3.1.1 Funções do *Hardware*

O *hardware* utilizado para a implementação do projeto foi o microcontrolador Motorola MC68HC908GR4 (composto por uma memória *flash* - para armazenar o programa em linguagem de máquina, memória RAM para dados temporários e o próprio microcontrolador que faz o processamento das notas da interface externa para o usuário realizar a avaliação direta via *hardware*, utilizando o teclado ou o potenciômetro deslizante, conforme o tipo de avaliação e duração do vídeo, e visualização via LCD da nota atribuída. A seguir resumem-se os componentes de *hardware* utilizados para no. Em SILVA (2009d) apresenta-se os *layouts* das placas de circuitos desenvolvidas para o projeto.

Hardware: microcontrolador Motorola MC68HC908GR4, porta serial padrão RS-232 para comunicação com o sistema de software, Visor de LCD, botões (total de cinco para avaliação e um para mudança de escala) e (um) potenciômetro, gabinete para montagem do dispositivo, cabos e fontes de alimentação.

A interface com o usuário é realizada através do hardware elaborado para a comunicação com o computador servidor.

Em SILVA (2009d) é apresentada a planilha com o investimento necessário para levar a efeito o projeto, levando-se em consideração os componentes e *softwares* utilizados. Não estão sendo considerados os custos de utilização dos equipamentos como: osciloscópio, fontes e outros, pois os mesmos estão disponibilizados no laboratório.

O projeto do *hardware* e suas especificações e os *layouts* das placas de circuito impresso do projeto encontram-se em SILVA (2009d).

4.3.2 Especificação de Software

Nesta seção apresenta-se a descrição do *software* utilizado para a implementação da ferramenta de avaliação que faz parte do SAASVD.

4.3.2.1 Linguagem e Ferramentas de Desenvolvimento

As ferramentas utilizadas para o desenvolvimento do software estão listados a seguir.

Software: Modelagem (UML) e Programação Orientada a Objetos (Java) para o *front-end*, Ferramenta de Modelagem UML *Judy Community*, Linguagem de Programação C para o *firmware*, Ambiente integrado de software livre Eclipse, Banco de Dados Relacional *Open Source* MySQL e *MySQL Workbench* para a geração dos *scripts* e modelo relacional da aplicação respectivamente, *Eagle CAD Layout Editor* para a construção dos *layouts* dos circuitos impressos e *GNUPlot* para a geração dos gráficos da aplicação.

4.4 ARQUITETURA DO SISTEMA PROPOSTO

Nesta seção descrevem-se a arquitetura do software desenvolvido. O objetivo é reconhecer a arquitetura de software do sistema e sua relação com o padrão ISO/IEEE 1471 (IEEE 1471, 2007), identificar as características e benefícios da arquitetura utilizada.

4.4.1 Motivação para desenvolver softwares melhores

Desenvolver *software* de qualidade não é uma tarefa trivial. É por esse motivo que vários projetos de *software* fracassam durante seu desenvolvimento. Entre esses maus resultados encontram-se os que custaram muito acima do orçamento, os incompletos e os que não

solucionam os problemas para os quais foram projetados (BOOCH, JACOBSON RUMBAUGH, 2006).

A complexidade do processo de desenvolvimento de *software* é um complicador. Além de lidar com a complexidade inerente ao problema em análise, deve-se também ter a preocupação de como o *software* resolve o problema. Assim, o *software* deve além de resolver o problema, resolvê-lo da forma esperada e com a qualidade devida.

4.4.2 Arquitetura do sistema proposto

No caso do sistema proposto optou-se pela arquitetura de software MVC (*Model, View, Controller* - Modelo, Visualização e Controle) (MVC, 2008).

A arquitetura MVC fornece uma maneira de dividir a funcionalidade envolvida na manutenção e apresentação dos dados de uma aplicação. A arquitetura MVC não é nova e foi originalmente desenvolvida para mapear as tarefas tradicionais de entrada, processamento e saída para o modelo de interação com o usuário. A Figura 35 ilustra a divisão das camadas da aplicação.

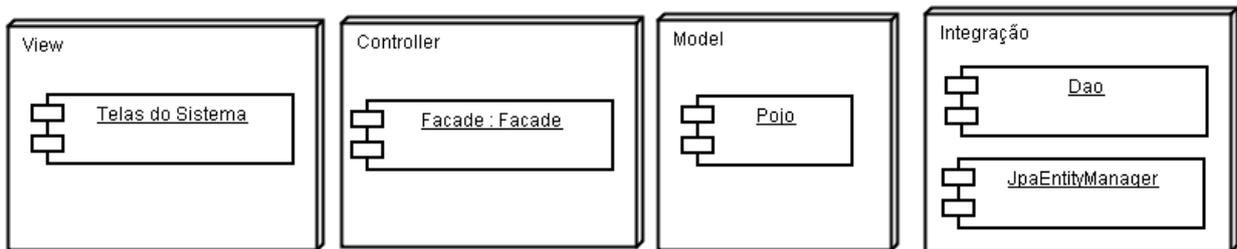


Figura 355 – Camadas da Aplicação

View. É a camada de visão da aplicação, responsável por tratar da interação com o usuário via apresentação de telas e execução de lógica de interface visual.

Controller. É a camada de controle da aplicação. Esta camada é chamada pela camada de visão para executar a lógica de negócio da aplicação, como cálculos ou persistência de informação em banco de dados.

Model. É a camada que mantém o estado da aplicação, os dados que são transmitidos e apresentados para o usuário.

Integração. É a camada que possui os repositórios de dados e as aplicações ligadas com as quais a aplicação interage.

A seguir apresentam-se os diagramas de arquitetura do Sistema proposto. A Figura 36 apresenta o diagrama de classes da arquitetura.

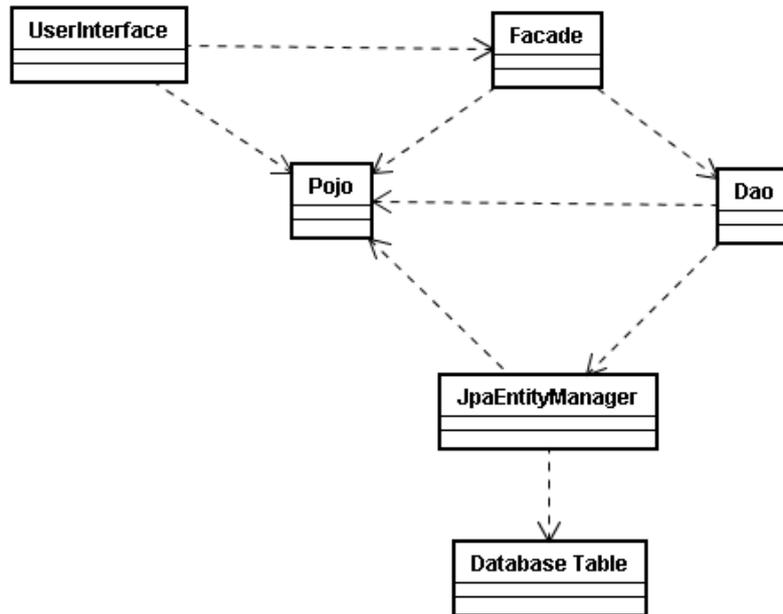


Figura 366 – Diagrama de classes da arquitetura

Segue uma breve descrição da arquitetura proposta:

UserInterface (Interface com o usuário). São as classes responsáveis pela interface gráfica do sistema.

FAÇADE. É o ponto de chamada para utilizar as funcionalidades de negócio da aplicação. Por exemplo, a funcionalidade de calcular os valores das métricas objetivas implementadas no sistema, será representado por um método em um *façade* apropriado para implementar e executar este cálculo.

A vantagem de se utilizar *façades* é encapsular e isolar o código de negócio das classes que cuidam da interação com o usuário ou com outros sistemas. Facilitando a reutilização de uma mesma funcionalidade de negócio por duas ou mais telas, ou por um ou mais sistemas ou componentes deste.

O *Facade* é um Padrão de Projeto definido pelo GoF (*Gang of Four*) (GAMMAN et al, 1994) como uma boa prática para implementar, encapsular e reutilizar código de negócio.

POJO (Plain Old Java Object). Trata-se de um objeto Java utilizado para implementar o modelo de domínio da aplicação. A sua utilização serve para trafegar um conjunto de valores entre as

camadas e as classes da aplicação, permitindo assim eliminar o uso de *Transfer Objects* ou *Value Objects*, simplificando a aplicação e melhorando a produtividade e manutenibilidade.

DAO. O DAO (*Data Access Object*) é o ponto de chamada para utilizar os recursos do banco de dados, tanto para consulta (*select*), quanto para manipulação de dados (inserção, atualização, deleção).

A vantagem de se utilizar DAO's é permitir a separação inteligente de código de negócio do código da manipulação do banco de dados. Isso possibilita ter independência de dados de tal forma que a mudança do mesmo não gera manutenção nas classes de negócio.

O DAO é uma boa prática catalogada no *Sun Blueprints Design Patterns* (BLUEPRINTS, 2008) como uma forma de implementar código de acesso a recursos de banco de dados e legados que fornecem e consomem informação.

JpaEntityManager/Tabela do banco de Dados (*Database Table*). O *JpaEntityManager* é o gerenciador de persistência de banco de dados definido na especificação JPA (*Java Persistence API*) (JPA, 2008). É um componente que permite acesso objeto-relacional aos recursos do banco. Ao invés de salvar dados de uma classe, criando um comando “*insert into*”, por exemplo, é solicitado ao *JpaEntityManager* que grave a classe diretamente em banco de dados, sem que haja necessidade de escrever código SQL para inserir na tabela do banco de dados.

A seguir descrevem-se as tecnologias envolvidas na construção do sistema, conforme apresentado na Figura 37.



Figura 377 – Tecnologias envolvidas na implementação do sistema

Swing. É uma API (*Application Programming Interface* - Interface de Programação de Aplicativos) Java para interfaces gráficas. Ela é compatível com a API AWT (*Abstract Window Toolkit* – Caixa de ferramentas abstratas para Windows), mas trabalha de uma maneira

totalmente diferente. A API *Swing* desenha os componentes sem a necessidade de ocupar o sistema operacional (DEITEL, 2005).

Spring. É um *framework* de desenvolvimento Java centrado nos conceitos de desenvolvimento leve e ágil, baseada em objetos POJO. O *Spring* permite a implementação simplificada de IoC, Aspectos e injeção de código, além de facilitar a integração com diversas soluções no mundo Java. Possui dezenas de classes de utilitários que permitem agilidade no desenvolvimento.

Na arquitetura proposta o *Spring* será utilizado na camada de negócios e de persistência (*Hibernate*), bem como poderá servir para realizar a integração entre as camadas *web* (JSF) numa possível extensão do sistema.

JPA/Hibernate. O JPA é a mais recente especificação criada para permitir persistência objeto-relacional com bancos de dados relacionais. Ao invés de salvar dados em tabelas, o código do sistema pede pela persistência de classes carregadas com os valores que se quer salvar. Assim, os produtos que implementam JPA, como o *Hibernate* (HIBERNATE, 2008), transformam as requisições de salvar e consultar via classes, em comandos SQL que são enviados ao banco de dados.

Eclipse. A plataforma Eclipse é um ambiente integrado de desenvolvimento (IDEs), onde a indústria de *software* possa desenvolver diversos programas, aplicativos e ferramentas, de forma otimizada e padronizada, baseando-se nas iniciativas de software livre (ECLIPSE, 2008).

A seguir descreve-se o projeto do sistema em detalhes, destacando os diagramas UML que foram desenvolvidos.

4.5 PROJETO

O projeto abrange a forma de implementação do *software* e *hardware* conforme descrito nas seções anteriores. Aqui, apresentam-se os diagramas de casos de uso, sequência e de classe (BOOCH, RAMBAUGH e JACOBSON, 2000; LARMAN, 2004), o MER (Modelo Entidade-Relacionamento), o Modelo Relacional, o fluxograma do *firmware* com seus diagramas de estados do projeto completo. Na Tabela 3 são descritas as configurações do computador que foi utilizado no projeto.

Tabela 3 – Configurações do microcomputador utilizado nos testes

Componente	Configuração
Processador	Pentium IV 1.8 GHz
Memória RAM	1 GB
Placa de vídeo	NVIDIA GeForce 7300 SE

A seguir descrevem-se os diagramas de modelagem do sistema que foram gerados a partir da análise de requisitos do sistema. Com o intuito de manter a clareza e a concisão no desenvolvimento desse trabalho, foram apresentados somente os diagramas principais e a documentação completa da modelagem e projeto do sistema se encontra no Relatório Técnico desenvolvido por (SILVA, 2009c).

Casos de Uso

Casos de Uso são descrições de interações típicas entre os usuários de um sistema e o sistema propriamente dito. Eles representam a interface externa do sistema e especificam um conjunto de exigências do que o sistema deve fazer. Diagramas de Caso de Uso descrevem relacionamentos e dependências entre um grupo de Caso de Uso e os Atores participantes no processo. A Figura 38 apresenta o diagrama de caso de uso da aplicação:

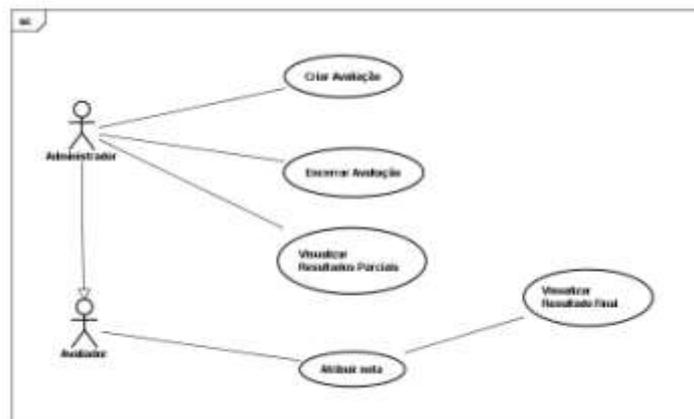


Figura 388 – Diagrama de Casos de Uso da aplicação

Existem dois atores nesse diagrama. Um deles é o administrador que criará e encerrará a avaliação, bem como poderá visualizar resultados parciais estatísticos. O segundo ator (avaliador) terá somente a função de realizar a avaliação criada pelo administrador. Note que o

administrador também pode ter a função de realizar a avaliação, porém o usuário não terá acesso a nenhuma função do administrador.

Diagrama de Classe

Um diagrama de classes mostra um conjunto de classes, interfaces e colaborações e seus relacionamentos. Usa-se esse diagrama para ilustrar a visão estática do projeto de um sistema. O diagrama de classe do sistema é mostrado na Figura 39:

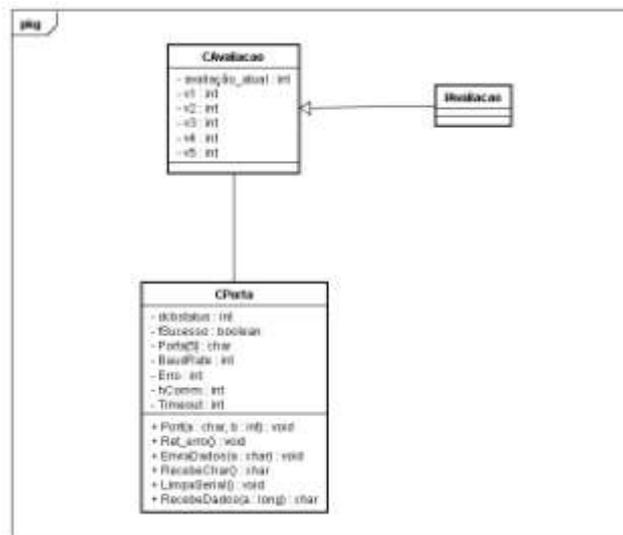


Figura 39 – Diagrama de Classe da aplicação

A classe CAvaliacao é a classe base para as demais classes. Ela fará o controle do que passa pela porta serial do microcontrolador, as chamadas das funções da classe CPorta e todas as funções de interface.

A classe CPorta fará a comunicação das notas recebidas via interface direta com o usuário (teclado ou potenciômetro), sendo assim enviadas para o banco de dados.

A classe IAvaliacao será a interface direta com o administrador para o início e conclusão da avaliação, além da opção de pesquisa temporária que poderá realizar.

Diagramas de Sequência

Um diagrama de sequência dá ênfase à ordenação temporal das mensagens.

A seguir são apresentados os diagramas de sequência na Figura 40, baseados no diagrama de casos de uso da aplicação:

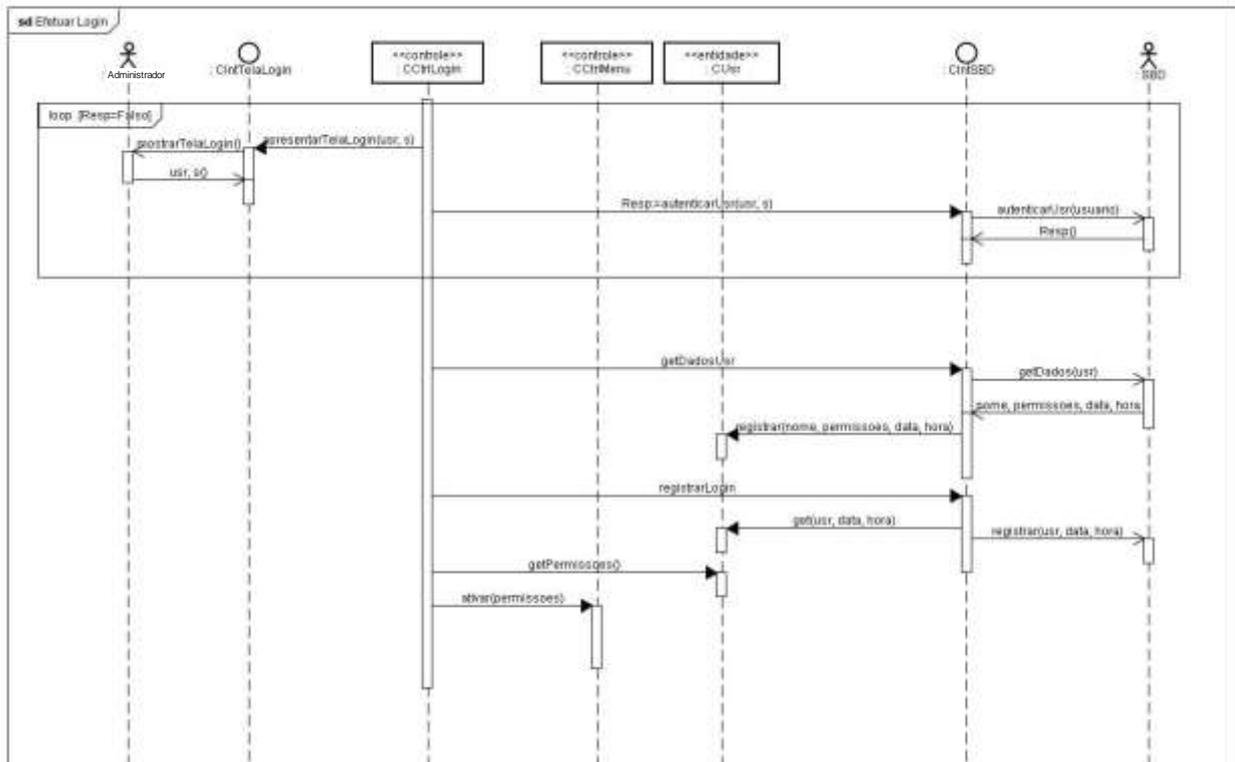


Figura 400 – Diagramas de Sequência

Diagrama de Estados

Os diagramas de estados são empregados para a modelagem dos aspectos dinâmicos de um sistema. Na Figura 41, pode-se observar o funcionamento do *firmware* através de um DTE.

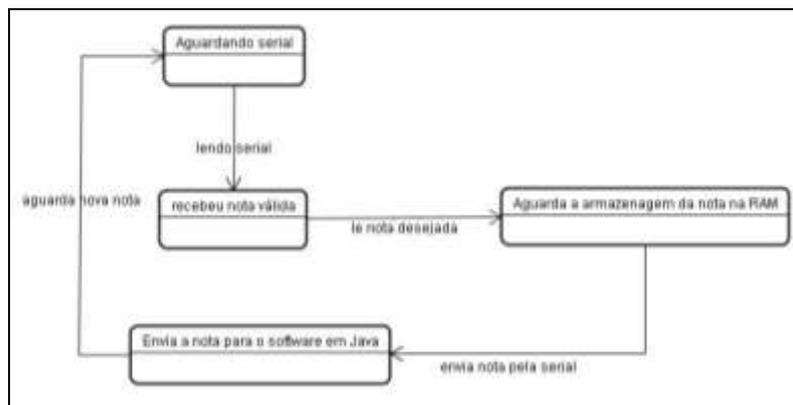


Figura 411 – Diagrama de Estados da aplicação

Interface do Software

A Figura 42 apresenta a interface principal da aplicação:

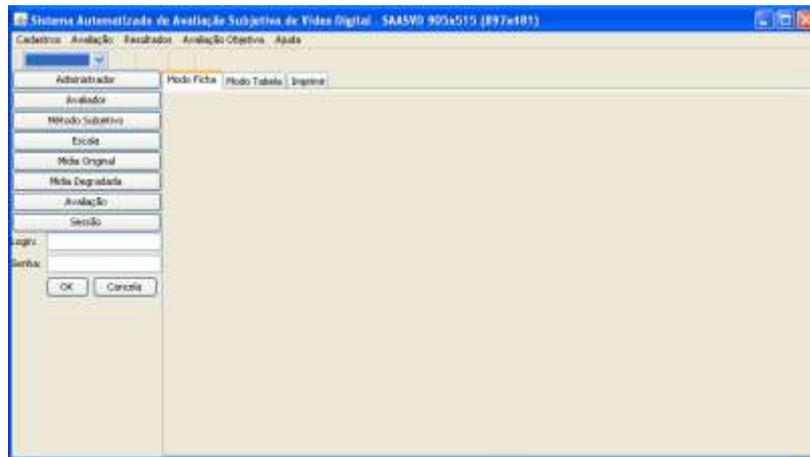


Figura 422 – Interface principal da Aplicação

4.6 PROJETO DO BANCO DE DADOS

Um projeto de banco de dados consiste nas seguintes partes: Modelagem conceitual, Modelagem Lógica e Modelagem Física (implementação) (ELSMARI e NAVATHE, 2005).

4.6.1 Requisitos do Sistema de Banco de Dados do SAASVD

O objetivo geral do sistema de Banco de Dados do SAASVD será de armazenar e permitir a manipulação dos dados das avaliações subjetivas/objetivas em estudo pelo grupo de pesquisa em TV Digital. De forma sucinta seguem os requisitos que foram contemplados no SAASVD:

- Permitir a avaliação de um ou mais vídeos digitais e, de acordo com uma escala, definida pela norma ITU-R dar uma nota pelo vídeo assistido ou por trechos do vídeo assistido de acordo com o tipo de avaliação utilizada e o tempo de exibição.
- As avaliações subjetivas podem ser de tipos diferentes (discretas ou contínuas), com referências ou sem.
- Cada sessão de avaliação durará um tempo finito e não pode passar de 30 minutos incluindo as explicações e exemplos.

- Todos os vídeos deverão ser sorteados aleatoriamente e não há repetição de sequência de apresentação de uma sessão para outra.
- O avaliador responde a um questionário onde informa seus dados (o nome, se tem experiência com vídeo, faixa etária e faz um teste de acuidade visual).
- Os vídeos são apresentados e o avaliador observa e dá uma nota dentro da escala da avaliação.
- Cada arquivo de vídeo possui um formato e resolução específicos.
- Emitir relatórios e estatísticas sobre as avaliações objetivas testadas e sua relação com o MOS das avaliações subjetivas realizadas.
- Calcular Métricas Objetivas que sejam objeto de estudo do grupo de pesquisa em TV Digital da UTFPR.
- Permitir a inserção de várias métricas objetivas a fim de testar a sua qualidade.
- Calcular o desvio padrão, média e variância das amostras coletadas.

Com base nessas especificações de requisitos foi modelado o Modelo Entidade-Relacionamento (MER) da aplicação e posterior mapeamento para o Modelo Relacional do SAASVD.

4.6.2 Modelo Entidade-Relacionamento e Modelo Relacional

Para o armazenamento das notas realizadas pelo módulo de avaliação, foram modeladas algumas entidades que representam o esquema de dados, que inclui: armazenamento das notas, controle do usuário, controle de notas dos usuários para evitar erros de avaliação. A Figura 43 apresenta o modelo entidade-relacionamento (MER) da aplicação.

O MER tornou-se padrão para modelagem conceitual de Banco de Dados por sua simplicidade (entidades, relacionamentos, atributos) e representação gráfica (retângulos, losangos, elipses e linhas) de fácil compreensão. Um esquema conceitual de dados é também chamado de esquema ER ou digrama ER (ELSMARI e NAVATHE, 2005).

A partir do modelo ER gerado da análise de requisitos do banco de dados gerou-se, via *MySQL Workbench*, o modelo Relacional do sistema. A Figura 44 apresenta o modelo relacional do sistema de avaliação.

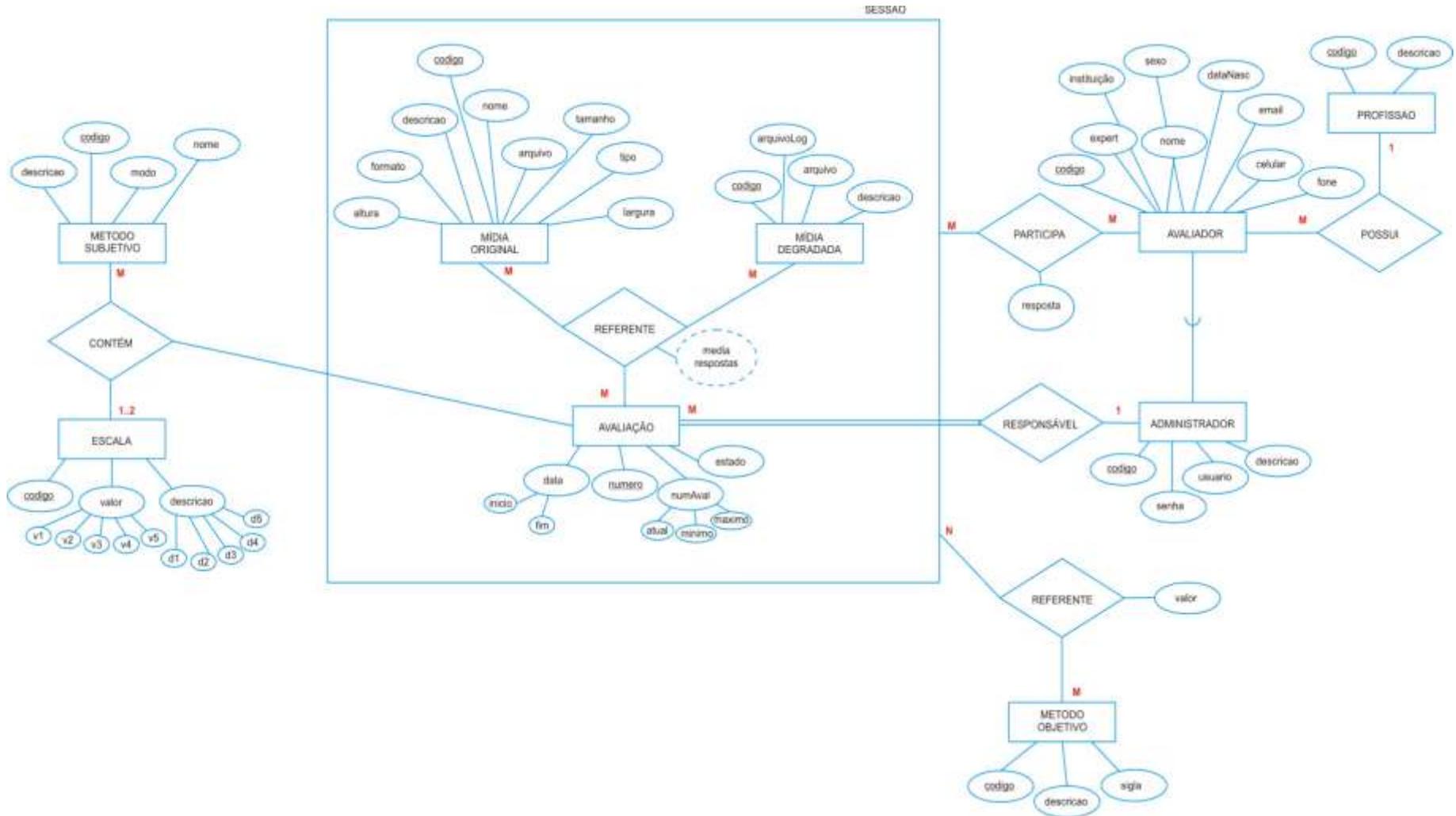


Figura 433 – Modelo Entidade-Relacionamento da Aplicação

O MER da aplicação representa a modelagem dos requisitos do SAASVD e a partir dele, tem-se uma visão dos dados que serão armazenados na aplicação.

A leitura do MER apresenta as seguintes relações e entidades:

- Os métodos subjetivos possuem até duas escalas e uma escala pode ser utilizada por mais de um método subjetivo.
- Uma avaliação está relacionada com métodos subjetivos e uma escala vinculada a esse método.
- Uma sessão é composta das mídias (original e degradado) e sua respectiva avaliação.
- Um avaliador participa de uma sessão e cada esta possuirá os mesmos vídeos apresentados em sequências diferentes.
- O administrador será responsável por controlar a avaliação.
- Para cada vídeo será feito o cálculo da métrica objetiva utilizada no experimento.

Após a modelagem conceitual, definição das entidades, relacionamentos e atributos dessas entidades, mapeou-se este MER para o seu respectivo Modelo Relacional, representado-se assim, as entidades, relacionamento e atributos em tabelas e campos que foram armazenadas no banco de dados MySQL, como mostrado na Figura 44.

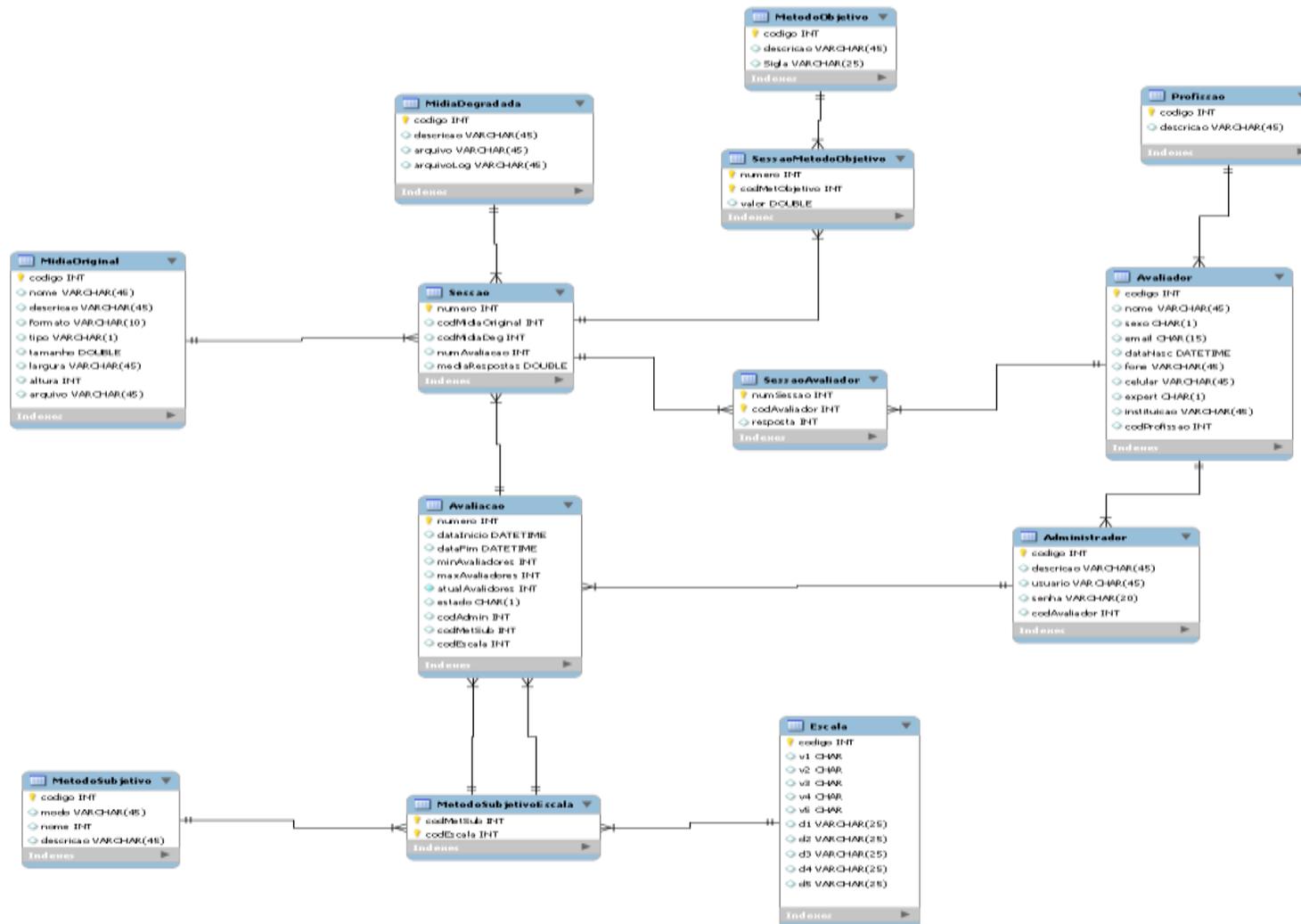


Figura 444 – Modelo Relacional da Aplicação

A partir desses dados foi implementado o projeto físico do banco de dados da aplicação. Os *scripts* de criação do banco de dados encontram-se em SILVA (2009c).

4.7 PROJETO DE *FIRMWARE*

O *firmware* é o programa que roda na memória EPROM do microcontrolador para funcionamento dos botões, do potenciômetro e da serial para transferência dos dados do *kit* do microcontrolador para o banco de dados no servidor. Na Figura 45, mostra-se o funcionamento do mesmo:

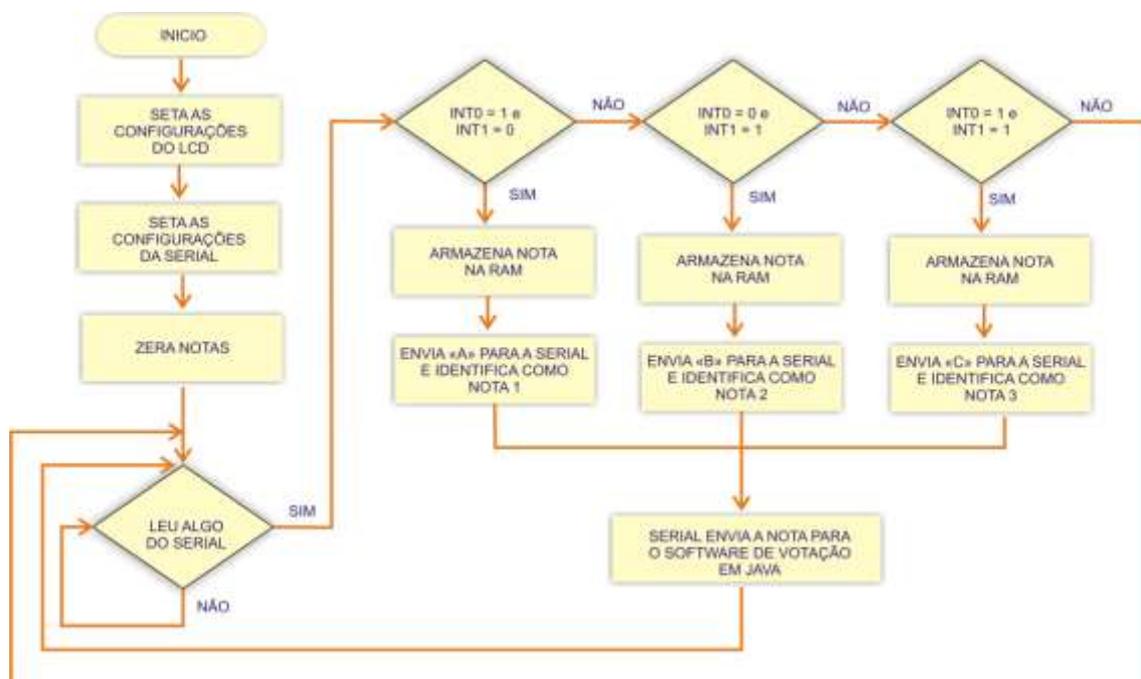


Figura 455 – Diagrama em bloco do *Firmware*

O fluxograma acima descreve as rotinas referentes à interface com botões: é estabelecida a configuração do LCD, os níveis de configuração dos pinos do microcontrolador, o endereço para escrita, etc. Após realizar a configuração do LCD, deve-se configurar a porta serial, que trabalhará com 9600 bps. Para garantir que nenhum dos dados está armazenado na RAM, opcionalmente é inserido um ZeraNotas para configurar a variável da serial para zero. Se for lido algo da serial, ela verificará qual botão (caso seja utilizado a interface de botões) foi pressionado e a respectiva nota será enviada para a RAM, que logo envia via serial para o *software* de

administração dos votos. Esse processo continuará até que seja encerrada a avaliação. Mais detalhes do projeto de *firmware* encontram-se em SILVA (2009c).

4.8 ESPECIFICAÇÃO DE TESTES DE VALIDAÇÃO DO PROJETO

Neste trabalho nos limitamos a um conjunto de testes funcionais, tendo em vista o escopo e os objetivos desse trabalho. O teste do software em Java foi realizado a fim de verificar-se a comunicação entre o microcontrolador MC68HC908GR4 (via porta serial) com o Sistema Gerenciador de Banco de Dados - SGBD (MySQL) para armazenamento das notas. Durante os testes foram verificadas a comunicação com o banco de dados, usabilidade e facilidade de uso pelo usuário, gerenciamento do software e extração de resultados do sistema.

Uma vez verificados a integração do sistema e o teste de suas funcionalidades, fez-se os testes com os avaliadores de forma a validar os testes subjetivos, conforme descrito no capítulo 5. Os testes com o *hardware* desenvolvido estão descritos em SILVA (2009c).

4.9 DIAGRAMA DE BLOCOS FINAL DO SAASVD

Fazendo-se, um paralelo com o Diagrama apresentado na seção 3.2 (Figura 26), pode-se notar claramente que vários passos na metodologia foram substituídos por processos automatizados. Dessa forma, conseguiu-se maior agilidade na verificação e validação dos resultados com o sistema proposto. Propiciando aos pesquisadores analisar as métricas objetivas estudadas e sua relação com as métricas subjetivas.

Desde o cadastro dos avaliadores até a finalização das avaliações com a geração dos gráficos, o processo inteiro foi automatizado, permitindo ganho em tempo em torno de 95% em relação ao tempo para obter os resultados no processo inicial que serviu como base para a extração de requisitos.

A Figura 46 mostra o processo de avaliação dos vídeos no ambiente proposto.

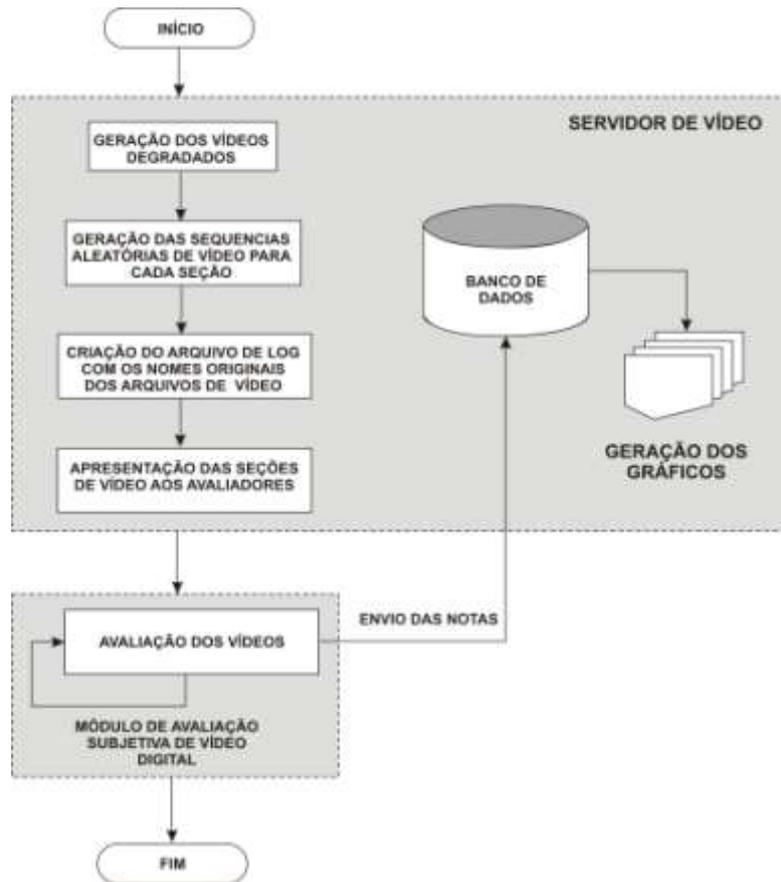


Figura 466 – Processo de avaliação dos vídeos no ambiente proposto

Em relação ao experimento preliminar apresentado na seção 3.2 (Figura 26, pp. 56), foram implementados todas as tarefas manuais descritas no diagrama de blocos da Figura 26 da referida seção.

Além disso, a ferramenta de *hardware* implementada possibilitou o descarte do questionário de avaliação para a anotação das notas dos avaliadores, bem como permitiu o envio automático das mesmas para o SGBD (Sistema Gerenciador de Banco de Dados) implementado para armazenar as notas subjetivas e valores objetivos calculados pelo sistema.

Seguindo essa tendência de automatização, no bloco 3 foi substituído o programa externo de geração de gráficos por uma chamada direta via script para um gerador de gráficos construído em *software* livre e de configuração aberta.

A seção a seguir, apresenta uma visão geral do ambiente de avaliação subjetiva construído.

4.10 VISÃO GERAL DO AMBIENTE DE AVALIAÇÃO SUBJETIVA

O sistema foi instalado em sala apropriada, conforme ilustrado na Figura 47. Na construção do ambiente de avaliação subjetiva estudos sobre iluminação (NBR 5413, 1992), distância dos observadores, cores das paredes foram realizados. Antenas de recepção do sinal digital foram instaladas, além das instalações elétricas e lógicas.

Algumas características são comuns a todos os procedimentos subjetivos para medir a qualidade de imagem de televisão, embora existam algumas diferenças entre condições de visualização para televisão com resolução convencional e para televisão de alta definição como podem ser vistos em (ITU-R, 2002; ITU-T, 2008; VQEG, 2008).

Ajustes e estudos anteriores foram utilizados para fazer adequações necessárias. Todos os parâmetros descritos têm influencia sobre a avaliação e devem ser mantidos tanto quanto possível próximos aos valores padronizados para permitir a comparação entre resultados obtidos por diferentes laboratórios.

Dessa forma, o ambiente foi equipado e os parâmetros foram mantidos dentro do estabelecido pelas normas vigentes conforme apresentado na seção 3.1.1 (Capítulo 3).



Figura 477 – Visão geral do ambiente de avaliação

No ambiente de Avaliação Subjetiva de Vídeo Digital estão à disposição uma TV LCD 42 polegadas *Full HD*, um tocador de *blu-ray*, dois conversores de TV digital, uma antena UHF,

um microcomputador conectado à TV para avaliação dos vídeos e processamento dos vídeos e o protótipo da cadeira de avaliação. A Figura 48 apresenta um avaliador utilizando a cadeira de avaliação.



Figura 488 – Avaliador fazendo uma avaliação utilizando o método SDSCE de vídeos no formato CIF.

A cadeira está equipada com um braço que foi desenvolvido para atender às necessidades referentes à avaliação subjetiva e faz parte do SAASVD proposto nesta dissertação.

A Figura 49 apresenta a concepção artística do sistema implementado, destacando os principais componentes distribuídos no ambiente de avaliação.

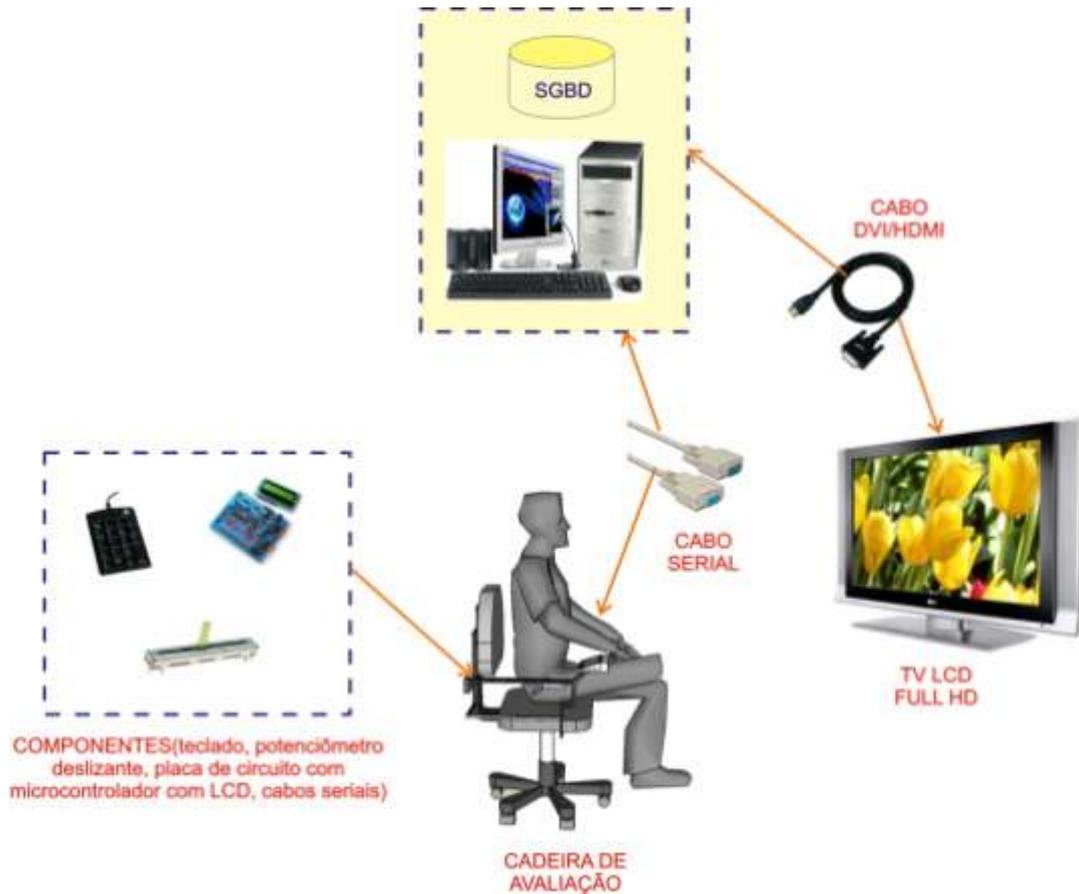


Figura 49 – Visão geral da avaliação no ambiente proposto

Conforme os objetivos iniciais, o sistema e o ambiente de avaliação foram implementados de forma a conduzir avaliações com qualidade, num ambiente planejado para dar conforto e construído dentro dos padrões e normas vigentes no que diz respeito às avaliações subjetivas de vídeo digital. Dessa forma, o grupo de pesquisa em componentes e sistemas de telecomunicações poderá conduzir avaliações de qualidade de vídeo de forma padronizada e condizente com o que está sendo desenvolvido pelas instituições de pesquisa que trabalham com essa temática de pesquisa.

4.11 RESUMO E CONCLUSÃO DO CAPÍTULO

Neste capítulo apresentou-se a modelagem das informações coletadas no Capítulo 3 e foi desenvolvida a especificação do ambiente composto de ferramentas de *hardware* e *software* implementadas para auxiliar na coleta dos dados e geração dos resultados estatísticos decorrentes das avaliações subjetivas.

Os diagramas de modelagem de dados, o projeto e implementação do banco de dados para armazenar os resultados das avaliações subjetivas, o projeto do *firmware* e os parâmetros para a construção do ambiente foram apresentados aqui.

Conclui-se o capítulo com uma visão geral do ambiente proposto e das ferramentas de forma que se pode comparar com o Diagrama presente na Seção 3.2 (Figura 26, pp. 56) e concluir que o processo, conforme descrito na Recomendação ITU-R (ITU-R, 2002), não foi alterado. No entanto, o mesmo foi aperfeiçoado no que diz respeito à velocidade e praticidade no armazenamento de dados subjetivos e objetivos, comparação de métricas e maior velocidade na geração dos resultados dos testes subjetivos.

A seguir apresentam-se os resultados gerados pela utilização do sistema proposto e a análise desses resultados no que tange à análise estatística dos resultados dos testes.

5 VALIDAÇÃO DO AMBIENTE E DE SUAS FERRAMENTAS

Neste capítulo, o foco se concentra na validação do ambiente proposto e suas ferramentas. No capítulo anterior, foi descrita a metodologia utilizada para o desenvolvimento e implementação do ambiente e suas ferramentas. A utilização das ferramentas permitiu uma redução no tempo de resposta das avaliações em torno de 95% em relação ao procedimento de aquisição dos resultados realizado na avaliação subjetiva preliminar apresentado na seção 3.2, uma vez que o tempo de avaliação continua o mesmo devido às restrições da própria norma (ITU-R, 2002). Este ganho reduz o tempo total do processo que envolve ainda a análise dos resultados dos testes efetuados com as métricas subjetivas e a validação das métricas objetivas estudadas.

Na seção que segue apresentam-se os resultados obtidos com os experimentos efetuados de avaliação subjetiva.

5.1 RESULTADOS PRELIMINARES COM AVALIAÇÃO MANUAL

Após os testes subjetivos e a recuperação das amostras de vídeos avaliados conforme a classificação adotada na nomenclatura dos arquivos e já com os valores de PSNR obtidos em etapa anterior partiu-se para a análise dos dados coletados a partir dos avaliadores. A Figura 50 apresenta uma imagem das seções de avaliação subjetiva utilizando avaliação manual.



Figura 50 – Sessão de avaliação Subjetiva utilizando avaliação manual

Foram *plotados* dois grupos de gráficos. O primeiro grupo (Figuras 51 e 52) relaciona o PSNR e o MOS obtido a partir das médias das avaliações com duas das degradações estudadas em (ALBINI, 2009) e o segundo apresenta a relação do PSNR com o MOS para cada tipo de vídeo observado.

As Figuras 51 e 52, respectivamente, apresentam os gráficos da PSNR pelo MOS de todos os vídeos analisados no experimento de avaliação subjetiva preliminar presente na seção 3.2 para os artefatos de borramento e ruído gaussiano.

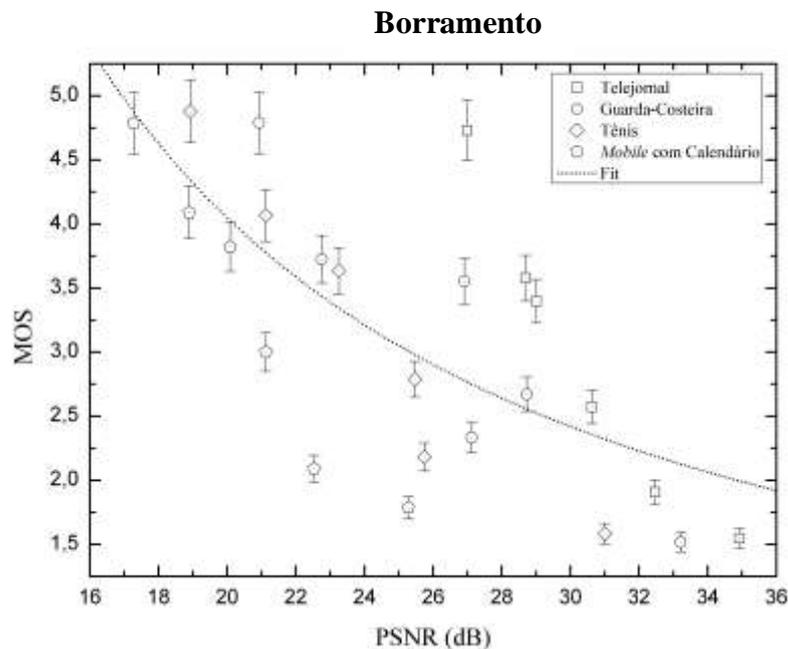


Figura 51 - Relacionamento entre o MOS e o PSNR com base nas degradações para o artefato de borramento

O Gráfico apresentado na Figura 51 apresenta a relação entre o PSNR e o MOS. O artefato aplicado aos vídeos analisados neste gráfico é o borramento, conforme mostrado na seção 2.5.3 e com máscaras de degradação conforme a seção 3.2.4 (Quadro 3). Vale salientar que a escala utilizada foi do método SDSCE (seção 2.3.1.1) e nessa escala, quanto mais próximo de 1 melhor a qualidade do vídeo observado.

O gráfico apresentado na Figura 52 apresenta a relação entre PSNR e o MOS e o artefato observado é o Ruído Gaussiano.

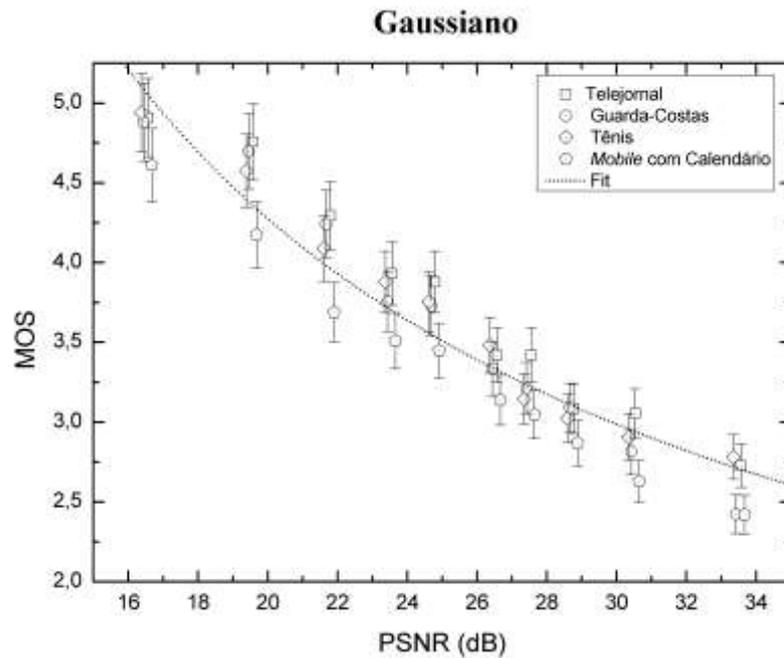
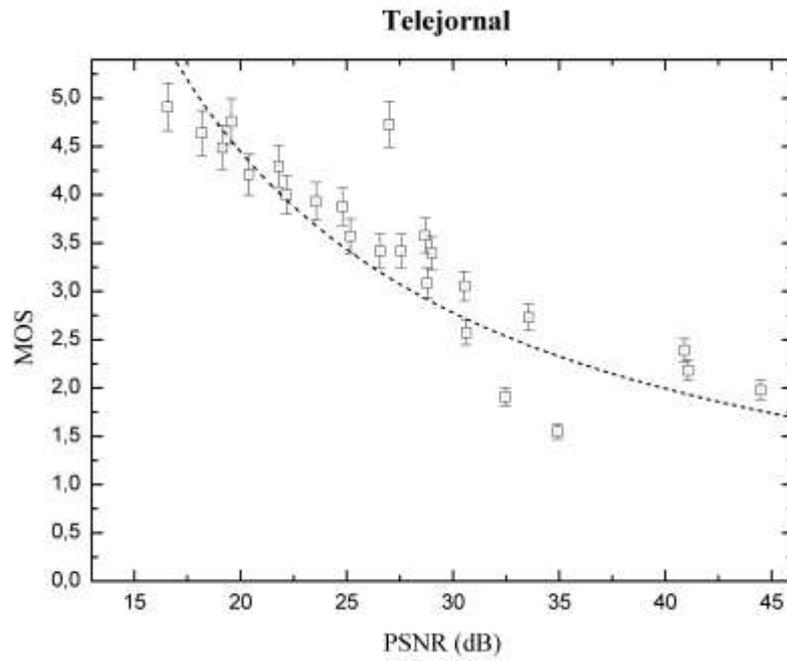


Figura 522 - Relacionamento entre o MOS e o PSNR com base nas degradações para o artefato de Ruído Gaussiano

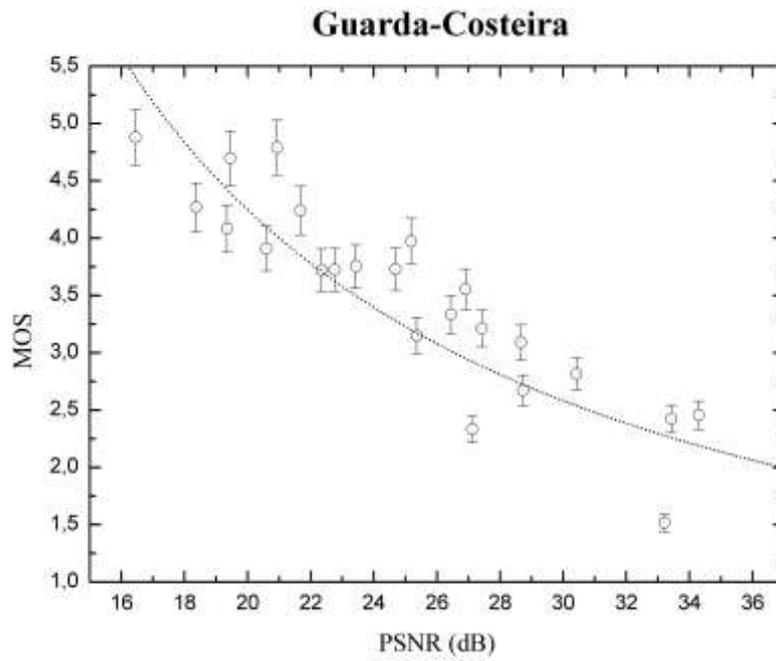
Foi observado durante as sessões que algumas degradações foram consideradas inaceitáveis do ponto de vista do SVH conforme relato dos próprios avaliadores.

Para exemplificar, com relação ao ruído branco gaussiano, observou-se que esse tipo de degradação, mesmo em pequenas proporções é tolerável pelos avaliadores, no entanto, conforme aumenta a sua incidência, este se torna extremamente intolerável. Foi notado também, que muitos avaliadores imediatamente após exibir o vídeo degradado com esse artefato que os mesmos já davam uma nota ruim para o vídeo, sem nem mesmo esperar pela finalização do tempo de exibição.

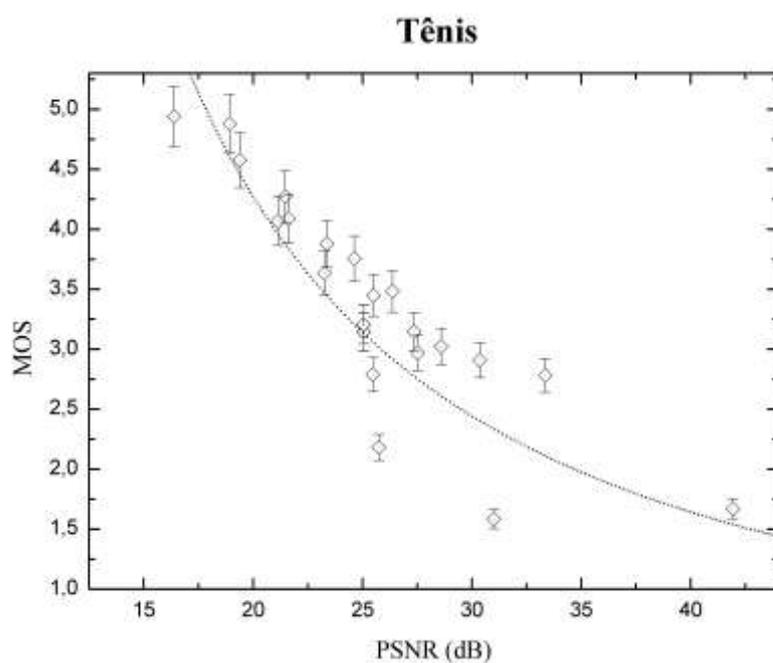
Os gráficos apresentados na Figura 53 representam os dados do PSNR x MOS de todas as degradações estudadas para cada tipo de vídeo utilizado. O intervalo de confiança utilizado foi de 95% para todos os gráficos *plotados* neste capítulo.



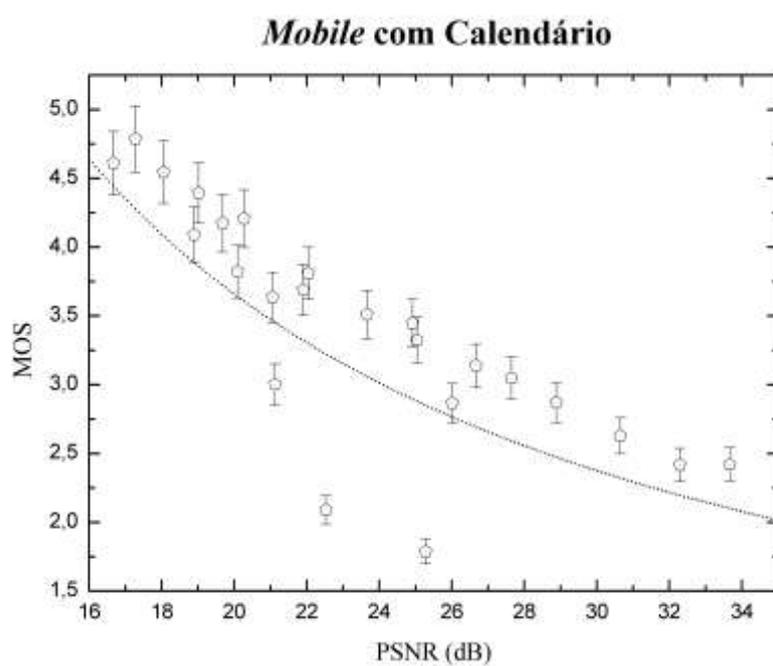
(a) Vídeo Telejornal



(b) Vídeo Guarda-Costeira



(c) vídeo Tenista



(d) Vídeo Mobile com calendário

Figura 53 - Relacionamento entre o MOS e o PSNR com base nos vídeos utilizados no experimento (a, b, c e d).

A tolerância a determinados tipos de degradação também é dependente dos avaliadores, pois alguns não suportam o travamento, no entanto, toleram a blocagem se a taxa for pequena e apresentada em um curto espaço de tempo.

O mesmo tipo de degradação aplicado a vídeos distintos gerou notas diferentes, como pode-se observar nos gráficos apresentados na Figura 53. Logo, conclui-se que o valor do MOS não acompanha o de PSNR quando os vídeos são diferentes.

Os resultados, as entrevistas e a impressão visual dos avaliadores pelo administrador das avaliações, permitiram desenvolver um sistema para gerar melhores ganhos em velocidade na realização das análises dos resultados.

5.2 ANÁLISE DOS RESULTADOS OBTIDOS COM O NOVO AMBIENTE

Nos testes preliminares para especificação do ambiente, realizados neste trabalho, utilizou-se o PSNR e o MSE por serem métricas tradicionais e exaustivamente utilizadas nas pesquisas sobre qualidade de vídeo. Posteriormente, procedeu-se a implementação de avaliações subjetivas para validar os valores do SSIM calculados por meio do sistema proposto para a mesma amostra de vídeos estudados na avaliação subjetiva preliminar.

O SSIM é um método para calcular a similaridade entre duas imagens/vídeos, conforme foi descrito na seção 2.5.2.3. A Figura 54 ilustra os procedimentos de avaliação subjetiva utilizando o ambiente e ferramentas propostas.

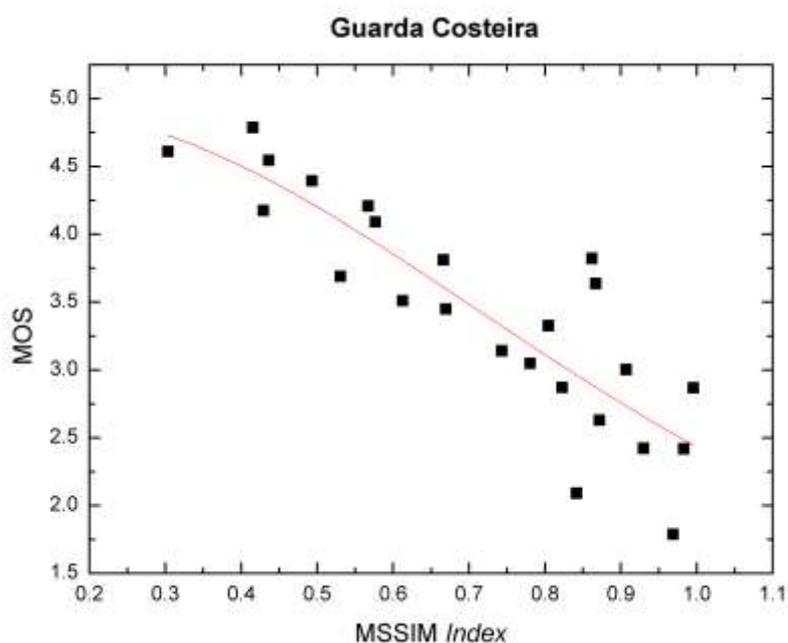


Figura 54 – Avaliações subjetivas utilizando o ambiente de avaliação e suas ferramentas

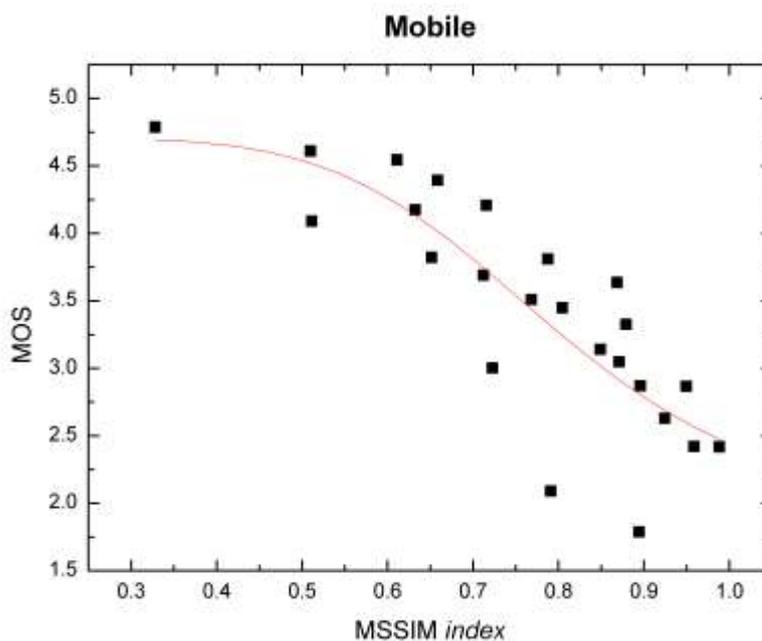
A utilização do ambiente permitiu uma redução no tempo de resposta das avaliações, pois nos testes preliminares com coleta de dados manual, para o tamanho da amostra utilizada, levou-se em torno de 4 horas, desde a tabulação dos resultados até, finalmente, a geração dos gráficos.

A utilização do sistema permite a tabulação e a geração dos gráficos automaticamente. Este ganho reduz o tempo total do processo que envolve ainda a análise dos resultados dos testes efetuados com as métricas subjetivas e a validação das métricas objetivas estudadas. Cuidados foram tomados para atender aos itens previstos na norma (ITU-R, 2002) e atenderam aos critérios estabelecidos pelo VQEG (VQEG, 2008) para os testes realizados além dos trabalhos relacionados utilizados como referência que constam da Seção 3.3, uma vez que se utilizou um painel de LCD para realizar as avaliações.

Os gráficos apresentados na Figura 55 foram obtidos a partir da utilização do ambiente de avaliação subjetiva construído e das ferramentas desenvolvidas para a finalidade de reduzir o tempo da análise dos resultados das avaliações subjetivas.



(a) Gráfico do MOS x MSSIM para o vídeo Guarda-Costeira

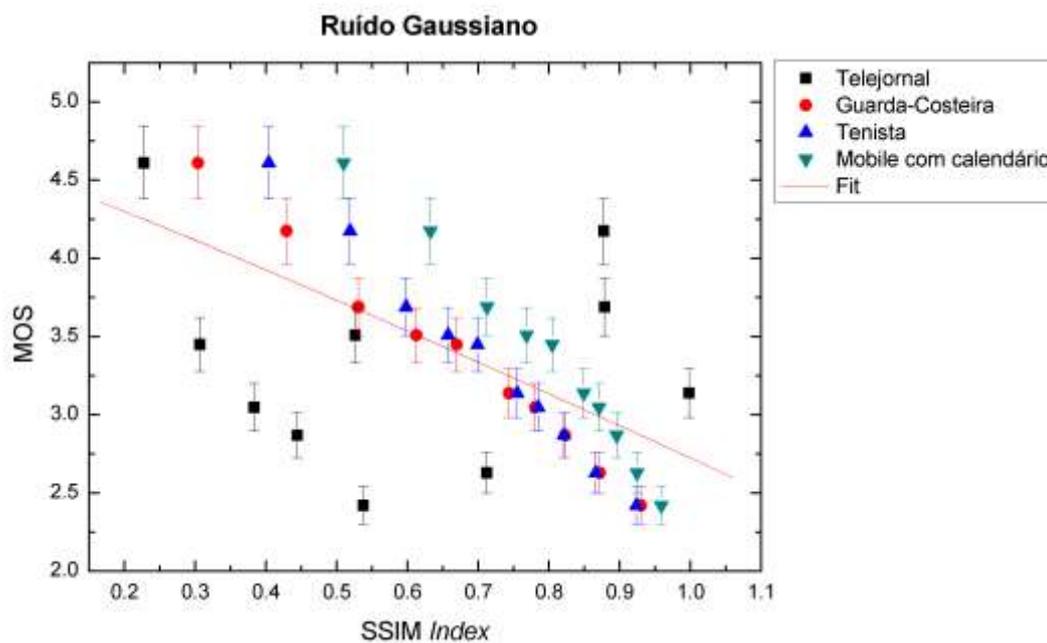


(b) Gráfico do MOS x MSSIM para o vídeo Móbile

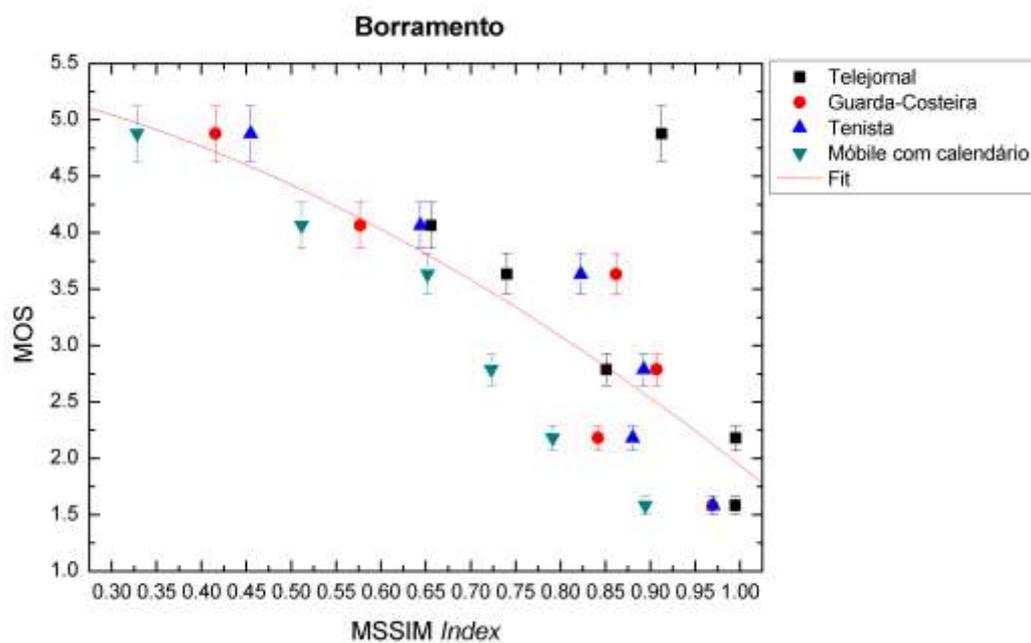
Figura 55 – Gráficos do MOS x MSSIM dos vídeos (a) e (b)

Nos gráficos apresentados na Figura 55 os vídeos são medidos com a métrica objetiva MSSIM e comparados ao MOS dos vídeos avaliados. O MSSIM que é a média dos valores SSIM de cada quadro dos vídeos avaliados. Nesta métrica, quanto mais próximo o valor de 1, melhor a qualidade do vídeo avaliado. A escala utilizada para a comparação com o MOS é a SDSCE, onde os valores variam de 1 a 5, sendo 1 o valor de maior qualidade e 5 o pior resultado em termos de qualidade do vídeo observado.

De forma a manter a relação com o que foi apresentado na avaliação subjetiva preliminar, gerou-se os gráficos apresentados na Figura 56, utilizando-se o MSSIM como métrica objetiva em vez do PSNR com as mesmas degradações apresentadas na Seção 5.1



(a) Ruído Gaussiano



(b) Borramento

Figura 56 - Relacionamento entre o MOS e o MSSIM com base nas degradações para o artefactos de Ruído Gaussiano (a) e Borramento (b)

Observa-se nos gráficos acima que a relação entre o MSSIM e o MOS também varia de acordo com a degradação e depende do tipo de vídeo avaliado, conforme descreve a literatura.

Observa-se que no gráfico 55 (a), o vídeo Telejornal apresenta maior dispersão, isso ocorre devido o vídeo apresentar uma imagem quase estática e o ruído gaussiano se faz notar com maior evidência, já nos demais devido à maior dinâmica dos mesmos, o ruído não causou maior variação das notas atribuídas, ou seja, o coeficiente de correlação é mais baixo para esse vídeo.

Com relação às ferramentas utilizadas para realizar a avaliação subjetiva, os avaliadores se sentiram bastante confortáveis em utilizá-las e foi adotada a metodologia de reutilizar os avaliadores que participaram da primeira avaliação manual para que os mesmos relatassem sua experiência e comparassem com o procedimento anterior. Para isso, foi distribuído um questionário sobre a usabilidade do sistema ao final das avaliações e os resultados dos mesmos encontram-se resumidos na Tabela 4:

Tabela 4 – Resultado dos testes de usabilidade

Questões avaliadas	Excelente	Bom	Regular	Pobre	Ruim
Simplicidade de utilização	70%	30%			
Nível Informações retornadas pelo sistema	35%	30%	30%	5%	
Utilização geral do dispositivo	35%	65%			

Os resultados do teste de usabilidade para os participantes do experimento levaram a melhoria do nível de resposta dada pelo sistema aos avaliadores, de forma que, após os primeiros testes fizeram-se as adequações necessárias para que o nível de informação retornada fosse a mais abrangente possível.

Essas adequações foram realizadas logo após os primeiros 15 avaliadores terem dado seus pareceres, uma vez que esse número atende a norma.

Observa-se que 70% dos avaliadores consideraram o sistema de avaliação como excelente e o restante anotaram como bom.

5.3 RESUMO E CONCLUSÃO DO CAPÍTULO

Neste capítulo apresentaram-se os gráficos gerados no processo de avaliação subjetiva de forma a validar o ambiente e suas ferramentas. São apresentados os gráficos do PSNR x MOS e do MSSIM x MOS gerados na avaliação subjetiva por meio das ferramentas desenvolvidas para o ambiente proposto.

A seguir, encerra-se o trabalho com as conclusões finais e trabalhos futuros.

6 CONCLUSÕES E TRABALHOS FUTUROS

6.1 CONCLUSÕES

A avaliação subjetiva de qualidade de imagem/vídeo é a forma mais efetiva de se avaliar de forma confiável a qualidade, especialmente porque medidas amplamente utilizadas como o MSE, PSNR, não provêm, em geral, boa indicação da qualidade percebida da imagem/vídeo. A Recomendação ITU-R (ITU-R, 2002) é o padrão internacionalmente reconhecido para a avaliação subjetiva da qualidade de imagens/vídeos em televisão.

Com a realização deste trabalho foi possível observar, diante dos resultados obtidos, que a utilização do ambiente de avaliação permite um ganho significativo em termos de tempo para a obtenção dos resultados das avaliações subjetivas, através da geração das estatísticas e gráficos, e seu relacionamento com as métricas objetivas estudadas. Além disso, devido à modularidade empregada na construção do SAASVD, existe a possibilidade de utilização de outras métricas objetivas existentes, bem como de novas métricas que estão sendo desenvolvidas no laboratório.

Conforme se pode observar, os resultados do teste de usabilidade para os participantes do experimento foram excelentes, porém algumas sugestões dadas pelos avaliadores foram levadas a efeito, para melhorar o nível de resposta dada pelo sistema aos avaliadores.

Foi observado que existe uma correlação do MOS e dos valores do PSNR e MSE, obtidos por meio da implementação das respectivas fórmulas que representam tais métricas objetivas, conforme já havia sido validado por (VQEG, 2000; SILVA, 2008), mas que essa correlação é fraca e mesmo sendo as métricas mais utilizadas, muito esforço em pesquisa é dispendido para se implementar métricas com maior correlação com as medidas subjetivas.

Dessa forma, partindo-se de uma experiência já comprovada e validada, pode-se notar que as degradações geradas pelos algoritmos de (ALBINI, 2009) possibilitaram um intervalo de validade para as métricas utilizadas.

Tendo em vista a necessidade de se utilizar uma métrica que melhor representa as características do SVH conforme (WANG et al, 2004), apresentando, assim, melhor correlação das medidas subjetivas e objetivas, implementou-se o SSIM. A partir das medidas do SSIM calculou-se os valores do MSSIM para cada vídeo estudado utilizando-se o sistema automatizado a fim de verificar e validar as especificações do sistema proposto e se o mesmo cumpre os requisitos (Seção 4.1.1 e 4.6, respectivamente) e está de acordo com as normas vigentes no que diz respeito a avaliação da qualidade de vídeo digital (ITU-T, 2008; ITU-R, 2002).

Os estudos experimentais realizados mostram a necessidade de casamento entre o tipo de degradação presente, os objetivos da avaliação e a metodologia empregada (ITU-R, 2002). O foco do trabalho foi no entendimento da norma implementada para a realização de avaliações subjetivas e na extração dos requisitos necessários à criação das ferramentas que permitisse agilizar a geração dos resultados.

Em avaliações subjetivas os elementos relacionados ao projeto de teste, tais como: considerações de visualização, escolha dos avaliadores, considerações de referência, duração dos trechos de vídeo em teste, seleção do faixa de degradações e da escala de avaliação, tem impacto direto sobre o resultado obtido. Escolher as cenas para um teste subjetivo é de fundamental importância para a experimentação. O ponto de partida de um teste é sempre a forma como as cenas serão analisadas. Com base na análise comparativa dos resultados de diversos experimentos observados na literatura corrente, concluiu-se que o modo de analisar as cenas não dependerá apenas dos objetivos do teste, conforme recomendação da norma (ITU-R, 2002), mas também do nível de degradação presente.

Observou-se também a importância dos fatores psicológicos no que diz respeito à influência do contexto e da cena observada. O entendimento e a quantificação destes efeitos são indispensáveis no projeto de melhoria das técnicas de avaliação.

A utilização da avaliação subjetiva traz vantagens uma vez que permite a obtenção de resultados válidos para os sistemas convencionais e sistemas de compressão sobre uma ampla gama de aplicações para cenas paradas e em movimento. Em contrapartida, uma grande variedade de métodos e parâmetros de teste precisa ser considerada, requerendo um arranjo cuidadoso de instalação e controle, diversos avaliadores são envolvidos, o que a torna bastante dispendiosa e complexa. Os custos inerentes à avaliação subjetiva motivaram o desenvolvimento desse conjunto de ferramentas, uma vez que a avaliação subjetiva é considerado parâmetro de qualidade para validar avaliações objetivas segundo VQEG (VQEG, 2008). A utilização de métricas objetivas de qualidade de vídeo envolveu a condução simultânea de testes objetivos e subjetivos para determinar como os parâmetros objetivos podem ser usados para prever as respostas subjetivas dos avaliadores e podem ser usados para motivar a implementação de métricas objetivas mais consistentes e que reflitam os valores das métricas subjetivas.

Assim, o ambiente de avaliação subjetiva de vídeo digital está equipado e pronto para receber os mais variados tipos de avaliação subjetiva previstas nas normas citadas (seção 2.5.1) e devido à arquitetura utilizada para a construção do sistema, o mesmo poderá ser acrescido de novas funcionalidades se assim se fizer necessário.

6.2 TRABALHOS FUTUROS

Como proposta de continuidade deste trabalho, pretende-se investigar as métricas subjetivas com testes comparativos, conforme visto em (PINSON, 2003). Para isso, novos conjuntos de testes estão sendo desenvolvidos. O intuito é melhorar as medidas estatísticas de forma a conseguir atribuir valores para a adequação objetivo-subjetiva, afim de implementar métricas objetivas com maior correlação com as medidas subjetivas.

Foi utilizado para correlacionar os valores obtidos através das métricas objetivas o valor do MOS, no entanto o sistema permitirá o cálculo do DMOS (ITU-R, 2002) para futuras avaliações.

Pretende-se também, integrar ao sistema já implementado, outros módulos, tais como: *codec* customizável para a apresentação de vídeos em vários formatos, geração de artefatos controlada para realizar testes com vídeos segmentados, detecção automática de degradações, rotinas para suavização dos artefatos e sua respectiva avaliação.

Dentre os trabalhos futuros destacam-se a implementação de avaliações em tempo real via *Web*. Além disso, poderá ser implementado um dispositivo semelhante ao descrito neste trabalho com a utilização da tecnologia WPAN (*Wireless Personal Area Network*).

Acrescentar o modelo de avaliação subjetiva proposta pelo ITU (ITU-T, 1999) e do SAMVIQ – *Subjective Assessment Methodology for Video Quality* na ferramenta para possibilitar medições e validações de métricas objetivas.

REFERÊNCIAS

- ALBINI, F. L. P., **Geração de Artefatos em Vídeo Digital**, Dissertação de Mestrado, CPGEI-UTFPR, 2009.
- ALENCAR, M. S., **Televisão Digital**, São Paulo: Érica, 2007.
- AUTRUSSEAU F.; CALLET P. LE. **Subjective quality assessment ivc database**, 2005, <http://www2.irccyn.ec-nantes.fr/ivcdb/>
- BASSO, A.; DALGIC, I.; TOBAGI, F.A.; LAMBRECHT, C.J.V.D.B. **Study of MPEG-2 Coding Performance based on a Perceptual Quality Metric**, Proceedings of PCS'96, pp 263-268, IEEE, 1996.
- BLUEPRINTS. Disponível em: <<http://java.sun.com/reference/blueprints/>>. Acesso em: 03 de março de 2008.
- BOOCH, G., RUMBAUGH, J., JACOBSON, I., **UML – Guia do usuário**, Rio de Janeiro: Campus, 2006.
- BOVIK, A., **Handbook of Image & Video Processing**, Academic Press, Canada. 2000.
- BROWN, A. G. **The Importance of Viewing Environment Conditions in a Reference Display System**, CINEMAQUEST, 2007.
- BRUNNSTRÖM, K.; HANDS, D.; SPERANZA, F.; WEBSTER, A. **VQEG Validation and ITU Standardization of Objective Perceptual Video Quality Metrics**, IEEE SIGNAL PROCESSING MAGAZINE, maio de 2009.
- CHOI, K.; CHOI, J. K.; HONG, J. H.; MIN, G. J.; LEE, J. **Comparison of Video Streaming Quality Measurement Methodologies**, 10TH International Conference on Advanced Communication Technology, - ICACT 2008, vol 2, pp. 993 - 996, fevereiro de 2008.
- CIE. Disponível em <http://www.cie.co.at/index_ie.html>. Acesso em: 05 junho de 2009.
- DE ANGELIS, A.; MOSCHITTAL, A.; RUSSO, F.; CARBONE, P. **Image Quality Assessment: an Overview and Some Methodological Considerations**. IEEE Amuem 2007 –

International Workshop on advanced Methods for Uncertainty Estimation in Measurement, Itália, 2007.

DEITEL, HARVEY M. e DEITEL, PAUL J. **Java como programar**, 6ª edição, São Paulo: Prentice-Hall, 2005.

ECLIPSE. Disponível em <<http://www.eclipse.org>>. Acesso em: 05 março de 2008.

EDEN, A. **No-Reference Image Quality Analysis for Compressed Video Sequences**, IEEE Transactions on Broadcasting, vol. 54, no. 3, setembro de 2008.

ELSMASRI, R E.; NAVATHE, S. **Sistemas de Banco de Dados**, São Paulo: Addison Wesley, 2005.

ESKICIOGLU A.M.; FISHER P.S., **Image Quality Measures and their Performance**, IEEE Transactions on Communications, vol. 43, no.12, pp 2959-2965, Dec. 1995.

FARIA E SILVA, A.E., **Avaliação de Sistemas de Redução de Taxa de Bits de Vídeo com Enfoque na Percepção do Usuário**, Tese de Mestrado, DECOM-FEECUNICAMP, fevereiro de 1998.

GAMMAN, E.; HELM, R.; JOHNSON, R.; VLISSIDES, J. (the GangOfFour). **Design Patterns: Elements of Reusable Object-Oriented**. Addison-Wesley. Professional, 1994.

GLASS, P.; **Development of the visual system and implications for early intervention**. *Infants and Young Children*, 15, 1-10, 2002.

GOLDSTEIN, E. B., **Blackwell Handbook of Sensation & Perception**, Blackwell Publishing: USA, 2005.

HANDS, D.S.; BROTHRTON, M.D.; BOURRET, A.; BAYART, D. **Subjective Quality Assessment for Objective Quality Model Development**, IEEE Electron, vol. 41, pp. 408 - 409, março de 2005.

HIBERNATE. Disponível em <<https://www.hibernate.org/>>. Acesso em: 05 Março de 2008.

HISATAKE, Y.; ITO, H.; KAWATA, Y.; MURAYAMA, A. **Subjective Evaluation Based on Analysis of Correlation between Physical Properties and Visualization of Various Motion Blurs on LCDs**, IEICE Transactions on Electronics, fevereiro de 2007.

HUYNH-THU, Q; GHAMBARI, M. **Scope of validate of PSNR in image/video quality assessment**, IEEE - Electronic letters, vol. 44, no. 13, junho de 2008.

HUYNH-YHU, Q.; GHANBARI, M. **Impact of Jitter and Jerkness on Perceived Video Quality**, in Proc. of the Second International Workshop on Video Processing and Quality Metrics for Consumer Electronics, B. Li, Ed., Jan. 2006.

IEEE 829. **IEEE Std 829: Standard for Software Test Documentation**. New York: IEEE Computer Society, September, 1998.

IEEE e ISO/IEC. **System and Software Engineering - Recommended Practice for Architectural Description of Software-Intensive Systems**. ISO/IEC 42010 IEEE Std 1471-2000 First Edition.

ITU-R BT.710-4, **Subjective Assessment Methods of Quality Images in Systems of High Definition Television**, 1998.

ITU-R BT-500.11, **Methodology for the Subjective Assessment for the Television Pictures**, 2002.

ITU-T P.910, **Telephone Transmission Quality, Telephone, Installations, Local Line Networks**, 2008.

ITU-R Handbook, **Subjective assessment methodology in television**, 1996.

JPA. Disponível em: <<http://java.sun.com/developer/technicalArticles/J2EE/jpa/>>. Acesso em: 10 de março de 2008.

JUMISKO, S. H.; ILVONEN, V. P.; VAANANEN-VAINIO-MATILLA, K. A. **Effect of TV Content in Subjective Assessment of Video Quality on Mobile Devices**, SPIE Proceedings on Multimedia and mobile devices, vol. 5684, pp. 243 - 254, 2005.

KONING, T.C.M.; VELDHOVEN, P.; KNOCHE, H.; KOUIJ, R.E. **Of MOS and men: bridging the gap between objective and subjective quality measurements in mobile TV**,

Proceedings of Multimedia on Mobile Devices 2007, IS&T/SPIE Symposium on Electronic Imaging , January 2007, San Jose, CA, USA.

KONNIK, MIKHAIL V., **Small survey of Objective Image Quality metrics**. Disponível em <<http://www.mvkonnik.info/2009/03/small-survey-of-objective-image-quality.html>>. Acessado em abril de 2009.

KOUMARAS, H.; PLIAKAS, T.; KOURTIS, A. **A Novel Method for Pre-Encoding Video Quality Prediction**. Mobile and Wireless Communications Summit, 2007. , 2007, p.1- 4.

KUO, Wen-Hung; LIN, Po-Hung; HWANG, Sheue-Ling; **A framework of perceptual quality assessment on LCD-TV**, Elsevier CODEN DISPDP, vol 28, no. 1, pp. 35 - 43, 2007.

LAMBRECHT, C. Van Den Branden; **Special Issue on Image and Video Quality Metrics**, Signal Processing, vol. 70, Nov. 1998.

LARMAN C., **Utilizando UML e Padrões – Uma introdução à análise e ao projeto orientados a objeto e ao Processo Unificado**, 2ª Ed., Porto Alegre: Bookman, 2004.

LIU, S.; BOVIK, A. C., **Efficient DCT-Domain Blind Measurement and Reduction of Blocking Artifacts**, em IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 12, Dezembro, 2002 p. 1139-1149.

MOLINARI, Leonardo. **Testes de Software**. São Paulo: Érica, 2003.

MONTEIRO, J.N.; NUNES, M. S. **A Subjective Quality Estimation Tool for the Evaluation of Video Communications System**, IEEE - Symposium on Computers and Communications - ISCC 2007, Portugal, julho de 2007.

MONTEZ, C., BECKER, V., **TV Digital Interativa: Conceitos e Tecnologias**, em Anais do WebMídia & LA-Web, Ribeirão Preto – SP p. 39-77. 2004.

MVC. Disponível em <<http://java.sun.com/blueprints/patterns/MVC.html>>. Acesso em: 15 de março de 2008.

PAPPAS, T.N.; SAFRANEK, R.J. **Perceptual criteria for image quality evaluation**, in Handbook of Image and Video Processing (A. Bovik,ed.), Academic Press, May 2000.

PEDRINI, H.; SCHWARTS, W. R., **Análise de imagens digitais: princípios, algoritmos e aplicações**, São Paulo: Thomson Learning, 2008.

PETROVIC, W. **Subjective Tests for Image Fusion Evaluation and Objective Metric Validation**, Elsevier Publishers, vol. 8, pp. 208 - 216, 2007.

PINSON, M. H.; WOLF, S. **The Impact of Monitor Resolution and Type on Subjective Video Quality Testing**, NTIA Technical Memorandum tm-04-412, U.S. Department of Commerce - National Telecommunications and Information Administration, 2004.

PINSON, M.; WOLF, S. **Comparing Subjective Video Quality Testing Methodologies**, SPIE Visual Communications and Image Processing, vol. 5150, pp. 573 - 582, julho de 2003.

PLANO DE TESTES SUBJETIVOS. Disponível em <<http://www.its.bldrdoc.gov/vqeg/projects/multimedia/index.php>>. Acesso em: 10 de dezembro 2008.

PRESSMAN, Roger S. **“Software Engineering: a practitioner’s approach”**. New York: McGraw Hill, 6th ed., 860p. 2006.

RICHARDSON, I. E. G. **H.264 and MPEG-4 Video Compression**, John Wiley & Sons ltd., England, 2003.

ROUSE, D. M.; HEMAMI, S. S. **Analyzing the Role of Visual Structure in the Recognition of Natural Image Content with Multi-Scale SSIM**, SPIE Proceedings on Human Vision and Electronic Imaging, fevereiro de 2008.

SESHADRINATHAN, K.; BOVIK, A.C., **A Structural Similarity Metric for Video Based on Motion Models**. Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on Volume 1, 2007.

SHANIKA A., KARUNASEKERA; NICK G., **A distortion measure for blocking artifacts in images based on human visual sensivity**, IEEE Transaction on Image Processing, vol. 4, no 6, june 1995.

SHEIKH, H.R.; SABIR, M.F.; BOVIK, A.C. **A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms**, IEEE Transactions on Image Processing, vol. 15, pp. 3440 - 3451, novembro de 2006.

SILVA, E. S. R.; FONSECA, Keiko V. O.; POHL, Alexandre de A. P.; BEZERRA, Sérgio A. C.; SILVA, Wyllian B. da. **Sistema de Auxílio à Avaliação Subjetiva de Vídeo Digital**. WebMedia2009, Simpósio Brasileiro de Sistemas Multimídia e Web, outubro de 2009a.

SILVA, E.S.R.; POHL, Alexandre de A. P.; ALBINI, F.L.P.; FONSECA, Keiko V. O.; **Projeto e Construção de um Ambiente de Auxílio à Avaliação Subjetiva de Vídeo**. SET2009, Sociedade Brasileira de Engenharia de Televisão, agosto de 2009b.

SILVA, E.S.R.; **Especificação do Ambiente de Avaliação de Vídeo Digital**, Relatório Técnico, CPGEI, UTFPR, 2008. Disponível em: <http://lcd.cpgei.ct.utfpr.edu.br/tv/?page_id=12>

_____, **Documentação do Software - Sistema de Automatização de Avaliação Subjetiva de Vídeo Digital (SAASVD)**, Relatório Técnico, CPGEI, UTFPR, 2009c. Disponível em: <http://lcd.cpgei.ct.utfpr.edu.br/tv/?page_id=12>

_____, **Projeto de Hardware - Sistema de Automatização de Avaliação Subjetiva de Vídeo Digital (SAASVD)**, Relatório Técnico, CPGEI, UTFPR, 2009d. Disponível em: <http://lcd.cpgei.ct.utfpr.edu.br/tv/?page_id=12>

SOUK, J. H.; LEE, J. **Recent Picture Quality Enhancement Technology Based on Human Visual Perception in LCD TVs**, Journal of Display Technology, vol. 3, pp. 371 - 376, dezembro de 2007.

SUKUMAR, V.; WARNER, D.; DOHERTY, P.; HESS, H.; NOREN, K; KRONE, S. **Image Quality Assessment - Error Visibility using Gradual Modulation Statistics**, IEEE Workshop on Microelectronics and Electron Devices, pp. 59 - 61, abril de 2007.

TESTE DE ACUIDADE VISUAL. Disponível em: <<http://www.vejam.com.br/baixavisao-teste/>>. Acesso em: 10 de maio de 2008.

TESTE DE ISHIHARA. Disponível em: <<http://colorvisiontesting.com/ishihara.htm>>. Acesso em: 10 de maio de 2008.

TRAÇOS DE VÍDEO. Disponível em: **Video Traces Research Groups**, <<http://trace.eas.asu.edu/index.html>>. Acesso em: 10 de Dezembro de 2008.

VQEG, **Final Report from the Video Quality Experts Group on the validation of objective models of Video Quality Assessment – Phase I**, disponível: <http://www.vqeg.org>, mar. 2000.

VQEG, **Final Report from the Video Quality Experts Group on the validation of objective models of Video Quality Assessment – Phase II**, disponível: <http://www.vqeg.org>, mar. 2003.

VQEG, **Final Report from the Video Quality Experts Group on the Validation of Objective Models of Multimedia Quality Assessment, Phase I ©2008 VQEG**, disponível: <http://www.vqeg.org>, setembro de 2008.

WANDELL, B.A. **Foundations of Vision**, Sinauer Associates, Inc., 1995.

WANG, Z., BOVIK, A. C., SHEIKH, H. R., SIMONCELLI, E. P., **Image Quality Assessment: From Error Visibility to Structural Similarity**, em IEEE Transactions on Image Processing, Vol. 13, No.4, Janeiro 2004 p. 600-612. 2004

WANG, Z.; BOVIK, A. C. **Why is image quality assessment so difficult?**, IEEE - International Conference on Acoustics, Speech, and Signal Processing, 2002.

WATSON, A.B; HU, J.; MCGOWAN, J.F.; MULLIGA, J.B. **Design and Performance os a Digital Video Quality Metric**, Proceedings of SPIE Human Vision and Electronic Imaging. vol. 3644, pp. 168-174, San José, CA, 1999.

WESTERINK, JOYCE H. D. M.; ROUFS, JAQUES A. J., **Subjective image quality as a function of viewing distance, resolution and picture size**, SMTPE Journal, pp. 113-119, fevereiro, 1989.

WINKLER, S., **“Digital Video Quality”**, John Wiley & Sons, Ltda., Chichester, 2005.

WINKLER, S.; CAMPOS, R. **Video Quality for Internet Streaming Applications**, SPIE Proceedings Series - Human vision and electronic imaging, Santa Clara, CA, vol. 5007, pp. 104 - 115, 2003.

WU, H. R., RAO, K. R., **Digital Video Image Quality and Perceptual Coding**, CRC Press, New York, 2006.

XIONG, B., ZHU, C., **A new multiplication-free block matching criterion**, IEEE Transaction on Circuits and Systems for Video Technology, vol. 18, no. 10, outubro de 2008.

YE, S.; SU, K.; XIAO, C. **Video Quality Assessment Based on Edge Structural Similarity**, IEEE Proceedings of the 2008 Congress on Image and Signal Processing, vol. 3, pp. 445 - 448, 2008.

YU, Z.; WU, H.R. **Human Visual System based Objective Digital Video Quality Metrics**, Signal Proceedings, WCCC-ICSP 2000, 5th International Conference, vol. 2, pp. 1088 - 1095, agosto de 2000.

YUAN, F.; HUANG YAN YAO, L. **An Improved PSNR Algorithm for Objective Quality Evaluation**, Control Conference, CCC 2007, pp. 376-380, junho de 2007.

ZAMPOULO R.F.; GOMES D.A; SEARA R., **Avaliação e Comparação de Métricas de Referência Completa na Caracterização de Limiares de Detecção em Imagens**, XXVI Simpósio Brasileiro de Telecomunicações – SBRT08, setembro 2008, Rio de Janeiro, RJ.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)