

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
XX CURSO DE ESPECIALIZAÇÃO EM TELEINFORMÁTICA E REDES DE
COMPUTADORES**

EMILENA AIRI SHIMAKO

Avaliação Subjetiva e Objetiva de Qualidade de Vídeo Digital

MONOGRAFIA DE ESPECIALIZAÇÃO

**CURITIBA
2012**

EMILENA AIRI SHIMAKO

Avaliação Subjetiva e Objetiva de Qualidade de Vídeo Digital

Monografia apresentada como requisito parcial para obtenção do título de Especialista em Teleinformática e Redes de Computadores da Universidade Tecnológica Federal do Paraná, UTFPR.

Orientador: Keiko Verônica Ono Fonseca, Profa. Dra.

**CURITIBA
2012**



TERMO DE APROVAÇÃO

Título da Monografia

Avaliação Subjetiva e Objetiva de Qualidade de Vídeo Digital

por

Emilena Airi Shimako

Esta monografia foi apresentada às 16:00h do dia 28 de fevereiro de 2012 como requisito parcial para a obtenção do título de ESPECIALISTA EM TELEINFORMÁTICA E REDES DE COMPUTADORES, Universidade Tecnológica Federal do Paraná. O candidato foi argüido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado com a nota 10,0 (DEB INTEIROS)

Profa. Dra. Keiko Verônica Ono Fonseca
(UTFPR)

Prof. Dr. Walter Godoy Junior
(UTFPR)

Visto da Coordenação

Prof. Dr. Walter Godoy Júnior
Coordenador do Curso

RESUMO

Esta monografia aborda sobre vídeo digital, sua utilização, demonstrando o método de compressão MPEG (MPEG-1, MPEG-2, MPEG-4 e MPEG-4 part 10), para diminuir a redundância de imagens no vídeo, informando sobre as redundâncias temporal e espacial, como remover essas redundâncias através da transformada discreta de cosseno (DTC) e a aplicação da predição por compensação de movimentos, compressão com e sem perda de dados, subamostragem de cores utilizadas pelo padrão para melhorar a eficiência da codificação e a sequência de quadros com as informações para a codificação e a decodificação (GoP). Descreve o estudo sobre a avaliação subjetiva realizada de acordo com as recomendações estabelecidas pela ITU-R BT.500-12, como tipo e condições de ambiente, número mínimo de pessoas e características para a avaliação, duração de cada sessão e definições sobre as metodologias aplicadas como Double-Stimulus Impairment Scale (DSIS), Double-Stimulus Continuous Quality-Scale (DSCQS), Single-Stimulus (SS), Stimulus-comparison, Single Stimulus Continuous Quality (SSCQE) e Simultaneous Double Stimulus for Continuous Evaluation (SDSCE). Esta pesquisa mostra a avaliação objetiva com referência total, referência reduzida e sem referência, assim como as métricas de qualidade Mean Square Error (MSE), Peak-Signal-to-Noise-Ratio (PSNR), Signal-to-Noise-Ratio (SNR) e Structural SIMilarity (SSIM) e também alguns artefatos de degradação de vídeo como blocagem e borramento.

Palavras chave: Vídeo digital, MPEG, avaliação subjetiva, avaliação objetiva.

SUMÁRIO

1.	INTRODUÇÃO	1
2.	VÍDEO DIGITAL	3
2.1.	MPEG	4
2.1.1	MPEG-1 (ISO/IEC 11172)	10
2.1.2	MPEG-2 (ISO/IEC 13818)	10
2.1.3	MPEG-4 (ISO/IEC 14496)	11
2.1.4	MPEG-4 part 10 (ISO/IEC 14496-10).....	11
3.	AVALIAÇÃO SUBJETIVA DE VÍDEO	14
3.1.	Características comuns.....	14
3.2.	Metodologias para avaliação subjetiva de qualidade de vídeo.....	16
3.2.1.	DSIS (Double-Stimulus Impairment Scale)	16
3.2.2.	DSCQS (Double-Stimulus Continuous Quality-Scale)	18
3.3.	Métodos alternativos de avaliação	20
3.3.1.	SS (Single-Stimulus)	20
3.3.2.	Stimulus-comparison	21
3.3.3.	SSCQE (Single Stimulus Continuous Quality Evaluation)	21
3.3.4.	SDSCE (Simultaneous Double Stimulus for Continuous Evaluation)	24
4.	AVALIAÇÃO OBJETIVA DE VÍDEO	28
4.1.	Métricas de qualidade do vídeo	32
4.1.1.	MSE	32
4.1.2.	PSNR.....	33
4.1.3.	SNR	34
4.1.4.	SSIM.....	34
4.2.	Artefatos de degradação de vídeo	34
4.2.1.	Blocagem (Blocking).....	35
4.2.2.	Borramento (Blurring)	36
5.	CONCLUSÃO	38
6.	REFERÊNCIAS BIBLIOGRÁFICAS.....	39

LISTA DE FIGURAS

Figura 1. Divisão de um Quadro com Várias Partições	13
Figura 2. Variação I	16
Figura 3. Variação II	17
Figura 4. Estrutura de teste do método DSCQS.....	18
Figura 5. Classificação da qualidade da imagem pelo método DSCQS.....	19
Figura 6. Condição de teste	22
Figura 7. Média das notas das sequências de votação sobre o programa segmento Z	23
Figura 8. Exemplo do formato de visualização	25
Figura 9. Diagrama de blocos de referência total	29
Figura 10. Diagrama de blocos de referência reduzida	30
Figura 11. Diagrama de blocos sem referência	31
Figura 12. Imagem com artefato Blocagem	36
Figura 13. Imagem original.....	37
Figura 14. Imagem com artefato Borramento	37

LISTA DE ABREVIATURAS E SIGLAS

ATSC	<i>Advanced Television System Committee</i>
AVC	<i>Advanced Video Coding</i>
CD-ROM	<i>Compact Disc Read-only memory</i>
DSCQS	<i>Double-Stimulus Continuous Quality-Scale</i>
DSIS	<i>Double-Stimulus Impairment Scale</i>
DTC	<i>Discrete Cosine Transform</i>
DVB	<i>Digital Video Broadcasting</i>
DVD	<i>Digital Versatile Disc</i>
ES	<i>Elementary Stream</i>
FFT	<i>Fast Fourier Transform</i>
FR	<i>Full-Reference</i>
GoP	<i>Group of Pictures</i>
HDTV	<i>High-Definition Television</i>
HVS	<i>Human Visual System</i>
IDCT	<i>Inverse Discrete Cosine Transform</i>
IEC	<i>International Electrotechnical Commission</i>
IP	<i>Internet Protocol</i>
IPTV	<i>Internet Protocol Television</i>
ISDB-T	<i>Integrated Services Digital Broadcasting Terrestrial</i>
ISO	<i>International Organization for Standardization</i>
ITU	<i>International Telecommunications Union</i>
JVT	<i>Joint Video Team</i>
LHS	<i>Local Harmonic Strength</i>
MPEG	<i>Moving Picture Experts Group</i>
MSE	<i>Mean Square Error</i>
NR	<i>No-Reference</i>

PSNR	<i>Peak-Signal-to-Noise-Ratio</i>
PVD	<i>Preferred Viewing Distance</i>
QoS	<i>Quality of Service</i>
RR	<i>Reduced-Reference</i>
SDSCE	<i>Simultaneous Double-Stimulus for Continuous Evaluation</i>
SIF	<i>Source Input Format</i>
SNR	<i>Signal-to-Noise-Ratio</i>
SS	<i>Single Stimulus</i>
SSCQE	<i>Single Stimulus Continuous Quality Evaluation</i>
SSIM	<i>Structural SIMilarity</i>
VCEG	<i>Video Coding Experts Group</i>
VHS	<i>Video Home System</i>

1. INTRODUÇÃO

Nos meios de comunicação como *Internet*, TV e celulares, o vídeo digital tem tido um crescimento expressivo. Ele tem sido usado para diversas finalidades, destacando-se o entretenimento, a publicidade e o uso pessoal através de redes sociais. O aumento da utilização de vídeos gera a necessidade da melhora na qualidade de transmissão e da ampliação dos sistemas. [FERGUTZ, 2010].

Um sistema de informação pode ser considerado qualquer tecnologia que gere, manipule, transporte ou armazene a informação, seja ela na forma de textos, números, áudio ou vídeo. Considerando especificamente a área de redes de computadores – onde a informação é transportada – um dos maiores desafios tem sido encontrar mecanismos que integrem de forma eficiente diversas mídias de informação. Serviços multimídia tendem a demandar uma capacidade maior da rede, expressa em termos de uma alta taxa de transferência de dados. Devido ao tráfego multimídia apresentar um grande nível de redundância, técnicas de compressão ou compactação de dados podem minimizar substancialmente o impacto provocado pelo tráfego dessa informação em uma rede qualquer [VENDRAMIN, 2004]. Porém, durante o processo de aquisição, compressão, processamento, transmissão e reprodução, os vídeos distribuídos através de redes de comunicação sofrem vários tipos de distorções. Por exemplo, técnicas de compressão de vídeo com perdas, que são quase sempre usadas para reduzir a largura de banda necessária para armazenar ou transmitir dados de vídeo, podem degradar o sinal durante o processo de quantização. Como outro exemplo, os fluxos de vídeo digital transmitidos através de canais sujeitos a erros, como os canais sem fios, podem ser recebidos de forma imperfeita, devido a danos ocorridos durante a transmissão. Redes de comunicação baseadas em troca de pacotes, tais como a Internet, podem provocar perda ou atraso considerável dos pacotes de dados recebidos, dependendo das condições da rede e da qualidade dos serviços. Todos estes erros de transmissão podem resultar em distorções no vídeo recebido. Como na maioria dessas aplicações o vídeo processado é destinado ao consumo humano, estes irão, em última instância, decidir se a operação foi bem sucedida ou não. Portanto, é essencial para um sistema de serviço de vídeo ser capaz de reconhecer e quantificar a degradação na qualidade

do vídeo que ocorre no sistema, para que ele possa manter, controlar e, eventualmente, melhorar a qualidade deste sinal [ESTRADA, 2011].

Um dos problemas que surge com a introdução de sistemas de compressão de vídeo digital nos mais diversos serviços (distribuições primária e secundária de televisão digital, vídeo sob demanda, videofone, videoconferência, etc.) é o de avaliação de qualidade [ARTHUR, 2002]. Pode-se medir a qualidade do vídeo através do método de avaliação subjetiva ou objetiva.

Este trabalho está organizado da seguinte forma: no capítulo 2 apresenta-se a definição sobre vídeo digital e o método de compressão MPEG. No capítulo 3 demonstra-se a forma de avaliação subjetiva. As métricas subjetivas são obtidas através de avaliações envolvendo seres humanos, que usualmente são instruídos a visualizar uma série de vídeos e atribuir uma nota à cada um de acordo com a sua percepção de qualidade. No capítulo 4 descreve-se a forma de avaliação objetiva de qualidade de vídeo. Os métodos objetivos são aplicados por ferramentas automatizadas, que analisam o vídeo de entrada e o vídeo de referência e resultam em determinados valores, que correspondem à qualidade estimada para o vídeo de entrada. Apesar das dificuldades de criação de uma técnica objetiva que apresente resultados precisos para avaliação de qualidade de vídeo, os resultados da aplicação delas são obtidas de maneira muito mais simples, enquanto a aplicação de metodologias subjetivas normalmente requer mais tempo, esforço e investimento. Porém, se bem aplicadas, as avaliações subjetivas geralmente apresentam resultados confiáveis e precisos [KOZAMERNIK et al., 2005 *apud* DARONCO, 2009]. Por último serão apresentadas a conclusão no capítulo 5 e as referências bibliográficas no capítulo 6.

2. VÍDEO DIGITAL

O vídeo digital, uma sequência de imagens digitais (quadro ou *frames*), surgiu há mais de uma década, com grande parte dos produtos baseados no padrão de vídeo MPEG-2 [IBRAHIM, 1998a *apud* FERGUTZ, 2010]. Em seguida, buscando maior eficiência e melhor compressão, o padrão MPEG-4 ganhou espaço sendo adotado em sistemas de TV digital [CHANGGUI, 1998 *apud* FERGUTZ, 2010]. Vídeos e filmes são compostos de uma sequência temporal de frames que são projetadas a uma taxa adequada (24 fps para filmes e 30 fps para TV) para criar uma ilusão de movimento [THYAGARAJAN, 2011]. A resolução da imagem é diretamente proporcional à quantidade de *pixels* usada em sua representação. Já a sua qualidade está ligada ao número de bits utilizados para representar um *pixel* [VENDRAMIN, 2004].

Algumas vantagens do vídeo digital [CHAGAS, 2011]:

- Possibilidade de Transmissão sem degradação ou com aceitável e até controlável razão sinal-ruído;
- Capacidade de melhoria e flexibilização da resolução na recepção;
- Capacidade de manter a qualidade de reconstrução, reduzindo ao máximo a quantidade de dados armazenados ou transmitidos;
- Capacidade de implementação de multimídia, fazendo a interação da informação do vídeo com dados como tabelas, gráficos, jogos, etc.;
- Capacidade de implementação de técnicas de tratamento do vídeo através de programas computacionais.

A compressão de vídeo permite reduzir a quantidade de dados necessária para representar vídeos digitais, diminuindo os custos de transmissão e armazenamento dos mesmos.

As principais aplicações do vídeo digital são [MANOEL, 2007]:

- Armazenamento e distribuição de filmes.

- Difusão de TV digital aberta e por assinatura.
- Videoconferência e vídeo-mensagem.

Atualmente, estas aplicações utilizam as seguintes tecnologias como meios de armazenamento e comunicação:

- Mídias óticas de acesso aleatório, tais como CD, DVD e os novos formatos HD-DVD e Blue-Ray. Essas mídias são geralmente usadas em equipamentos com hardware dedicado, conectados a um sistema de TV.
- Internet e redes IP (fluxo contínuo e multicasting) – um meio adequado para videoconferências em tempo real, principalmente com redes que possuem garantia de qualidade de serviço (QoS). Também utilizada para distribuição on-line e off-line de filmes e programas de TV, cuja demanda tem aumentado muito com a crescente utilização de conexões domésticas de banda larga.
- Radiodifusão (wireless) – A maior parte da transmissão de vídeo entre satélites como também as transmissões diretas via satélite atualmente são digitais. Em médio prazo, a radiodifusão terrestre de vídeo analógico também deverá ser suplantada pela radiodifusão digital. O meio de telefonia celular, que já provê transmissão de dados digitais, também tem possibilitado a utilização de vídeo digital.

2.1. MPEG

Um vídeo é formado por uma sequência de imagens, chamadas de quadro. Para que o vídeo chegue mais próximo do real, ele deve possuir um grande número de quadros, com isso aumenta o tamanho do vídeo. Para diminuir o tamanho sem a perda muito grande da qualidade são utilizadas técnicas de compressão, que removem as redundâncias do vídeo original. Uma das técnicas de compressão utilizada é a família de padrões de compressão de vídeo digital MPEG.

O MPEG é um acrônimo para *Moving Picture Experts Group* é um padrão que foi estabelecido em 1988 pelo grupo de trabalho ISO (*International Organization for Standardization*)/ IEC(*International Electrotechnical Commission*) [FERGUTZ, 2010], [HANZO; CHERRIMAN; STREIT, 2007], [THYAGARAJAN, 2011]. Os codificadores MPEG apresentam três vantagens: compatibilidade universal, grandes taxas de compressão e perda aceitável de qualidade na imagem final exibida [BER05 *apud* BORTOLETO, 2005].

Para que um sinal de áudio e vídeo possa ser transmitido em uma rede, é necessário que ele seja codificado para otimizar o uso de recursos da rede e minimizar perdas de dados na transmissão. O fluxo de um sinal de vídeo gerado por uma aplicação típica pode ser considerado de grande volume, se comparado aos fluxos de texto e gráficos, por isso, a compressão de dados, especialmente de áudio e vídeo, é necessária para se aperfeiçoar a utilização da largura de banda existente nas redes atuais e limitar a demanda por maior capacidade de armazenamento e transferência. A compressão, que ocorre no codificador do transmissor de vídeo digital, serve para reduzir e remover as informações redundantes ou de menor importância do vídeo original, para facilitar o envio através da rede e armazenamento em disco, mas com isso tem a perda da qualidade visual [BORTOLETO, 2005]. Existe a compressão com e sem perda de dados. Com perda são normalmente conhecidos como quantizadores, pois a informação original é submetida a um processo de quantização, permitindo altas taxas de compressão, ao custo da perda de fidelidade de informação [BRUNO, 2002], também é a mais usada e difundida, já que certas distorções podem ser imperceptíveis ao olho humano, ou mesmo toleradas. Este é o tipo de compressão utilizado pelo padrão MPEG [MANOEL, 2007]. Já sem perda é baseada em técnicas que garantem uma cópia exata do fluxo de dados de entrada depois de um ciclo de compressão/expansão [LYRA et al., 2003], tentando reduzir a redundância da informação, de forma às mensagens ocuparem o menor espaço possível (em bytes), sempre limitada pela entropia [BRUNO, 2002]. Em um sinal de vídeo normalmente possui redundância espacial e redundância temporal. A redundância espacial são as informações redundantes que aparecem em uma mesma imagem, como uma cor de fundo, por exemplo. Já a redundância temporal utiliza a comparação entre dois quadros

consecutivos, o quadro atual e o quadro passado, por exemplo, quando o fundo permanece o mesmo e algum objeto muda de posição na imagem [BORTOLETO, 2005].

O sistema visual humano é mais sensível à variação de brilho (luminância) do que à variação de cor (crominância). Assim, os padrões de compressão de vídeo podem fazer uso dessa característica humana para melhorar a eficiência da codificação reduzindo a quantidade de componentes de crominância em relação aos componentes de luminância [RICHARDSON, 2002 *apud* TIMM et al., 2010], chamada de subamostragem de cores é utilizado pelo padrão MPEG. O padrão utiliza o modelo YUV ou YCbCr, onde Y é uma componente que representa os níveis de cinza ou de luminância, que está relacionada à percepção de brilho da imagem e U e V ou Cr (crominância vermelha) e Cb (crominância azul), componentes que estão associadas à percepção de saturação e matriz das cores, crominância [CHAGAS, 2011].

Existem várias formas de relacionar os componentes de crominância com o componente de luminância para realizar a subamostragem. Os formatos mais comuns são o 4:4:4, o 4:2:2 e o 4:2:0. No formato 4:4:4, para cada quatro amostras de Y, existem quatro amostras de Cb e quatro amostras de Cr. Por isso, os três componentes de cor possuem a mesma resolução e existe uma amostra de cada elemento de cor para cada pixel da imagem e, assim, a subamostragem não é aplicada. No formato 4:2:2, para cada quatro amostras de Y na direção horizontal, existem apenas duas amostras de Cb e duas amostras de Cr. Neste caso, as amostras de crominância possuem a mesma resolução vertical das amostras de luminância, mas possuem metade da resolução horizontal. No formato 4:2:0, para cada quatro amostras de Y, existe apenas uma amostra de Cb e uma amostra de Cr. Neste caso, as amostras de crominância possuem metade da resolução horizontal e metade da resolução vertical do que as amostras de luminância. A nomenclatura 4:2:0 é usada por motivos históricos, pois os números não representam a relação lógica entre os componentes de cor, a qual deveria ser 4:1:1 [RICHARDSON, 2003 *apud* SILVA, A. , 2007].

Para remover a redundância espacial aplica-se a transformada discreta de cosseno (DTC - *Discrete Cosine Transform*), processo matemático de conversão amplitude/frequência, sem perdas e reversível, que converte os dados da amplitude espacial em dados de frequência espacial [SILVA, A., 2008]. A DCT efetua o mapeamento entre a representação de uma imagem bidimensional e a sua representação no domínio da frequência em cada um dos blocos de 8X8 pixels. O bloco resultante da transformação é um bloco também 8X8, mas composto por coeficientes transformados [BORTOLETO, 2005].

A Transformada DCT é um processo reversível (IDCT - *Inverse Discrete Cosine Transform*) que efetua o mapeamento entre a representação de uma imagem bidimensional e a sua representação no domínio da frequência espacial. A DCT não reduz diretamente o número de bits requerido para representar o bloco. A redução no número de bits vem do fato de que a transformação tende a concentrar a energia nos coeficientes de baixa frequência. Os demais coeficientes possuem valor próximo a zero, podendo ser descartados. A distribuição não uniforme dos coeficientes transformados é um resultado da redundância espacial presente no bloco original [CASTRO, F.; CASTRO, M., 2001].

Para remover a redundância temporal aplica-se a predição por compensação de movimento. Este tipo de predição busca compensar movimentos de translação que tenha ocorrido entre o bloco que está sendo codificado e o respectivo bloco no frame de referência que é usado como ponto de partida para predição [CASTRO, F.; CASTRO, M., 2001]. Pela comparação dos macroblocos de um quadro com os de outro, verifica-se se a correlação entre os pixels destes quadros é alta, ou seja, se eles têm conteúdo similar ou idêntico. As coincidências podem ser ignoradas, pois é possível codificar um quadro através do cálculo do vetor de deslocamento em relação ao anterior, reduzindo-se significativamente a quantidade de informação a ser armazenada. Apenas um vetor de movimento é estimado, codificado e transmitido para cada um destes blocos. O erro de estimação, ou seja, a diferença entre um quadro e sua estimação também é transmitido [PER99 *apud* BORTOLETO, 2005]. Outra forma de predição utilizada no padrão MPEG é a denominada predição bidirecional, feita a partir de um quadro prévio (*forward*) e de um quadro

subsequente (*backward*). Esta predição estabelece uma combinação linear destes dois quadros, interpolando os dois deslocamentos [CASTRO, F.; CASTRO, M., 2001].

O principal componente de um MPEG é o GoP (Group of Pictures), sequência de quadros que contêm todas as informações para que esses quadros sejam codificados e depois decodificados. A distância temporal das imagens é o tempo ou o número de imagens entre tipos específicos de imagens em um vídeo digital. "M" é a distância entre sucessivos quadros P e "N" é a distância entre sucessivos quadros I [CHAGAS, 2011]. Dentro do GoP, cada quadro recebe um número que determina o momento em que este deve ser reproduzido. O GoP compreende um quadro I (*Intracoded*), que são imagens completas codificadas individualmente, e seus referentes: P (*Predictive*), quadros codificados com predição relativa ao último quadro, e B (*Bidirectional*), quadros que carregam as diferenças entre o último e o próximo quadro.

Os quadros I são os mais importantes do vídeo. Eles são os únicos completos no fluxo do MPEG e possuem a informação completa, o que os qualifica como pontos de entrada no fluxo, através de acesso randômico. Devem aparecer periodicamente no fluxo de saída já que outros quadros dependem dele. Se apenas um quadro I fosse transmitido, todos os demais quadros do vídeo seriam decodificados com base nele, impossibilitando a decodificação se o mesmo fosse perdido. Operações de avanço rápido e retrocesso também seriam impossíveis.

Na formação dos quadros P, o MPEG utiliza macroblocos. Para compor o quadro atual, o codificador tenta localizar um macrobloco semelhante no quadro anterior, codificando, então, a diferença entre eles. Como apenas as mudanças entre o novo quadro e o quadro de referência precisam ser armazenadas, estes quadros são menores que os quadros do tipo 'I'. Os quadros P são necessários para decodificação de quadros do tipo 'B'.

A composição do quadro B ocorre de forma semelhante à dos quadros P. A única diferença é que o macrobloco de referência pode ser localizado tanto no quadro anterior, como no quadro seguinte. O objetivo desse processo é obter uma

compensação de movimentos. Para a decodificação de um quadro B é necessário o armazenamento de, pelo menos, três quadros: atual, o anterior e o futuro. Como contém muito pouca informação, quadros B nunca são usados como quadro de referência para os demais [BORTOLETO, 2005], [VENDRAMIN, 2004].

Abaixo um resumo das informações sobre alguns padrões [FERGUTZ, 2010]:

- MPEG-1
 - Criado em 1991 para codificar vídeos inteiros;
 - O MP3 é a camada 3 de áudio do MPEG-1;
 - Taxas de bits de 1,5 Mbps;
 - Resolução de 352x240 pixels;
 - 2 horas de gravação gera 1 GB de dados.
- MPEG-2
 - Padrão utilizado em aparelhos de DVD;
 - Taxa de bits variando entre 3 Mbps a 100 Mbps;
 - Resolução de até 1920x1024 pixels.
- MPEG-4
 - Surgiu para ajustar-se melhor à Internet;
 - Qualidade superior com uma taxa de *bits* muito menor;
 - Elementos transmitidos separadamente e combinados para formar quadro do vídeo
- MPEG-4 part 10
 - Conhecido como H.264/AVC;
 - Alta variabilidade;

- 50 kbps a 20 Mbps.

2.1.1 MPEG-1 (ISO/IEC 11172)

Codificação de imagens em movimento e áudio associados à até aproximadamente de 1,5 Mbit / s, é o primeiro padrão MPEG. Ele é compreendido em 5 partes [THYAGARAJAN, 2011]:

- Parte 1: Sistema;
- Parte 2: Vídeo;
- Parte 3: Áudio;
- Parte 4: Teste de conformidade;
- Parte 5: Software de simulação.

O MPEG-1 oferece uma qualidade de imagem semelhante ao VHS (Video Home System). Atualmente ainda é utilizado em armazenamento de vídeo em CD-ROM (Compact Disc Read-only memory).

2.1.2 MPEG-2 (ISO/IEC 13818)

MPEG-2 é um método de codificação, o qual é utilizado de televisão digital como Advanced Television System Committee (ATSC), Digital Video Broadcasting (DVB), ISDB-T [BUENO et al., 2010].

É o padrão utilizado em aparelhos de DVD e possui sinal de qualidade de TV digital de alta definição (HDTV), superando os padrões analógicos.

O padrão incorpora três técnicas de compressão: codificação preditiva, codificação baseada em transformada e codificação de entropia [FERGUTZ, 2010].

2.1.3 MPEG-4 (ISO/IEC 14496)

Este padrão é indicado para projetos que exigem multimídia interativa, especialmente vídeo conferência, distribuição sobre Internet e aplicações similares utilizando pequena largura de banda, com máximo 4 Mbps [CHODOREK, 2002 *apud* FERGUTZ, 2010]. Considerado um padrão de compressão para gráficos e vídeos, foi desenvolvido para transmitir sinais multimídia em redes com pouca largura de banda disponível (como dispositivos móveis, por exemplo). Além do vídeo e do áudio, o MPEG-4 pode combinar texto, gráficos e elementos 2D e 3D, pois usa um conceito de camadas. Uma das funcionalidades mais interessantes diz respeito à escalabilidade, onde um mesmo objeto de vídeo (chamado ES – Elementary Stream) pode ser transmitido em uma camada de qualidade básica acrescida de uma ou mais camadas complementares. Assim, um dispositivo com display de baixa resolução (um celular, por exemplo) receberia apenas a camada básica, enquanto uma TV de alta resolução teria condições de receber todas as camadas do ES [VENDRAMIN, 2004].

Uma das principais características do padrão é oferecer uma maior flexibilidade e reusabilidade na produção de conteúdos, ele proporciona funcionalidades de construção de páginas web e animação gráfica [BARCIK, 2006].

2.1.4 MPEG-4 part 10 (ISO/IEC 14496-10)

Com a redução do custo do poder de processamento e memória, a diversificação do suporte de rede para dados de vídeo codificados, e os avanços na tecnologia de codificação de vídeo, surgiu à necessidade de um padrão para compressão de vídeo com aumento na eficiência da codificação e robustez melhorada para ambientes de rede. Para estes fins os grupos de estudo ITU-T Video Coding Experts Group (VCEG) e ISO/IEC Moving Picture Experts Group (MPEG), em 2001, formaram o Joint Video Team (JVT), grupo de especialistas em codificação de vídeo, para o desenvolvimento de uma nova recomendação [ITU-T, 2005], o MPEG-4 part 10/AVC (Advanced Video Coding), também conhecido como H.264. Esse padrão é atualmente utilizado para armazenamento e transmissão de vídeo digital, vídeo conferência, cinema digital, TV digital e é o sistema utilizado no Brasil.

Desenvolvido em resposta à crescente necessidade de maior compressão de imagens em movimento para várias aplicações, como videoconferências, armazenamento de mídia digital, televisão, internet e comunicação. Também é projetado para permitir a utilização da representação de vídeo codificado de uma forma flexível para uma grande variedade de ambientes de rede. Projetado para ser genérico, no sentido de que serve várias aplicações, taxas de bits, resoluções, qualidades e serviços. As aplicações devem cobrir, entre outras coisas, meios de armazenamento digital, televisão e comunicação em tempo real. No decorrer da criação, vários requisitos de aplicações típicas foram considerados, elementos algoritmos necessários foram desenvolvidos, e estes foram integrados em uma única sintaxe. Assim, esta especificação irá facilitar a troca de dados de vídeo entre diferentes aplicativos [ITU-T, 2005].

Existem muitas aplicações potenciais para codecs H.264/AVC, que vão de celulares à televisão digital e, por isso, a indústria está extremamente ativa nesta área e algumas soluções para HDTV já estão disponíveis, principalmente para decodificadores (que são mais simples) [SILVA, A. M., 2007].

O padrão H.264/AVC é o primeiro que prevê a utilização de macroblocos com tamanhos variáveis (16x16, 16x8, 8x16 ou 8x8 pixels). Uma partição 8x8 é chamada sub-macrobloco, um sub-macrobloco com partição 8x8 pode ser particionado em blocos de 8x8, 8x4, 4x8 ou 4x4 pixels [PURI; CHEN; LUTHRA, 2004 *apud* SILVA, A. M., 2007]. A Figura 1 mostra a divisão do quadro em várias partições.

Para um CODEC H.264/AVC os elementos funcionais básicos (predição, transformação, quantização e codificação de entropia) são sensivelmente diferentes dos demais padrões anteriores. Foram essas significativas mudanças que garantiram os ganhos em eficiência de codificação alcançados pelo padrão H.264/AVC [REGIS, 2009].

O padrão H.264 apresenta um ganho de eficiência de codificação em relação ao MPEG-2 em pelos menos duas vezes, sem aumento significativo no custo final da tecnologia. Por esta razão, o H.264 é o padrão de compressão geralmente utilizado para transmissão para receptores móveis [ALENCAR, 2007 *apud* REGIS, 2009].

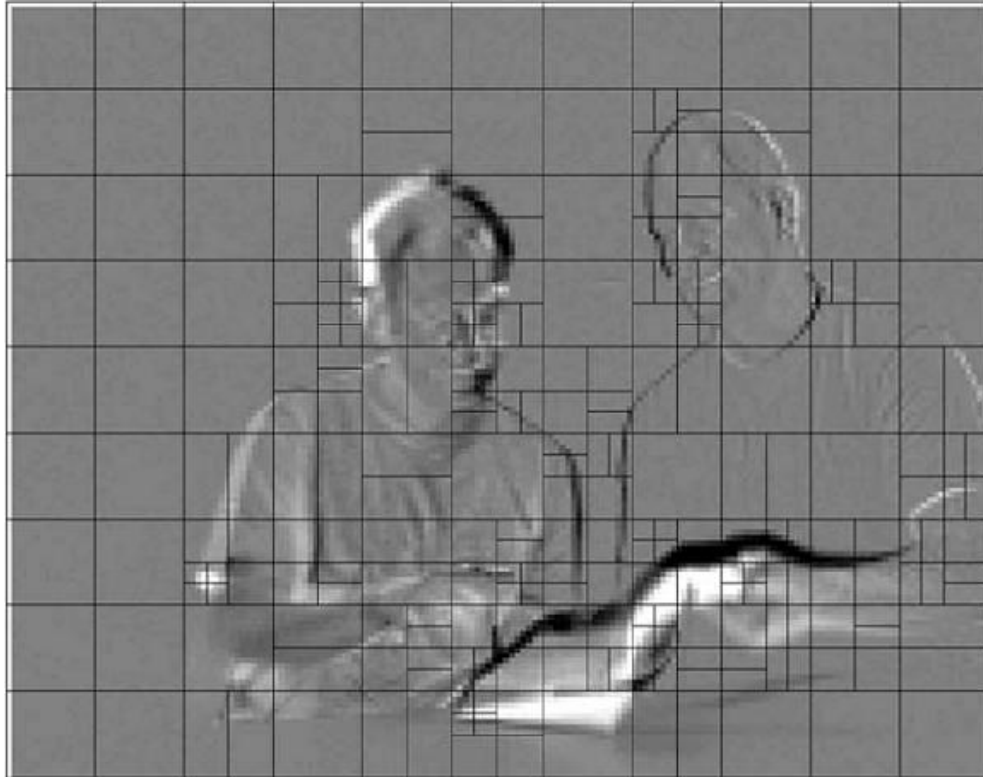


Figura 1. Divisão de um Quadro com Várias Partições [SILVA, A. M., 2007]

3. AVALIAÇÃO SUBJETIVA DE VÍDEO

Os métodos subjetivos de avaliação da qualidade de vídeo são utilizados para estabelecer o desempenho dos sistemas de televisão com medidas que antecipam as reações dos telespectadores dos sistemas testados. Não é possível caracterizar completamente o desempenho por meios objetivos, conseqüentemente, é necessário completar a medidas objetivas com as medições subjetivas.

A avaliação, descrita abaixo, deve estar de acordo com o padrão de avaliação de qualidade estabelecido pela ITU-R BT.500-12 [ITU-R, 2009].

3.1. Características comuns

Diferentes ambientes com diferentes condições de visualização fazem diferença na análise dos resultados. Em um ambiente laboratorial são fornecidas condições mais críticas para verificar o sistema. Já em ambiente residencial fornece um meio de avaliação de qualidade com rede de TV, parâmetros selecionados para definir um ambiente um pouco mais crítico do que as situações de visualização normal. Em ambos os ambientes são testados a iluminação da sala, condições da tela. No caso do ambiente residencial é avaliada a distância de visualização e o tamanho da tela satisfazendo PVD (*Preferred Viewing Distance*), que depende do tamanho da tela.

É essencial a inclusão de material crítico, para que seja possível obter uma interpretação dos resultados, mas não é possível a avaliação a partir de um material não crítico. Nos casos em que conteúdo da cena ou sequência afeta os resultados, o material deve ser escolhido para ser "crítico, mas não excessivamente" para o sistema em teste.

Pelo menos 15 observadores devem participar. São pessoas não especialistas, isto é, não ter como trabalho direto interesse na qualidade da imagem da televisão, não ser avaliador experiente, observadores não especialistas podem produzir resultados mais críticos com a exposição à transmissão de maior qualidade e tecnologias de visualização. Antes da sessão, os observadores selecionados devem ter uma acuidade visual normal e uma visão de cores normal. O número de

observadores necessários depende da sensibilidade e a confiabilidade do procedimento de ensaio adotado e sobre o tamanho esperado do efeito pretendido. Diferentes níveis de habilidade entre os diferentes grupos de avaliadores não especialistas de diferentes laboratórios de teste podem ocasionar uma diferença sistemática entre os resultados obtidos, segundo um estudo de consistência. Porém mais pesquisas precisam ser realizadas para comprovar este resultado. Algumas possíveis características nos painéis de avaliação podem ser adicionadas para facilitar o estudo, como a profissão (por exemplo, estudante universitário, funcionários de uma empresa, etc.), sexo e faixa etária.

Para realizar a avaliação, os avaliadores devem ser cuidadosamente introduzidos ao método de avaliação, aos tipos de deficiência ou fatores de qualidade possíveis de ocorrer, a escala de classificação, a sequência e ao tempo. A sequência que demonstra a variação e o tipo das deficiências a serem avaliados devem utilizar imagens diferentes das utilizadas no teste, mas de sensibilidade comparável. A avaliação de qualidade pode ser definida como para consistir de atributos específicos de percepção.

A sessão de teste deve durar até meia hora. No início da primeira sessão, cerca de cinco apresentações modelo devem ser introduzidos para estabilizar a opinião dos observadores, os dados obtidos nestas apresentações não devem ser consideradas como resultado do teste. Caso tenha várias sessões, cerca de três apresentações modelo são necessárias no início da sessão seguinte. A apresentação deve ser mostrada de forma aleatória.

Para realizar a análise de resultados devem-se obter as informações sobre os detalhes da configuração do teste; detalhes dos materiais de teste; tipo de fonte de imagem e monitores, pois há algumas evidências de que o tamanho do monitor pode influenciar os resultados de avaliações subjetivas; número e tipo de avaliadores; sistemas de referência utilizados; pontuação média geral do experimento; pontuações médias originais e ajustadas e intervalo de confiança de 95%.

3.2. Metodologias para avaliação subjetiva de qualidade de vídeo [ITU-R, 2009]

3.2.1. DSIS (Double-Stimulus Impairment Scale)

A sessão de teste compreende em número de apresentações. Possui duas variações para a estrutura de apresentações. A variação I a imagem de referência ou sequência e a imagem de teste ou sequência são apresentadas apenas uma vez como é mostrado na Figura 2. Na variação II a imagem de referência ou a sequência e teste a imagem ou sequência são apresentadas duas vezes como é mostrado na Figura 3. A variação II, que é mais demorado do que a variação I, pode ser aplicado se a imagem possuir uma distorção pequena, quase imperceptível ou se a sequência em movimento estão em teste.

No início de cada sessão, uma explicação sobre o tipo de avaliação, escala de classificação, sequência e tempo (imagem de referência, cinza, imagem de teste, período de votação) é dada aos observadores. A escala e o tipo de distorção a serem avaliados devem ser ilustrados em imagens diferentes das usadas nos testes, mas de sensibilidade comparável. Nem sempre a pior qualidade vista necessariamente corresponde ao menor grau subjetivo.

É solicitado aos observadores uma avaliação sobre a impressão geral dada pela imagem e definir o desempenho em termos do conteúdo utilizado para definir a escala subjetiva. Os observadores devem olhar para a imagem pelo tempo descrito em T1 e T3. A votação será realizada durante o período T4.

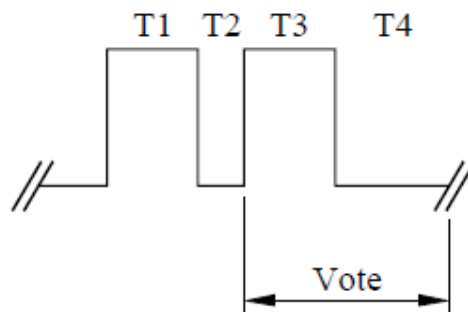


Figura 2. Variação I

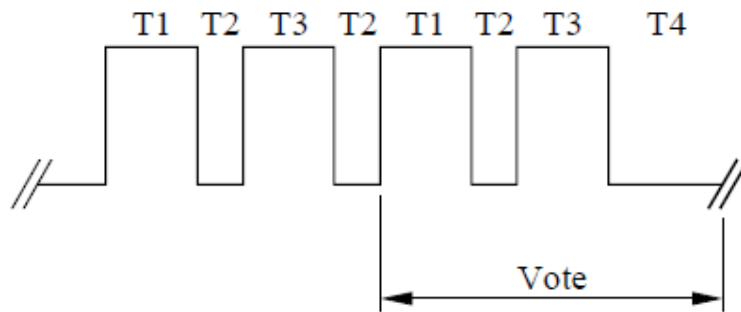


Figura 3. Variação II

Fases de apresentação:

T1 = 10 s Imagem de referência

T2 = 3 s Cinza médio produzido por um nível de vídeo de cerca de 200 mV

T3 = 10 s Condição de testes

T4 = 5-11 s Cinza médio

A experiência sugere que estender o período de T1 e T3 além de 10 segundos não irá melhorar a capacidade dos avaliadores para classificar as imagens ou as sequências.

As imagens de referência e as imagens modificadas são apresentadas em uma sequência randômica e de preferência com a sequência diferente em cada sessão. Em qualquer ocasião a mesma imagem de teste ou sequência não deve ser apresentada em duas ocasiões sucessivas com o mesmo ou diferente nível de alteração. A escala de imperfeições deve ser escolhida pelo grau utilizado pela maioria dos observadores, a pontuação média se caracteriza pela média geral tomada no experimento.

A escala de classificação possui cinco graus de distorção:

- 5 – imperceptíveis;
- 4 - perceptível, mas não irritante;
- 3 - ligeiramente irritante;
- 2 – irritantes;

- 1 - muito irritante.

A sessão não deve durar mais de meia hora, incluindo as explicações e preliminares, a seqüência de testes poderiam começar com algumas fotos indicativas da escala de distorção e a avaliação destas imagens não seriam levadas em conta no resultado final.

3.2.2. DSCQS (Double-Stimulus Continuous Quality-Scale)

O método double-stimulus é especialmente útil quando não é possível fornecer condições de teste de estímulo que mostra toda a variação de qualidade. O avaliador é questionado para ver par de imagens e avaliar a qualidade de ambos, cada um de mesma origem, mas um deles através do processo em análise e outro diretamente da origem. Nas sessões, que duram até meia hora, o avaliador é apresentado a uma série de pares de imagens em ordem randômica e com distorções aleatórias abrangendo todas as necessidades de combinações. No final da sessão, a nota média para cada condição de teste e teste da imagem são calculadas.

O número de repetições depende do comprimento da seqüência de teste. Para imagens estáticas uma seqüência de 3 a 4 segundos com cinco repetições, com votação realizada nas últimas duas imagens, pode ser o apropriado. Já para imagens em movimento uma seqüência de 10 segundos com duas repetições, votação durante o tempo, pode ser mais apropriado. A estrutura será representada na Figura 4.

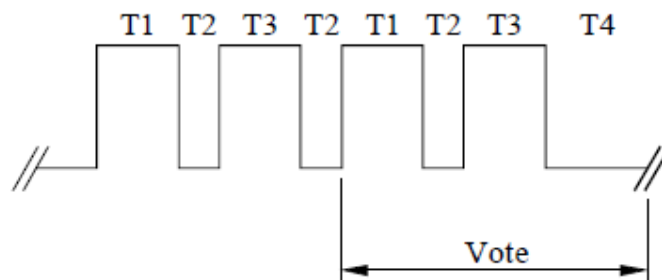


Figura 4. Estrutura de teste do método DSCQS

O método requer a avaliação de duas versões de cada imagem de teste. Um em cada par de testes imagens não possui defeito enquanto a outra apresentação pode ou não conter um defeito. Uma imagem sem deficiência é adicionado para servir de referência sem que seja avisado ao observador. Na série de testes a posição da imagem de referência é alterada de maneira aleatória. Solicita-se aos observadores a realização de uma avaliação geral sobre a qualidade de cada apresentação inserindo a marca em uma escala vertical. A escala vertical é impressa em pares para adaptar as duas apresentações de cada imagem de teste. A escala fornece um sistema de avaliação contínuo para evitar erros de quantização, porém são divididos com os cinco pontos da escala de qualidade como está representada na Figura 5.

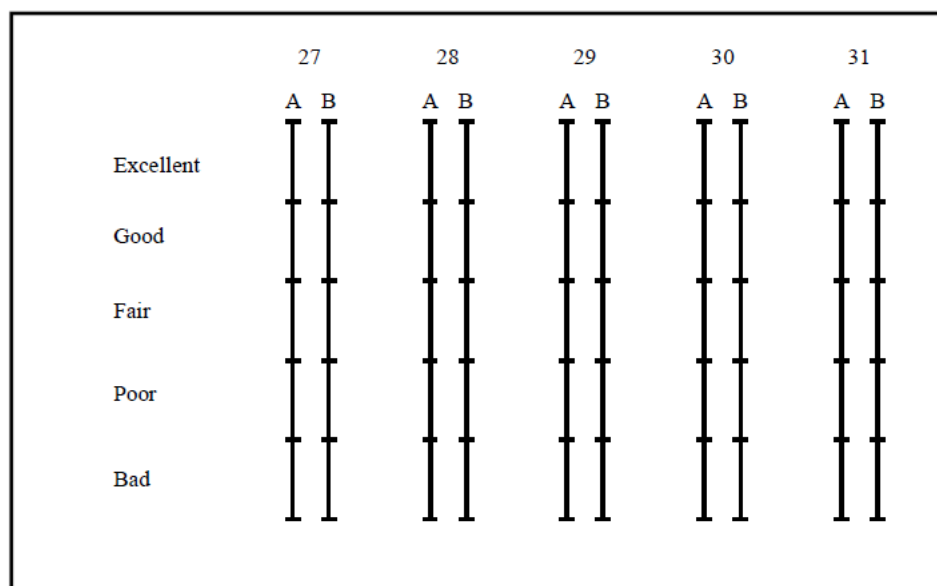


Figura 5. Classificação da qualidade da imagem pelo método DSCQS

A análise dos resultados é realizada com os pares de avaliações (referência e teste). Para cada condição é convertido, a partir da medida de comprimento, na folha de pontuação normalizada no intervalo de 0 a 100. A diferença entre as duas avaliações, de referência e a de teste, são calculadas.

Utilizar o método DSCQS para concluir sobre a qualidade das condições de teste não é muito recomendado associando valores numéricos DSCQS com adjetivos vindos de outro protocolo de teste, como por exemplo, imperceptível, perceptível mas não irritante, entre outros, provenientes do método DSIS. Observa-se que o

resultado obtido através do método DSCQS não deve ser tratado como uma pontuação absoluta, mas com a diferença entre a pontuação entre a condição de referência e a condição de teste.

3.3. Métodos alternativos de avaliação [ITU-R, 2009]

3.3.1. SS (Single-Stimulus)

No método SS, uma única imagem ou uma sequência de imagens é apresentada, e o avaliador fornece um índice de toda apresentação. O material de teste pode incluir somente sequência de testes ou pode incluir as duas sequências de testes e corresponder a sequência de referência. Neste último caso, a sequência de referência é apresentada como um estímulo independente para avaliação como qualquer outro teste de estímulo.

A sessão de teste consiste em uma série de ensaios de avaliação. Estas deverão ser apresentadas em ordem aleatória e, de preferência, em uma sequência aleatória diferente para cada observador. Quando uma única ordem aleatória de sequências é usado há duas variantes para as estruturas de apresentações SS e SSMR (single stimulus with multiple repetition), conforme listado abaixo:

a) As imagens de teste ou sequências são apresentadas apenas uma vez na sessão de teste; no início das primeiras sessões algumas sequências fictícias devem ser introduzidas; experimento normalmente garante que a mesma imagem não seja apresentada duas vezes em sucessão com o mesmo nível de comprometimento.

b) As imagens de teste ou sequências são apresentadas três vezes organizando a sessão de testes em três apresentações, cada um deles incluindo todas as imagens ou sequências a serem testados uma única vez; no início de cada apresentação é anunciada uma mensagem no monitor (Apresentação 1), a primeira apresentação é usado para estabilizar a opinião do observador, o dados emitidos a partir desta apresentação não deve ser utilizadas nos resultados do teste; a pontuação atribuída às imagens ou sequências são obtidos tomando a média dos

dados emitidas a partir da segunda e terceira apresentação, o experimento normalmente garante que as seguintes limitações para a ordem aleatória das imagens ou sequências no interior de cada apresentação sejam aplicadas:

- Uma determinada imagem ou sequência não está localizada na mesma posição em outras apresentações;
- Uma determinada imagem ou sequência não é diretamente localizada antes da mesma imagem ou sequência em outras apresentações.

3.3.2. Stimulus-comparison

No método stimulus-comparison duas imagens ou sequências de imagens são exibidas e o espectador fornece um índice da relação entre as duas apresentações.

O experimento de avaliação irá usar um monitor ou dois monitores coerentes e que geralmente prossegue como em casos de SS. Se um monitor é usado, o experimento envolverá um campo de estímulo adicional idênticas em duração com a primeira. Neste caso, é uma boa prática para garantir que, em ensaios, ambos os membros de um par ocorram igualmente muitas vezes em primeira e segunda posição. Se dois monitores são utilizados, os campos de estímulo são mostrados simultaneamente.

3.3.3. SSCQE (Single Stimulus Continuous Quality Evaluation)

A introdução de compressão de televisão digital produzirá falhas para a qualidade da imagem a qual são dependentes de cena e variáveis no tempo. Mesmo dentro de pequenas partes de vídeo digitalmente codificado, a qualidade pode variar largamente dependendo do conteúdo da cena, e os prejuízos podem ser muito, e de curta duração.

Como resultado, a nova técnica SSCQE tem sido desenvolvido e testado.

Um aparelho de gravação eletrônico conectado a um computador deve ser usado para a gravação da avaliação contínua de qualidade. Este dispositivo deve ter as

seguintes características: Mecanismo deslizante sem qualquer posição danificada escala linear com curso de 10 cm, posição fixa ou sobre a mesa, amostras gravadas duas vezes por segundo. Pelo menos 15 pessoas, não especialistas devem ser ocupados para o teste.

A apresentação de teste representa o desempenho total de um teste, podendo ser dividido em sessões, séries de uma ou mais diferentes combinações de um tipo de programa (duração de pelo menos 5 minutos) processado de acordo com o parâmetro de qualidade sob avaliação sem separação e ordenados em uma ordem pseudoaleatório, que devem durar entre 30 minutos e 60 minutos, para lidar com a máxima duração dos requisitos.

Os dados devem ser coletados a partir de todas as sessões de teste. Um único gráfico de classificação de qualidade média em função do tempo, $q(t)$, pode ser obtida como a média das pontuações de todos os observadores, classificação de qualidade por programa de segmento, parâmetro de qualidade ou pela sessão por inteiro, como mostra o exemplo na Figura 6.

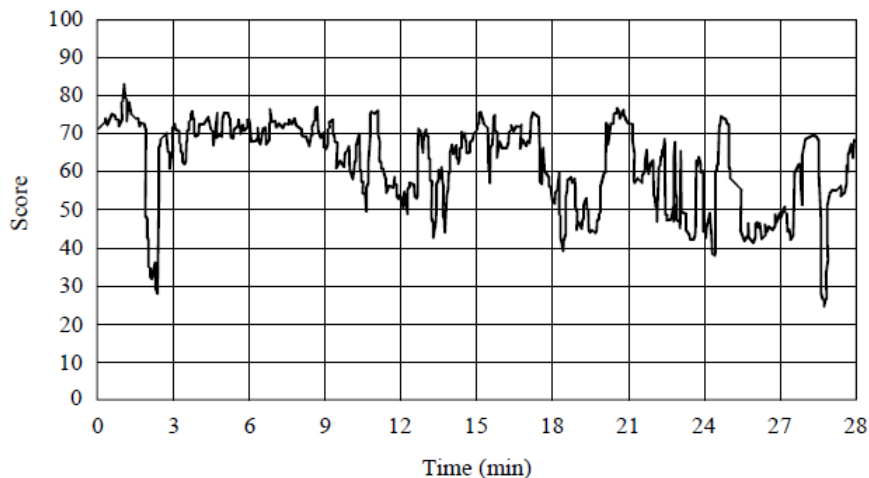


Figura 6. Condição de teste

No entanto, o atraso variando em horário diferente a resposta do observador pode influenciar no resultado da avaliação se apenas a média sobre um programa de segmento é calculado. Estudos estão sendo realizados para avaliar o impacto do

tempo de resposta de diferentes telespectadores sobre o grau de qualidade resultante.

Estes dados podem ser convertidos em um histograma de probabilidade, $P(q)$, da ocorrência do nível de qualidade q , como no exemplo representado na Figura 7.

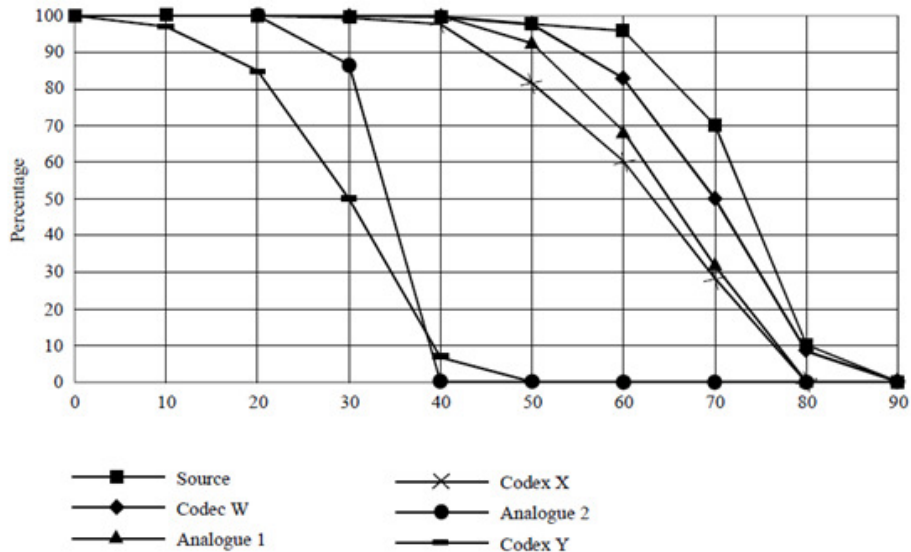


Figura 7. Média das notas das sequências de votação sobre o programa segmento Z

Embora tenha sido demonstrado que a memória baseada em tendência pode existir em DSCQS, recentemente verificou-se que tais efeitos não são significativos nas avaliações DSCQS em trechos de 10 s do vídeo. Conseqüentemente, uma possível segunda etapa no processo SSCQE, atualmente em estudo, pode verificar o histograma de qualidade utilizando o método existente DSCQS com representação de 10 segundos de amostras extraídos dos dados de histograma.

Já foi identificado que o efeito da memória humana pode distorcer a avaliação da qualidade se as falhas visíveis ocorrem no aproximadamente último 10-15 s da sequência. Uma possível terceira fase na metodologia SSCQE seria para processar avaliações de qualidade contínua, a fim de obter uma única equivalente medição da qualidade. Ainda a ser estudada.

3.3.4. SDSCE (Simultaneous Double Stimulus for Continuous Evaluation)

A ideia de uma avaliação contínua chegou a ITU-R porque os métodos anteriores apresentaram algumas inadequações para a medição da qualidade de vídeo de compressão digital. As principais desvantagens dos métodos anteriores padronizados estão vinculadas à ocorrência relacionadas ao contexto de artefatos nas exibições de imagens digitais. Nos protocolos anteriores, a duração do tempo de visualização nas sequências de vídeo sob avaliação é geralmente limitado a 10 segundos que obviamente não é suficiente para o observador ter uma opinião representativa do que poderia acontecer no serviço real. Artefatos digitais são extremamente dependentes do conteúdo espacial e temporal da imagem de origem. Com os métodos de padronização anteriores foi muito difícil escolher representação sequências de vídeo, ou pelo menos avaliar a sua representatividade. Por esta razão ITU-R introduziu o método SSCQE, que é capaz de medir a qualidade de vídeo em sequências mais longas, representante de conteúdos de vídeo e estatísticas de erro. A fim de reproduzir condições de visualização que são o mais próximo possível de situações reais, em que não há referências, são usadas em SSCQE.

Quando a fidelidade tem que ser avaliado, as condições de referência devem ser introduzidas. SDSCE foi desenvolvido a partir da SSCQE, fazendo pequenas modificações sobre a maneira de apresentar as imagens para os observadores e sobre a escala de avaliação. O método foi proposto para que o MPEG avalie a robustez de erro na taxa de bits muito baixa, mas pode ser adequadamente aplicado a todos os casos em que fidelidade da informação visual for afetada pelo tempo- variação da degradação avaliada.

Para a realização do teste um painel de telespectadores observam duas sequências ao mesmo tempo: uma é a referência, o outro um é a condição de teste. Se o formato das sequências é SIF (formato de imagem padrão) ou menor, as duas sequências podem ser exibidas lado a lado no mesmo monitor, caso contrário, dois monitores devem ser utilizados (ver Figura. 8).

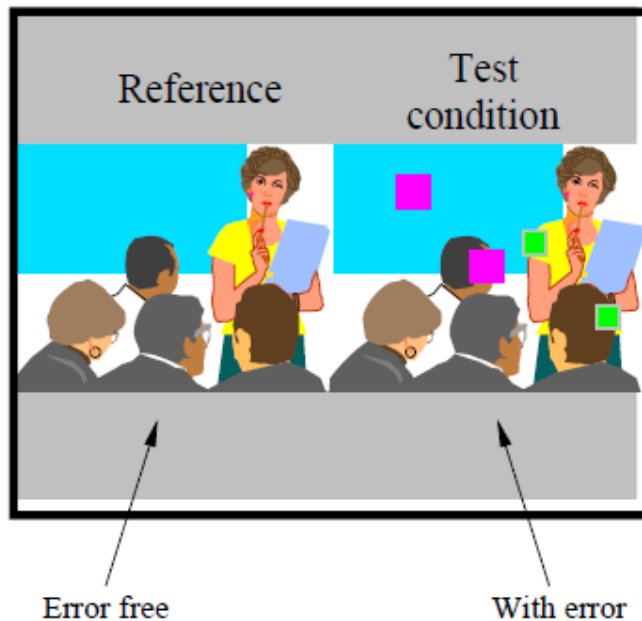


Figura 8. Exemplo do formato de visualização

É solicitado aos telespectadores para verificar as diferenças entre as duas sequências e julgar a fidelidade da informação de vídeo movendo o cursor de um dispositivo de aparelho de voto. Quando a fidelidade é perfeita, o controle deslizante deve estar no topo da faixa de escala (codificadas 100), quando a fidelidade é nula, o deslizante deve estar na parte inferior da escala (codificado 0). Os telespectadores estão cientes de que estão avaliando a imagem de referência enquanto assistem ao ciclo de sequência de imagens.

A fase de treinamento é uma parte crucial deste método de teste, uma vez que os telespectadores podem interpretar mal suas tarefas. Instruções escritas devem ser fornecidas para ter certeza de que todos os telespectadores recebem exatamente as mesmas informações. As instruções devem incluir uma explicação sobre o que os indivíduos estão para ver, o que eles têm de avaliar (ou seja, diferença de qualidade) e como eles expressam sua opinião. Qualquer pergunta do telespectador deve ser respondida a fim de evitar a possibilidade de qualquer opinião do administrador do teste. Após as instruções, uma sessão de demonstração deve ser executada. Desta forma os telespectadores obtêm o conhecimento dos processos de votação e tipo de deficiência.

Finalmente, um teste de simulação deve ser executado, onde uma série de condições representativas são mostradas. As sequências devem ser diferentes das utilizadas no teste e devem ser apresentadas uma após a outra sem qualquer interrupção. Quando o teste de simulação termina, o examinador deve verificar no caso da condição de teste serem iguais ao de referência, as avaliações estão perto de cem (ou seja, não tem diferença a ser observada); se os telespectadores declararem algumas diferenças o examinador deve repetir a explicação e o teste de simulação.

A confiabilidade dos telespectadores pode ser avaliada qualitativamente, verificando seu comportamento quando referência ou pares de referência são mostrados. Nestes casos, os telespectadores são esperados para dar avaliações muito perto de 100. Isto prova que, pelo menos, eles entenderam sua tarefa e eles não estão dando votos aleatórios.

No procedimento SDSCE, a confiabilidade dos votos depende de dois parâmetros:

- Systematic shifts (Mudanças sistemáticas): Durante um teste, um visualizador pode ser muito otimista ou pessimista demais, ou pode até mesmo não ter compreendido os procedimentos de votação (por exemplo, a escala de voto).
- Local inversions (Inversões Locais): Como em outros procedimentos de teste bem conhecidos, os observadores podem às vezes votar sem se preocupar muito em assistir e acompanhar a qualidade da sequência mostrada. Neste caso, a curva de toda votação pode estar relativamente dentro da média. Mas inversões locais podem, no entanto, ser observada.

Estes dois efeitos indesejáveis (comportamento atípico e inversões) poderiam ser evitados. O treinamento dos participantes é muito importante. O uso de uma ferramenta que permite detectar e, se necessário, descartar observadores inconsistentes deve ser utilizado.

A avaliação subjetiva é considerada uma avaliação confiável, pois é realizada por humanos. Mas possui algumas desvantagens, pois demanda de recursos sofisticados, grande número de participantes, sem experiência na área, é demorado, pois demanda um longo período de tempo para a conclusão, habilidade e experiência dos condutores da pesquisa.

4. AVALIAÇÃO OBJETIVA DE VÍDEO

Recentemente vários estudos demonstraram perspectivas no desenvolvimento de algoritmos com capacidade de simular e estimar as medidas subjetivas com grau de certeza cada vez mais elevado. Este é um dos motivos pelo qual são utilizados e desenvolvidos métodos objetivos, para contornar este problema [ALBINI, 2009].

No método de avaliação objetiva não é necessária a interação humana para avaliação, é baseado em modelos matemáticos que automatizam a avaliação do grau de degradação de um vídeo após algum processo degenerativo, tais como codificação (quantização, compressão etc.), transmissão e decodificação, medindo as características da imagem ou vídeo para estimar a média das opiniões dos usuários, aproximando a qualidade observada pelos telespectadores, levando em consideração as características do sistema visual humano HVS (Human Visual System). O uso da métrica resulta em uma economia de tempo e de custo. Os critérios de avaliação devem estar de acordo com os padrões estabelecidos pela ITU e ISO/IEC.

As medidas de avaliação objetiva são mais rápidas e de menor custo do que as soluções subjetivas e são fundamentais na organização de padrões. Mostram degradações ainda imperceptíveis sendo importante para o isolamento de sistemas defeituosos e na formação do conceito de limiares, a partir de que um determinado defeito passa a ser visível, e o quão longe o sistema em teste está deste limiar. Pelas suas características de imparcialidade, confiabilidade, possibilidade de repetição e de baixo custo, métodos objetivos aumentam a competição entre provedores de serviço e equipamentos, e facilitam a especificação e avaliação de novos sistemas. A realização de medidas objetivas em tempo real abre a perspectiva de monitoração contínua da qualidade do vídeo [ARTHUR, 2002].

A avaliação objetiva pode ser classificada de acordo com a disponibilidade do vídeo original que é considerado livre de distorção ou de alta qualidade [BEGAZO; RODRÍGUEZ; RAMÍREZ, 2011], em função da quantidade de informação necessária para o cálculo da qualidade do vídeo. Pode ser classificada em três categorias:

- Referência total (FR - Full-Reference), métricas que se baseiam quando ambos os sinais, do vídeo original e do vídeo processado, estão disponíveis para avaliação, envolvendo a avaliação dos algoritmos de compressão e de transmissão [ALBINI, 2009].

Em geral, esta métrica é a que tem o melhor desempenho e permite o uso de uma quantidade ilimitada de referência, pois possui a disponibilidade do vídeo de origem durante todo o processo. As medidas são realizadas sobre comparações pixel-a-pixel entre imagens original e degradada. Por isso, o objetivo deste parâmetro é reunir a maior quantidade de informação possível e útil, e que possam repercutir em um parâmetro de qualidade robusto e eficaz. [ARTHUR, 2002] [ESTRADA, 2011]. Este tipo de métrica possui geralmente um grau de complexidade baixo e uma implementação mais rápida [ALBINI, 2009]. A Figura 9 representa o diagrama de blocos da referência total.

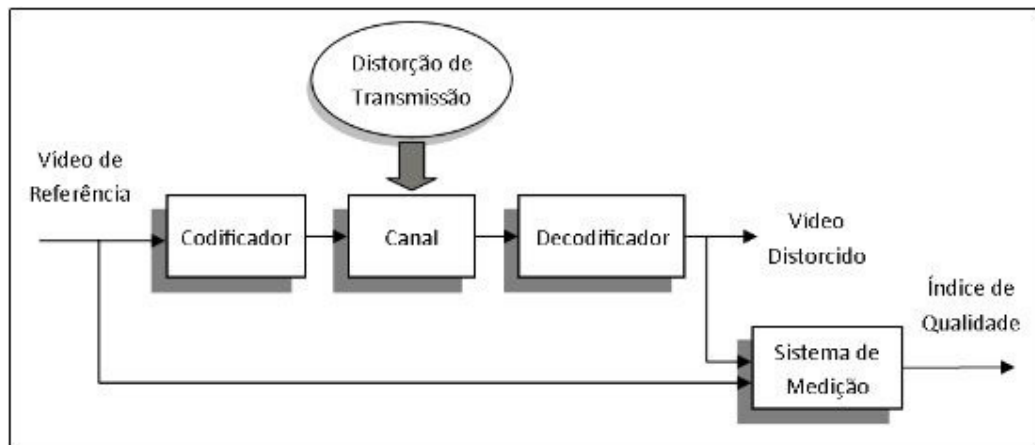


Figura 9. Diagrama de blocos de referência total [ESTRADA, 2011]

- Referência Reduzida (RR - Reduced Reference) utiliza-se de apenas algumas amostras ou certas características do vídeo original, uma quantidade limitada de informação, para efetuar a avaliação e são aplicadas, em geral, da mesma maneira que as métricas com referência total.

Uma das características interessantes desta métrica é a possibilidade de escolher a quantidade de informação secundária. Métricas desta classe podem

ser menos precisas do que as métricas de referência total, mas elas também são menos complexas, e fazem as implementações em tempo real mais acessíveis. No entanto, a sincronização entre os dados originais e distorcidos ainda é necessária [ESTRADA, 2011].

Quando o vídeo é transmitido, junto com ele são enviadas informações importantes sobre a sua imagem original, tais como, perda harmônica e ganho harmônico. Estes parâmetros são calculados utilizando-se a LHS – Local Harmonic Strength, que é uma técnica de referência reduzida, que usa apenas uma ferramenta para avaliar várias distorções. A perda harmônica e o ganho harmônico são informações adquiridas com base nas bordas detectadas na imagem, e com os coeficientes da Transformada rápida de Fourier (FFT – Fast Fourier Transform). Estes valores são calculados tanto para o vídeo original, como para o degradado e então comparados para avaliar a qualidade da imagem [GUNAWAN, e GHANBARI, 2007 apud ALBINI, 2009]. A Figura 10 representa o diagrama de blocos da referência reduzida.

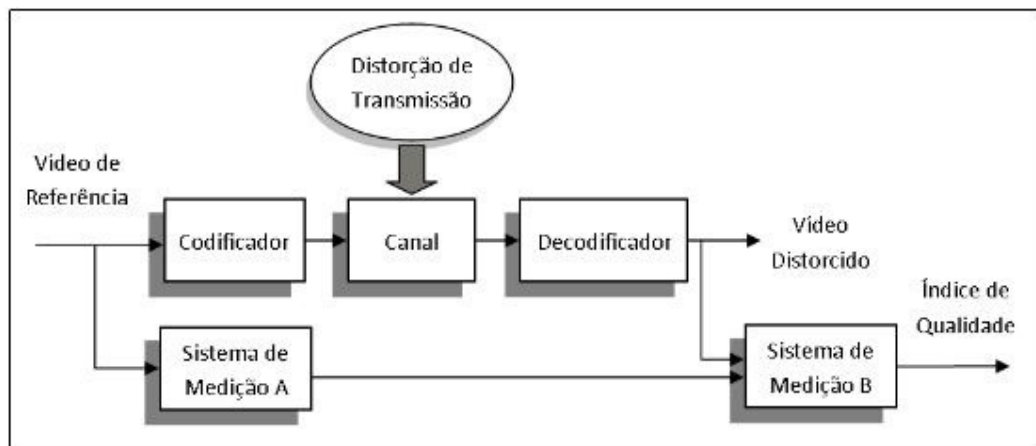


Figura 10. Diagrama de blocos de referência reduzida [ESTRADA, 2011]

- Sem Referência (NR - No-Reference) quando apenas o sinal processado, vídeo a ser avaliado está disponível, sem nenhuma informação do vídeo original, sendo que este tipo de métrica é o desejado para avaliação objetiva da qualidade de vídeo, embora seja muito difícil de ser implementado, pois para isso é

necessário simular inúmeras características do sistema visual humano, das quais muitas ainda são incompreendidas [ALBINI, 2009].

As métricas sem referência são métricas propostas para tentar prever a qualidade do vídeo com base apenas na imagem final recebida, ou seja, da mesma forma com que o telespectador faz a sua avaliação no método subjetivo. Este tipo de métrica seria a ideal para ser utilizada na medida da qualidade do vídeo. Estas, porém, são métricas muito complexas e difíceis de implementar e, talvez, o maior problema na área de avaliação de qualidade de imagem [WANG, BOVIK, 2006 apud ALBINI, 2009]. Ainda não se conhecem métricas sem referência que tenham sido totalmente eficientes e que possam ser utilizadas comercialmente para simular efetivamente o HVS dos espectadores [KOURMARAS, PLIAKAS e KOURTIS, 2007 apud ALBINI, 2009], [RICHARDSON, 2003 apud ALBINI, 2009]. Apesar de observadores humanos geralmente poderem avaliar a qualidade de um vídeo sem usar uma referência, criar uma métrica sem referência é uma tarefa muito difícil [ESTRADA, 2011]. A Figura 11 representa o diagrama de blocos sem referência.

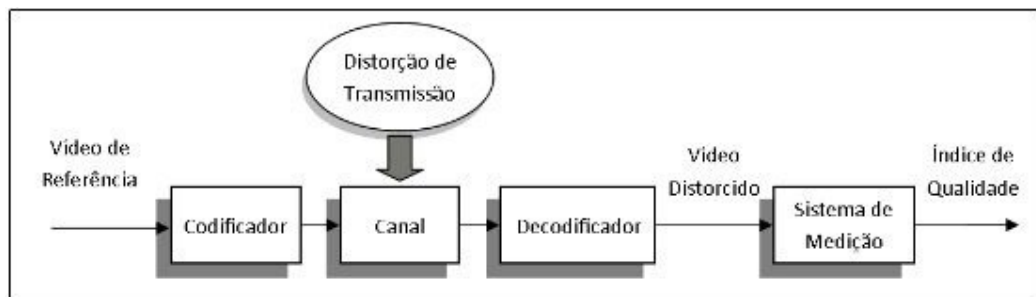


Figura 11. Diagrama de blocos sem referência [ESTRADA, 2011]

As métricas RR e NR são direcionadas a aplicações em tempo real, onde as limitações de complexidade computacional e da falta de acesso à referência são as principais restrições.

As métricas objetivas com referência reduzida ou com referência total possuem, porém, uma deficiência. Supondo o caso hipotético no qual o vídeo original possui uma qualidade muito baixa e o meio de transmissão é ideal, ou seja, não apresenta interferências, pode-se afirmar que o vídeo recebido é igual ao transmitido. Ao ser

avaliado por uma métrica com referência total ou reduzida, este vídeo recebido (o qual é igual ao original), será classificado como um vídeo perfeito, de qualidade muito boa, quando realmente não o é. Logicamente, se uma avaliação subjetiva fosse realizada, o seu resultado seria completamente divergente da avaliação objetiva. Isto se deve ao fato da referência utilizada ser um vídeo com uma qualidade baixa. Caso um telespectador fosse indagado quanto à qualidade do vídeo original, a sua opinião seria a mesma relativa ao vídeo recebido (pois estes são iguais). Com isso pode-se afirmar que as métricas com referência total ou reduzida são aplicáveis na predição da diferença entre a qualidade do vídeo original e do vídeo recebido (degradado), elas são úteis na avaliação do impacto gerado pelo meio de transmissão e/ou pelos algoritmos de compressão na qualidade da imagem, porém não funcionam sempre como uma boa projeção para a avaliação subjetiva que é realizada apenas com o vídeo recebido [ALBINI, 2009].

4.1. Métricas de qualidade do vídeo

A partir da imagem degradada, identifica-se visualmente o artefato que afeta a qualidade da imagem, o qual pode ser também, detectado objetivamente. As métricas MSE, PSNR e SNR, fazem medições em tempo real, comparando à imagem sob teste como a imagem original. Essas medidas mostram uma variação em relação à qualidade segundo observadores humanos, por consistirem de uma comparação matemática entre o sinal original e o sinal processado [REGIS, 2009].

4.1.1. MSE

MSE (Mean Square Error) erro quadrático médio é calculado através da soma das diferenças quadráticas entre todos os pixels da imagem original e da imagem que está sendo avaliada. Este resultado é dividido pelo número de *pixels* totais (para fazer a média), observando o nível de luminância de cada pixel. O MSE é bastante popular, apresenta expressões matemáticas simples, facilitando a manipulação analítica. Mas testes mostram que há pouca correlação de medidas baseadas no MSE com as observações subjetivas [WANG et al., 2004 *apud* REGIS, 2009], pois a percepção humana das distorções de imagem/vídeo e artefatos não é levada em

consideração. Quanto menor o valor do MSE mais próxima da imagem original a imagem avaliada será. O MSE sempre possui valor positivo e o seu valor mínimo é zero. Quando o valor do MSE é zero isso significa que as duas imagens são iguais.

Esta métrica possui a vantagem de ser facilmente implementada. Seu cálculo não exige muita capacidade de processamento e é extremamente rápido. Por outro lado, sua desvantagem reflete no fato de seu valor nem sempre se correlacionar adequadamente com as medidas de qualidade realizadas subjetivamente. [ALBINI, 2009].

4.1.2. PSNR

A métrica PSNR (Peak-Signal-to-Noise-Ratio - Relação Sinal Ruído de Pico) define a relação entre a máxima energia possível de um sinal e o ruído que afeta a representação do sinal [BEGAZO; RODRÍGUEZ; RAMÍREZ, 2011], pode ser calculado fácil e rapidamente e por isso é uma métrica de qualidade bastante popular, largamente utilizada para comparar a qualidade de imagens de vídeo comprimidas e descomprimidas. Normalmente é utilizada para medir a qualidade da reconstrução da imagem ou vídeo após uma compressão com perdas [PINHEIRO, 2010]. A medida do PSNR apresenta várias limitações. Para comparação ela necessita da imagem original no mesmo instante de tempo. Porém isso pode não ocorrer durante todo o espaço de tempo da medida. Outro problema, diz respeito à imagem “original”, que pode não apresentar uma boa qualidade. Para uma imagem dada ou uma sequência de imagens, um PSNR alto geralmente indica uma alta qualidade e um PSNR baixo indica baixa qualidade. Porém, um valor de PSNR particular não necessariamente se compara com um valor “absoluto” da qualidade subjetiva. Devido às limitações de métricas como o PSNR, vários trabalhos foram realizados nos últimos anos para desenvolver um teste objetivo mais sofisticado que se aproxime mais dos resultados subjetivos [ALBINI, 2009].

O principal problema desta medida é não levar em consideração as limitações do sistema visual humano (HVS). Os algoritmos de compressão de imagens e os algoritmos de compressão de vídeo se utilizam destas limitações para atuar de forma eficiente na compressão destas imagens e vídeos [FONSECA, 2008].

4.1.3. SNR

A SNR (Signal-to-Noise-Ratio) é a relação sinal-ruído que contabiliza o quanto um sinal foi corrompido, através do cálculo da energia da diferença entre a imagem de referência e a imagem de teste [PINHEIRO, 2010].

4.1.4. SSIM

O SSIM (Structural SIMilarity) índice de similaridade estrutural, é um algoritmo que utiliza a estatística da imagem para a avaliação de qualidade de imagens usando janelas, subdividindo as imagens sob comparação em blocos menores obtendo o valor do SSIM para cada par de blocos usando o valor médio. Os índices SSIM são calculados pelas janelas deslizantes, que se movem pixel-a-pixel do topo esquerdo até o canto inferior direito do quadro [REGIS, 2009].

A métrica SSIM assume que um sinal de vídeo/imagem é fortemente estruturado, cuja percepção da visão humana é altamente adaptada para extração de informações estruturais de uma cena. Assim, uma medida de similaridade estrutural (ou distorção) pode prover boa aproximação para a qualidade perceptual da imagem [Shi et al. 2009; Wang e Bovik 2002; 2006; Wang et al. 2004 *apud* SILVA, W. et al., 2009].

4.2. Artefatos de degradação de vídeo

No processo de compressão a degradação do vídeo pode se tornar perceptíveis e incomodar o usuário que pode ser identificado visualmente o artefato que afeta a qualidade da imagem. Alguns exemplos de artefatos são: blocagem (tiling ou blocking), borramento (blurring), distorção localizada (áreas pouco nítidas da imagem) [PLATINUM, 2009 *apud* SILVA et al, 2009], estas degradações interferem diretamente na qualidade do vídeo, alterando os valores de luminância ou crominância dos pixels e suas posições no quadro [SILVA et al, 2009]. Os sistemas de vídeo digital produzem degradações que são fundamentalmente diferentes das produzidas pelos sistemas analógicos. Estas degradações podem ser introduzidas através de filtragem, conversão, transformação, quantização e/ou transmissão.

Perda de resolução, efeito de bloco, ruído de quantização e erros de bloco são alguns dos indesejados efeitos do processo de compressão. Embora o vídeo comprimido seja geralmente degradado em relação ao original, alguns processos podem reduzir o nível de degradação existente no material. Isto significa, portanto, que para alguns tipos de degradação a codificação pode significar uma melhora na imagem. É o caso em que o sinal original já está degradado e pode acontecer, por exemplo, do material comprimido ser menos ruidoso do que o sinal original [ARTHUR, 2002].

Ao observar a transmissão de sinais de televisão analógica e digital, é fácil identificar o aparecimento de degradações, também chamadas de artefatos. Chuviscos e fantasmas nas transmissões analógicas [DTV, 2008 *apud* ALBINI, 2009].

4.2.1. Blocagem (Blocking)

Degradação do tipo blocagem ou efeito de bloco é a deterioração que a imagem apresenta na forma de padrões retangulares, que ocorre com a mudança repentina de intensidade regiões de fronteira entre blocos adjacentes, que não estavam presentes na imagem original, como pode ser observada na Figura 12. As discontinuidades se tornam visíveis quando a quantização é pouco refinada. Este tipo de degradação está presente em cenas complexas ou quando elevadas taxas de compressão são necessárias. Nestas situações, o “buffer” do codificador tende a encher e a quantização é ajustada para ser menos refinada e evitar o transbordo do “buffer”. Para uma dada granulidade e quantização, a blocagem é mais perceptível em áreas mais planas da figura [ARTHUR, 2002].



Figura 12. Imagem com artefato Blocagem [WU e RAO, 2006 apud ALBINI, 2009]

4.2.2. Borramento (Blurring)

É a redução da definição em bordas e áreas com muitos detalhes espaciais. Em cenas de alta resolução espacial este efeito é consequência do compromisso na alocação de bits para a descrição dos detalhes de alta resolução e na descrição do movimento. É a etapa de pré-filtragem a maior responsável pela perda de resolução, onde ocorre a interferência entre os *pixels*. Se a imagem for convertida de um formato com menor número de linhas para um formato com mais linhas, então, ela aparecerá borrada, antes mesmo da compressão. A Figura 13 representa a imagem original e na Figura 14 está representada a imagem com o artefato de borramento.

Perda de resolução ou erros de cromaticidade em vídeo digital comprimido pode ser mais perceptível que os demais tipos de degradação, pois a maioria dos sistemas emprega uma subamostragem na cromaticidade. Na presença de objetos coloridos em movimento rápido, a subamostragem de cores pode levar a um desagrupamento entre luminância e cromaticidade. Essa perda de resolução torna as outras deteriorações mais visíveis [ARTHUR, 2002].



Figura 13. Imagem original [ARTHUR, 2002]



Figura 14. Imagem com artefato Borramento [ARTHUR, 2002]

5. CONCLUSÃO

Esta monografia apresentou os métodos que avaliam a qualidade de vídeo digital transmitido, como nas redes sociais, TV digital, esteja cada vez melhor. Como normalmente é utilizado o método de compressão de vídeo MPEG que remove as redundâncias de um vídeo, utilizando a transformada discreta de cosseno e predição por compensação de movimentos, algumas informações são perdidas fazendo com que diminua a qualidade. Para isso testes são realizados para uma melhora na qualidade, os métodos subjetivos e objetivos são aplicados para a melhoria do produto final transmitido.

O método subjetivo de avaliação de qualidade é realizado através de opiniões dadas através de interação humana ao qual são escolhidas através de características estabelecidas pela recomendação ITU-R BT.500-12. Existem várias metodologias para a avaliação, cada uma com aplicação de avaliação diferente. Este método pode ser considerado o melhor para a avaliação de qualidade, o problema é que demanda muito tempo para a realização de testes e o gasto também pode ser grande, pois além das pessoas, o ambiente e a preparação das imagens de teste são importantes para que a avaliação seja mais próxima do ideal.

Já o método objetivo é realizado através de modelos matemáticos que ajudam na avaliação da degradação, que podem ser imperceptíveis na imagem, comparando matematicamente imagens de teste com a imagem original. Diferentemente do método de avaliação subjetivo ele não demanda tempo e nem muito gasto.

O ideal seria uma avaliação que utiliza tanto o método subjetivo quanto o método objetivo.

A dificuldade em realizar esta monografia, foi a falta de conhecimento sobre o assunto, o que demandou mais tempo para a pesquisa. Porém a quantidade de materiais disponíveis ajudou na elaboração.

6. REFERÊNCIAS BIBLIOGRÁFICAS

[ALBINI, 2009] ALBINI, F. L. P. **Geração e Avaliação de Artefatos em Vídeo Digital**. Dissertação de Mestrado - Universidade Tecnológica Federal do Paraná, UTFPR, 2009.

[ANDRADE; CORDEBELLO, 2010] ANDRADE, L. A.; CORDEBELLO, P. D. **Estudos de Codificação de Vídeo Anaglífico em MPEG2 e MPEG4 para Distribuição em Massa de Conteúdos Estereoscópicos**. Revista de Radiodifusão - SET, volume 04, número 04, 2010.

[ARTHUR, 2002] ARTHUR, R. **Avaliação Objetiva de Codecs de Vídeo**. Dissertação de Mestrado – Universidade Estadual de Campinas, UNICAMP, 2002.

[ASCENSO, 2003] ASCENSO, J. M. D. **Codificação Escalável de Vídeo com Elevada Granularidade**. Dissertação de Mestrado - Universidade Técnica de Lisboa - Instituto Superior Técnico, 2003.

[BARCIK, 2006] BARCIK, L. A. **Sistemas de televisão digital e televisão sobre protocolo IP**. Monografia de Especialização - Universidade Tecnológica Federal do Paraná, UTFPR, 2006.

[BEGAZO; RODRÍGUEZ; RAMÍREZ, 2011] BEGAZO D. C; RODRÍGUEZ D. Z; RAMÍREZ M. A. **Avaliação de qualidade de vídeo sobre uma rede IP usando métricas objetivas**. In: Conferência Ibero-americana em Sistema, Cibernética e Informática CISCI 2011, 2011, Orlando, USA. Anais CISCI 2011 Décima Conferencia Ibero-americana em Sistemas, Cibernética e Informática. Winter Garden : International Institute of Informatics and Systemics, 2011. v. I. p. 226-229.

[BUENO et al., 2010] BUENO, E. S. et al. **Análise objetiva HDTV do H.264 x MPEG-2 com e sem perda de pacotes**. Revista de Radiodifusão - SET, volume 04, número 04, 2010.

[BORTOLETO, 2005] BORTOLETO C. M. **Multicast Semi-Confável para Aplicações Multimídia Distribuídas**. Dissertação de Mestrado - Pontifícia Universidade Católica do Paraná, PUCPR, 2005.

[BRUNO, 2002] BRUNO, S. V. B. **Compressão de Dados sem Perda de Informação Usando Algoritmos de Recorrência de Padrões**. Pontifícia Universidade Católica do Rio de Janeiro, PUC-Rio, 2002.

[CASTRO, F.; CASTRO, M., 2001] CASTRO, F. C. C.; CASTRO, M. C. F. **Introdução ao Sistema MPEG de Codificação de Vídeo**. Pontifícia Universidade do Rio Grande do Sul, PUCRS, 2001.

[CHAGAS, 2011] C. E. **Open DVC: Uma Ferramenta para Simulação e Avaliação de Codificação Distribuída de Vídeo**. Dissertação de Mestrado - Pontifícia Universidade Católica do Rio de Janeiro, PUC-Rio, 2011.

- [CHIKKERUR et al., 2011] CHIKKERUR, S. et al. **Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison**. IEEE Broadcast Technology Society. Volume 57, p. 165 - 182, 2011.
- [DARONCO, 2009] DARONCO L. C. **Avaliação Subjetiva de Qualidade Aplicada à Codificação de Vídeo Escalável**. Universidade Federal do Rio Grande do Sul - UFRGS, 2009.
- [ESTRADA, 2011] ESTRADA, C. R. D. **Avaliação Automática de Qualidade de Videoconferências de Alta Definição**. Dissertação de Mestrado. Universidade Federal do Rio de Janeiro - UFRJ, 2011.
- [FARIAS; RODRIGUES, 2010] FARIAS, C. M.; RODRIGUES, P. H. A. **Métricas objetivas para avaliação da qualidade de vídeo percebida**. In: Simpósio Brasileiro de Sistemas Multimídia e Web - Webmedia, 2010, Belo Horizonte. Anais do Webmedia 2010, 2010. p. 1-8.
- [FERGUTZ, 2010] FERGUTZ, L. **Avaliação do network Calculus e VCC na caracterização de vídeos MPEG para sistemas multimídia**. Dissertação de Mestrado - Universidade Tecnológica Federal do Paraná, UTFPR, 2010.
- [FONSECA, 2008] FONSECA, R. N. **Avaliação da Qualidade de Vídeo em Televisão Digital**. Dissertação de Mestrado - Universidade de São Paulo, USP, 2008.
- [HANZO; CHERRIMAN; STREIT, 2007] HANZO, L.; CHERRIMAN, P. J.; STREIT J. **Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers**. Segunda Edição. John Wiley & Sons, Ltd., 2007.
- [IMG, 2006] IMG - **Compressão de vídeo digital**. http://www.img.lx.it.pt/~fp/cav/ano2005_2006/Trabalho_2/3compressaodevideodigital.htm - Acessado em Setembro/2011.
- [ITU-R, 2009] International Telecommunications Union. **Methodology for the subjective assessment of the quality of television pictures**. Recommendation ITU-R BT.500-12, 2009.
- [ITU-T, 2005] International Telecommunications Union H.264. **SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS Infrastructure of audiovisual services – Coding of moving video: Advanced video coding for generic audiovisual services**. ITU T Recommendation H.264, 2005.
- [LYRA et al., 2003] LYRA, A. L. et al. **Compressão sem perda: Método de Huffman e Método de Lempel– ZIV**. Universidade Federal de São Carlos, UFSCar, 2003.
- [MANOEL, 2007] MANOEL, E. T. M. **Codificação de Vídeo H.264 - Estudo de Codificação Mista de Macroblocos**. Dissertação de Mestrado - Universidade Federal de Santa Catarina, UFSC, 2007.
- [MPEG] MPEG - **The Moving Picture Experts Group (MPEG): The MPEG Home Page**. <http://mpeg.chiariglione.org/> - Acessado em Janeiro/2012

[PANSANATO, 2005] PANSANATO G. C. **Vídeo Digital - Padrão MPEG**. Universidade Estadual de Londrina, UEL, 2005.

[PINHEIRO, 2010] PINHEIRO, F. S. **Caracterização da Legibilidade de Documentos Digitalizados**. Universidade Federal do Pará, UFPA, 2010.

[REGIS, 2009] REGIS, C. D. M. **Avaliação de Técnicas de Redução da Resolução Espacial de Vídeos para Dispositivos Móveis**. Dissertação de Mestrado - Universidade Federal de Campina Grande, UFCG, 2009.

[SILVA, A., 2008] SILVA, A. **Transição para HDTV – XXIII: Recordando a compressão MPEG-2**. Produção Profissional - OUTUBRO, 2008

[SILVA, A. M., 2007] SILVA, A. M. C. **Um Estudo Sobre o Padrão H.264/AVC de Compressão de Vídeo**. Universidade Católica de Pelotas, UCPel, 2007.

[SILVA, E. et al., 2009] SILVA, E. S. R. et al. **Projeto e Construção de um Ambiente de Auxílio à Avaliação Subjetiva de Vídeo**. Revista de Radiodifusão - SET, volume 03, número 03, 2009.

[SILVA, W. et al., 2009] SILVA, W. B. et al. **Um Modelo de Referência Completa Para Avaliação Objetiva da Qualidade de Vídeo em Dispositivos Móveis em Ambientes do Sistema Brasileiro de TV Digital (SBTVD)**. Revista de Radiodifusão - SET, volume 03, número 04, 2009.

[TIMM et al., 2010] TIMM, E. et al. **Módulo de Interpolação de Crominância com Precisão de Quarto de Pixel e Subamostragem 4:2:0 para o Padrão H.264/AVC**. XIX CIC - XII ENPOS - II Mostra Científica - Universidade Federal de Pelotas, UFPel, 2010.

[THYAGARAJAN, 2011] THYAGARAJAN K. S. **Still Image and Video Compression with MATLAB**. Primeira Edição. John Wiley & Sons, Ltd, 2011.

[VENDRAMIN, 2004] VENDRAMIN, J. T. **Um algoritmo de suavização de tráfego para transmissão de vídeo digital comprimido no formato MPEG: Estudo de técnicas de suavização e sua avaliação**. Dissertação de Mestrado - Universidade Tecnológica Federal do Paraná, UTFPR, 2004.