



**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CAMPUS CURITIBA**

GERÊNCIA DE PESQUISA E PÓS-GRADUAÇÃO

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL - CPGEI**

Fábio Luiz Pessoa Albini

Geração e Avaliação de Artefatos em Vídeo Digital

DISSERTAÇÃO DE MESTRADO

**CURITIBA
2009**

Livros Grátis

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

Milhares de livros grátis para download.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial

DISSERTAÇÃO
apresentada à UTFPR
para obtenção do grau de

MESTRE EM CIÊNCIAS

por

FÁBIO LUIZ PESSOA ALBINI

**GERAÇÃO E AVALIAÇÃO DE ARTEFATOS EM VÍDEO
DIGITAL**

Banca Examinadora:

Presidente e Orientador:

PROF. DR. ALEXANDRE DE A. PRADO POHL

UTFPR

Examinadores:

PROF. DR. EVELIO MARTÍN G. FERNÁNDEZ

UFPR

PROF. DR. HUGO VIEIRA NETO

UTFPR

PROFA. DRA. KEIKO VERÔNICA O. FONSECA

UTFPR

Curitiba, março, 2009.

FÁBIO LUIZ PESSOA ALBINI

GERAÇÃO DE ARTEFATOS EM VÍDEO DIGITAL

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. Alexandre de A. Prado Pohl

Curitiba

2009

Ficha catalográfica elaborada pela Biblioteca da UTFPR – Campus Curitiba

A336g Albini, Fábio Luiz Pessoa
Geração de artefatos em vídeo digital / Fábio Luiz Pessoa Albini. – 2009.
73 p. : il. ; 30 cm

Orientador: Alexandre de A. Prado Pohl
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná.
Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração: Telemática, Curitiba, 2009
Bibliografia: p. 69 - 73

1. Vídeo digital. 2. Processamento de imagens – Técnicas digitais.
I. Pohl, Alexandre de Almeida Prado, orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração em Telemática. III. Título.

CDD 621.3

à minha noiva Loyse...

à minha família...

ao meu sobrinho Pietro...

aos meus colegas e amigos...

AGRADECIMENTOS

Primeiramente agradeço a Deus que tornou toda a minha vida possível e pelas inúmeras vezes em que, mesmo sem achar que precisava, me ajudou. Além de nunca ter me abandonado ou desamparado nesta longa caminhada.

À minha família que sempre me deu o suporte e amparo necessários, desde a minha criação até os dias atuais e com certeza continuará a me amparar e ansiar pelo meu crescimento profissional.

À minha noiva Loyse pela força, dedicação e amor repassados em todas as vezes que mais precisei. Por sempre estar presente ao meu lado, embora, muitas vezes, não fisicamente próxima, mas com toda a certeza, dentro do meu coração auxiliando e motivando, dando suporte e passando confiança em todos os momentos difíceis.

Aos meus colegas do LCD (Laboratório de Comunicação de Dados), da UTFPR – campus Curitiba, que por vezes, foram xingados, mal-interpretados e até mesmo desprezados, por uma pessoa (eu) que de tão cheia de problemas em sua cabeça, agia alheia à sua vontade, e eles se mostraram compreensivos, de coração aberto e dispostos a perdoar, pois no íntimo sabiam que todos nós passamos pelas mesmas pressões e necessidades.

Ao meu orientador professor Alexandre de Almeida Prado Pohl pela atenção em todos os seus conselhos e pela sua dedicação em aceitar um aluno, até o momento, desconhecido e fazer dele um pesquisador. Além disso, devo agradecer pela confiança e vontade demonstrados objetivando que seu aluno consiga sucesso em seu trabalho e reconhecimento.

E não poderia esquecer de agradecer aos outros amigos que ganhei durante todo o mestrado, principalmente ao Rafael Betito e sua esposa Andréia, juntamente com seu filho Guilherme que foram de fundamental importância em vários momentos importantes da minha vida.

Obrigado...

*“Aap apne karmon ke parinam ke bare shayad kabhi na jano, lekin yadi aap kuch na karo to koi
parinam hi na ho”*

*“Você talvez nunca saiba o resultado de sua ação, mas se você nada fizer não haverão
resultados”
(Gandhi)*

*“Tum woh badlav bano jo badlav duniya mein dekhna chahte ho”
“Seja a mudança que você deseja ver no mundo”
(Gandhi)*

*“Never, never, never give up!”
“Nunca, nunca, nunca desista!”
(Winston Churchill)*

RESUMO

Este trabalho teve como objetivo o estudo e a criação controlada de artefatos em vídeo digital e a sua mensuração com o emprego de métricas objetivas. Foram criados artefatos do tipo ruído branco gaussiano, sal e pimenta, borramento, blocagem e travamento. Para avaliação de seu impacto no vídeo digital foram utilizadas as métricas MSE (*Mean Squared Error*) e PSNR (*Peak Signal-to-Noise Ratio*)

Palavras-Chave: Qualidade de vídeo. Métricas com referência total. PSNR. MSE. Geração de artefatos em vídeo. Vídeo digital. Avaliação objetiva da qualidade de vídeo.

ABSTRACT

Generation and Evaluation of Artifacts in Digital Video

The objective of this work was the study and the controlled creation of digital video artifacts and their evaluation using objective metrics. Artifacts like salt and pepper, blurring, blocking, freezing and generated by white gaussian noise were created. To assess the impact of these artifacts on digital video the MSE (Mean Squared Error) and PSNR (Peak Signal-to-Noise Ratio) parameters were used.

Keywords: Quality of Video. Full-reference metrics. PSNR. MSE. Video artifacts generation. Digital video. Objective video quality assessment.

LISTA DE SIGLAS

BPP	-	Bit Por Pixel
CIF	-	Common Intermediate Format
DCT	-	Discrete Cosine Transform
DVD	-	Digital Video Disc
FFT	-	Fast Fourier Transform
FR	-	Full-Reference
HDTV	-	High Definition Television
HVS	-	Human Visual System
ITU	-	International Telecommunication Union
LGN	-	Lateral Geniculate Nucleus
LHS	-	Local Harmonic Strength
MOS	-	Mean Opinion Score
MPEG	-	Moving Picture Experts Group
MSAD	-	Mean Square Absolute Difference
MSE	-	Mean Square Error
MSU	-	Moscow State University
NR	-	No-Reference
NTSC	-	National Television Systems Committee
PAL	-	Phase Alternate Line
PSNR	-	Peak Signal-to-Noise Ratio
RR	-	Reduced-Reference
SECAM	-	Séquentiel Couleur avec Mémoire
SSIM	-	Structural Similarity
VQM	-	Video Quality Measurement
VAG	-	Variável Aleatória Gaussiana
VAU	-	Variável Aleatória Uniforme
VQEG	-	Video Quality Experts Group
WGN	-	White Gaussian Noise

LISTA DE TABELAS

TABELA 1. FORMATO DOS QUADROS DE VÍDEO.	22
TABELA 2. PARÂMETROS DA ITU-R BT.601-5.....	23
TABELA 3. COMPARAÇÃO DE RESULTADOS PARA O VALOR DE PSNR ENTRE O ALGORITMO DESENVOLVIDO E O SOFTWARE MSU.	47

LISTA DE FIGURAS

FIGURA 1. DESENHO ESQUEMÁTICO DO SISTEMA VISUAL HUMANO.	19
FIGURA 2. SISTEMA ÓTICO DO OLHO.	20
FIGURA 3. BARRAS COM TONS DE CINZA COM A MESMA VARIAÇÃO ENTRE CADA UMA.	21
FIGURA 4. QUADRO DE UM VÍDEO AMOSTRADO NO ROL DAS RESOLUÇÕES.	22
FIGURA 5. PADRÕES DE AMOSTRAGEM (PROGRESSIVO) 4:2:0, 4:2:2 E 4:4:4.	25
FIGURA 6. ALOCAÇÃO DOS CAMPOS SUPERIOR E INFERIOR DA AMOSTRAGEM 4:2:0.	26
FIGURA 7. EXEMPLO DO “CHUVISCO” ENCONTRADO NA IMAGEM DA TELEVISÃO ANALÓGICA.	27
FIGURA 8. ILUSTRAÇÃO DO ARTEFATO BLOCAGEM.	28
FIGURA 9. EXEMPLO DE PSNR: (A) ORIGINAL, (B) 30,6 dB E (C) 28,3 dB.	31
FIGURA 10. IMAGEM COM O FUNDO BORRADO (PSNR = 27,7 dB).	32
FIGURA 11. DESENHO ESQUEMÁTICO DA DISPOSIÇÃO DOS EQUIPAMENTOS.	35
FIGURA 12. ESTRUTURA DOS ARQUIVOS YUV.	36
FIGURA 13. FLUXOGRAMA DO FUNCIONAMENTO BÁSICO DOS ALGORITMOS.	37
FIGURA 14. HISTOGRAMA DO GERADOR DE VARIÁVEIS ALEATÓRIAS COM DISTRIBUIÇÃO GAUSSIANA, COM MÉDIA 0 E DESVIO PADRÃO 40.	39
FIGURA 15. QUADRO DO VÍDEO “AKIYO” ONDE (A) OS <i>PIXELS</i> POSSUEM 2% DE PROBABILIDADE DE OCORRÊNCIA E (B) OS <i>PIXELS</i> POSSUEM 15% DE PROBABILIDADE DE OCORRÊNCIA.	39
FIGURA 16. QUADRO DO VÍDEO “MOBILE” COM ARTEFATO SAL E PIMENTA, ONDE EM (A) A PROBABILIDADE DE CADA <i>PIXEL</i> SER SELECIONADO É 1% E EM (B) A PROBABILIDADE DE CADA <i>PIXEL</i> SER SELECIONADO É 5%.	40
FIGURA 17. QUADRO DO VÍDEO “STEFAN” COM A APLICAÇÃO DO ARTEFATO BORRAMENTO EM (A) UTILIZANDO UMA MÁSCARA LINEAR DE 5 <i>PIXELS</i> E (B) UTILIZANDO UMA MÁSCARA MATRICIAL 7x7.	44
FIGURA 18. QUADRO DO VÍDEO “COASTGUARD” COM A APLICAÇÃO DO ALGORITMO GERADOR DE <i>BLOCKING</i> , COM OCORRÊNCIA DE 5% POR <i>PIXEL</i>	44
FIGURA 19. GRÁFICOS DE CORRELAÇÃO DA MOS X PSNR PARA ALGUNS DOS ARTEFATOS GERADOS E AVALIADOS.	49
FIGURA 20. GRÁFICO DA COMPARAÇÃO DA MÉTRICA SUBJETIVA (MOS) COM A MÉTRICA OBJETIVA (PSNR) PARA TODOS OS ARTEFATOS DE UM TRAÇO DE VÍDEO.	50
FIGURA 21. MSE (A) E PSNR (B) PARA 2% DE RUÍDO BRANCO GAUSSIANO.	52
FIGURA 22. MSE (A) E PSNR (B) PARA 50% DE RUÍDO BRANCO GAUSSIANO.	52
FIGURA 23. MSE (A) E PSNR (B) PARA 100% DE RUÍDO BRANCO GAUSSIANO.	53
FIGURA 24. QUADRO DO VÍDEO “COASTGUARD” APRESENTANDO O RUÍDO BRANCO GAUSSIANO, COM OCORRÊNCIA DE 2% (A), 50% (B) E 100% (C).	54
FIGURA 25. MSE (A) E PSNR (B) PARA 2% DE RUÍDO SAL E PIMENTA.	55
FIGURA 26. MSE (A) E PSNR (B) PARA 5% DE RUÍDO SAL E PIMENTA.	55
FIGURA 27. QUADRO DO VÍDEO “COASTGUARD” APRESENTANDO RUÍDO SAL E PIMENTA, COM OCORRÊNCIA DE 2% (A) E COM 5% (B).	56
FIGURA 28. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA LINEAR DE 3 <i>PIXELS</i>	57
FIGURA 29. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA LINEAR DE 5 <i>PIXELS</i>	57

FIGURA 30. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA LINEAR DE 7 PIXELS.	58
FIGURA 31. QUADRO DO VÍDEO “COASTGUARD” COM BORRAMENTO GERADO PELAS MÁSCARAS LINEARES DE 3 (A), 5 (B) E 7 (C) PIXELS.	59
FIGURA 32. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA MATRICIAL DE 3X3 PIXELS.	60
FIGURA 33. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA MATRICIAL DE 5X5 PIXELS.	60
FIGURA 34. MSE (A) E PSNR (B) PARA O BORRAMENTO UTILIZANDO A MÁSCARA MATRICIAL DE 7X7 PIXELS.	61
FIGURA 35. QUADRO DO VÍDEO “COASTGUARD” COM BORRAMENTO GERADO PELAS MÁSCARAS MATRICIAIS DE 3X3 PIXELS (A), 5X5 PIXELS (B), 7X7 PIXELS (C).....	62
FIGURA 36. MSE (A) E PSNR (B) PARA 1% DE OCORRÊNCIA DE BLOCAGEM, 20 QUADROS FORAM AFETADOS.	63
FIGURA 37. MSE (A) E PSNR (B) PARA 5% DE OCORRÊNCIA DE BLOCAGEM, 6 QUADROS FORAM AFETADOS.	64
FIGURA 38. QUADRO DO VÍDEO “COASTGUARD” APRESENTANDO O ARTEFATO BLOCAGEM COM OCORRÊNCIA DE 1% (A) E 5% (B).....	65
FIGURA 39. MSE (A) E PSNR (B) PARA O TRAVAMENTO.	66

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	MOTIVAÇÃO.....	13
1.2	OBJETIVOS	15
1.3	ESTRUTURA DA DISSERTAÇÃO	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	SISTEMA VISUAL HUMANO (HVS – <i>HUMAN VISUAL SYSTEM</i>)	18
2.2	FORMATOS DE VÍDEO	21
2.3	FORMATOS DE AMOSTRAGEM YCbCr	23
2.4	ARTEFATOS	26
2.5	MÉTRICAS DE QUALIDADE DE VÍDEO BASEADAS EM <i>PIXEL</i> COM REFERÊNCIA TOTAL	29
2.5.1	<i>MSE</i>	29
2.5.2	<i>PSNR</i>	30
3	METODOLOGIA	33
3.1	MATERIAIS	33
3.2	FORMATO DOS ARQUIVOS YUV	35
3.3	GERADOR DE ARTEFATOS	36
3.3.1	<i>Ruído Branco Gaussiano</i>	37
3.3.2	<i>Sal e Pimenta</i>	40
3.3.3	<i>Borramento</i>	41
3.3.4	<i>Blocagem</i>	43
3.3.5	<i>Travamento</i>	44
3.4	<i>MSU VIDEO QUALITY MEASUREMENT TOOL</i>	45
3.5	<i>PSNR</i>	46
3.5.1	<i>Validação do PSNR</i>	46
3.6	COMPARAÇÃO COM MEDIDAS SUBJETIVAS	47
4	RESULTADOS.....	51
4.1	RUÍDO BRANCO GAUSSIANO	51
4.2	SAL E PIMENTA	55
4.3	BORRAMENTO	56
4.4	BLOCAGEM	63
4.5	TRAVAMENTO	65
5	CONCLUSÃO E TRABALHOS FUTUROS	67

1 INTRODUÇÃO

1.1 MOTIVAÇÃO

A medida da qualidade de imagens e sequências de imagens digitalmente codificadas, tem se tornado um assunto muito importante e vem recebendo grande atenção da comunidade de processamento de sinais (LAMBRECHT, VERSCHEURE, 1996).

Avaliar a qualidade de vídeo é uma necessidade para todas as transmissoras de televisão. O conhecimento da qualidade da imagem transmitida permite aos operadores do sistema saber quais atributos envolvidos na transmissão devem ser alterados e qual impacto cada um deles tem no vídeo observado pelo telespectador.

O sistema visual humano é complexo e ainda mal-compreendido (WHITAKER, BENSON, 2003), por isso, o desenvolvimento de métodos automáticos de avaliação da qualidade visual é, ainda, uma tarefa difícil.

Efetivar a avaliação da qualidade da imagem de um vídeo utilizando-se de espectadores (métrica subjetiva) é uma tarefa árdua, dispendiosa e demorada. Com o objetivo de evitar esta demora e desgaste, são utilizadas métricas objetivas (automatizadas) como opção para estimar a sensação visual de uma pessoa que assiste a imagem transmitida.

As métricas objetivas são divididas em métricas com referência total, com referência reduzida e sem referência. Com referência total são as métricas que se baseiam no vídeo original para avaliar o vídeo degradado, sua aplicação envolve a avaliação dos algoritmos de compressão e de transmissão. As métricas sem referência se baseiam apenas no vídeo a ser avaliado sem nenhuma informação do vídeo original. 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. Já as métricas com referência reduzida utilizam-se de algumas (poucas) informações do vídeo original para efetuar a avaliação e são aplicadas, em geral, da mesma maneira que as métricas com referência total.

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, borramento, bloqueio e travamento nas transmissões digitais são alguns dos artefatos observados [DTV, 2008].

A causa das degradações em transmissões analógicas, em sua maioria, é a baixa relação sinal/ruído, problemas com antena e locais onde o sinal acaba sendo impossibilitado de chegar, são os principais causadores dessa baixa relação. Já nas transmissões digitais, a maioria dos artefatos são causados pela compressão com perda que é efetuada para redução da quantidade de informação a ser enviada. Através dessa compressão são gerados alguns artefatos como bloqueio e borramento, pois devido à perda relativa à compressão, alguns *pixels* possuem sua informação alterada. Por outro lado o travamento, geralmente, é decorrente da perda de informações, devido ao baixo nível de sinal ou perda de pacotes.

Antes de se desenvolver novas métricas de qualidade de imagem é necessário obter vários vídeos que apresentem os artefatos encontrados na transmissão real. Porém, ao tentar gerar estes artefatos de forma natural, como por exemplo: manipulando a potência de sinal, alterando características da antena de recepção ou adicionando interferência no meio de transmissão, encontram-se alguns empecilhos como: a combinação de artefatos (muitas vezes desconhecida), a falta de controle na quantidade de cada artefato que está sendo inserido e a difícil caracterização dos artefatos presentes no vídeo. Por este motivo é interessante a criação de algoritmos que gerem estes artefatos de maneira controlada. Com estes algoritmos será possível ter a certeza da quantidade de artefatos que está sendo inserida além de quais artefatos estão realmente presentes no vídeo.

A partir da imagem degradada, identifica-se visualmente o artefato que afeta a qualidade da imagem, o qual pode ser, também, detectado objetivamente. Para isso são utilizadas métricas objetivas com referência total, sendo elas o PSNR (*Peak Signal-to-Noise Ratio*) – razão sinal/ruído de pico e o MSE (*Mean Square Error*) erro quadrático médio, ambos calculados utilizando-se os valores dos *pixels* da imagem original e degradada. Além disso, é possível utilizar cada tipo de artefato para o desenvolvimento de novas métricas. Estas poderão ser embarcadas para que, por exemplo, o próprio decodificador detecte os artefatos e possa aplicar a rotina correta para corrigir a degradação observada, melhorando assim a qualidade da imagem e reduzindo o impacto visual que aquele “defeito” teria sobre o telespectador.

1.2 OBJETIVOS

O principal objetivo deste trabalho é a construção de algoritmos para a geração de artefatos em vídeo digital, tais como aqueles gerados pelo ruído branco gaussiano, ruído sal e pimenta, blocagem, borramento e travamento. A geração controlada de artefatos nos vídeos pode auxiliar no desenvolvimento futuro de novas métricas de qualidade de vídeo e sua detecção.

Outra questão importante é aprender e aplicar a métrica de avaliação da qualidade, nos vídeos degradados pelo gerador, utilizando-se métodos com referência total, com isso, procura-se estabelecer um limiar de PSNR, a partir do qual pode-se considerar o vídeo com qualidade ruim.

1.3 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está organizada em cinco capítulos. No Capítulo 2 é apresentada a fundamentação teórica, a qual apresenta conceitos relevantes para o entendimento do trabalho, tais como: características do sistema visual humano, formatos de vídeo, formatos de amostragem, tipos de artefatos e algumas das métricas de qualidade existentes.

O Capítulo 3 descreve em detalhes os materiais utilizados, bem como a metodologia adotada para realizar esta pesquisa, além de mostrar e ilustrar as formas e aplicações das ferramentas utilizadas e suas validações.

No Capítulo 4 relatam-se os resultados obtidos, apresentando gráficos e quadros dos vídeos gerados, seguidos por análises dos resultados. E, finalmente, o Capítulo 5 apresenta a discussão dos resultados, as conclusões do trabalho realizado e as propostas para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Para o observador de uma imagem ou um espectador de televisão a qualidade do que se está vendo é fundamental para apreciar o objeto visual em toda sua extensão e variedade. A transição do domínio analógico para digital provocou o advento de novas tecnologias de armazenamento e reprodução de vídeo, tornando a exigência sobre qualidade de vídeo por parte dos telespectadores cada vez maior. Portanto, medir a qualidade da imagem que está sendo transmitida é uma maneira de tentar prever a experiência do espectador. Além disso, conhecendo o impacto visual no usuário, é possível tomar decisões e manipular as variáveis do sistema de codificação e transmissão para que sejam obtidos melhores resultados quanto à qualidade de imagem final (recebida).

Existem basicamente duas formas de se avaliar a qualidade de uma imagem ou vídeo: de forma subjetiva ou objetiva. O método subjetivo utiliza-se de observadores (humanos), os quais assistem a um vídeo ou imagem e fornecem a sua opinião (uma nota) para a qualidade experimentada. Após isso, é feita uma análise estatística dos dados coletados, que resulta em uma nota chamada MOS (*Mean Opinion Score*) (ITU, 2000), a qual é utilizada para a avaliação da qualidade percebida. Segundo (WANG, et al. 2004), esta é a única forma “correta” de se quantificar a qualidade da imagem, pois neste método os próprios espectadores avaliam a percepção visual do vídeo, cabendo a eles quantificar a sua qualidade. E de acordo com (KOUMARAS, PLIAKAS e KOURTIS, 2007), a MOS é seguramente o método mais preciso de medida da qualidade visual, e tem sido aplicada na grande maioria das técnicas de medida subjetiva. Existem algumas recomendações para avaliação subjetiva. Uma das mais utilizadas é a BT.500 da ITU (ITU, 2000), onde são especificados procedimentos para avaliações subjetivas, tais como: distância ideal dos espectadores à imagem, número de espectadores por teste, método de exibição do vídeo, iluminação do laboratório de testes, seleção do material a ser utilizado nos testes, tempo de duração da avaliação e características da ficha do avaliador. Outra recomendação existente para avaliação subjetiva é a P.910 também da ITU (ITU, 2008). Nesta são apresentadas sugestões para: arranjo das condições de visualização, número de participantes, instruções e treinamento dos espectadores e análise estatística dos resultados. Todavia, o método subjetivo demanda muito tempo para a sua execução, precisa de um grande número de participantes e

sendo assim torna-se muito dispendioso para ser implantado e mantido por longos períodos de tempo (DURANT, 1998). Um dos motivos pelo qual são utilizados e desenvolvidos métodos objetivos é para contornar este problema.

Os métodos objetivos de avaliação da qualidade visual consistem em algoritmos que visam medir características da imagem ou vídeo para fazer uma previsão, aproximada, da qualidade observada pelos telespectadores (EDEN, 2008). Muitos trabalhos relacionados às métricas objetivas da qualidade de imagem são encontrados na literatura. Estes trabalhos sobre avaliação objetiva podem ser classificados em função da quantidade de informação necessária para o cálculo da qualidade do vídeo. Esta classificação se dá em três classes: sem referência (NR - *no-reference*), com referência reduzida (RR - *reduced-reference*) e com referência total (FR - *full-reference*) (WANG, BOVIK, 2006), (WU, RAO, 2006), (KOUMARAS, PLIAKAS, KOURTIS, 2007).

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). Ainda não se conhecem métricas sem referência que tenham sido totalmente eficientes e que possam ser utilizadas comercialmente para simular efetivamente o sistema visual humano (HVS – *Human Visual System*) dos espectadores (KOUMARAS, PLIAKAS e KOURTIS, 2007), (RICHARDSON, 2003).

Entretanto, as métricas com referência reduzida, ao contrário das sem referência, utilizam algumas informações do vídeo original, ou seja, quando o vídeo é transmitido, junto com ele são enviadas informações importantes sobre a sua imagem original, tais como, perda harmônica e ganho harmônico (GUNAWAN, e GHANBARI, 2003). 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 FFT – *Fast Fourier Transform*. Estes valores são calculados tanto para o vídeo original, como para o degradado e aí então comparados para avaliar a qualidade da imagem (GUNAWAN, e GHANBARI, 2007). Estas métricas também possuem um grau de complexidade

elevado e pouca correlação com a avaliação subjetiva (WANG, BOVIK, 2006). Existem poucas propostas de métricas com referência reduzida na literatura, devido a este tópico de pesquisa ser muito recente. Ainda não se tem conhecimento de alguma proposta que tenha se destacado quanto à eficiente correlação da avaliação objetiva de vídeos com o HVS e que possa ser utilizada comercialmente (WANG, BOVIK, 2006), (RICHARDSON, 2003).

Por outro lado, as métricas com referência total utilizam como base o vídeo original para a avaliação da qualidade. Ou seja, ao realizar comparações entre o vídeo original e o exibido ao telespectador estas métricas tentam predizer a opinião subjetiva dos usuários. Este tipo de métrica possui geralmente um grau de complexidade baixo e uma implementação mais rápida.

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.

2.1 Sistema Visual Humano (HVS – *Human Visual System*)

A meta mais importante para os sistemas de codificação de vídeo e imagem é que o resultado produzido deve ser aceitável e prazeroso ao observador, no caso o ser humano. Para

atingir este objetivo é necessário levar em conta a resposta do Sistema Visual Humano (HVS). O HVS é o sistema pelo qual um observador humano enxerga, interpreta e responde ao estímulo visual (WU, RAO, 2006).

O sistema visual humano é poderoso e excede o desempenho dos sistemas visuais artificiais em quase todas as áreas de comparação (WHITAKER e BENSON, 2003). Ele pode ser dividido em quatro estágios, o primeiro é o processamento ótico, seguido do processamento da retina, do processamento LGN (*Lateral Geniculate Nucleus*) e por último o processamento cortical (WANG e BOVIK, 2006). A Figura 1 ilustra a disposição desses estágios.

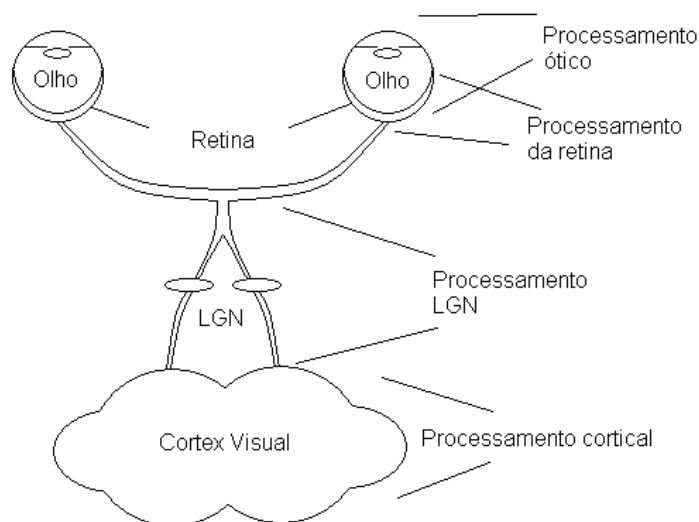


Figura 1. Desenho esquemático do sistema visual humano.

No primeiro estágio, a imagem de um objeto passa através da ótica ocular e é projetada sobre a retina, uma membrana foto-sensível no fundo do olho. O resultado é chamado imagem da retina. A ótica ocular é composta por três elementos básicos, ilustrados na Figura 2: a córnea, a pupila e a lente (o cristalino) (WANG e BOVIK, 2006).

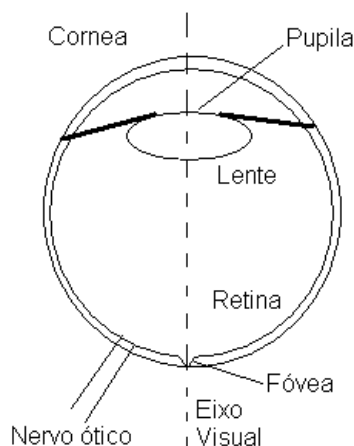


Figura 2. Sistema ótico do olho.

A retina é composta por várias camadas de neurônios. A última camada é formada por células fotoreceptoras que amostram a imagem da retina projetada sobre ela. Existem dois tipos de fotoreceptores: os cones e os bastonetes. Os cones são responsáveis pela visão em cores em condições de luz normal à alta, enquanto os bastonetes são responsáveis pela visão monocromática em condições de pouca luminosidade. Existem três tipos de cones categorizados de acordo com o espectro de sensibilidade aos comprimentos de luz, os cones G, cones M e cones P. Estes cones correspondem aos comprimentos de onda grandes (570nm), médios (540nm) e pequenos (440nm) respectivamente. Estes valores de comprimento de onda estão relacionados aos picos de sensibilidade espectral de cada tipo de cone. A percepção das cores é codificada pela atividade relativa destes três tipos de cones. Por outro lado, todos os bastonetes possuem a mesma sensibilidade espectral, (picos em 500 nm), e não codificam informações de cor. O sinal amostrado do foto-receptor passa por várias camadas de neurônios interconectados antes de ser transmitido para as células ganglionárias, que são responsáveis por transmitir a informação visual da retina ao cérebro (WANG e BOVIK, 2006).

O observador humano é o avaliador final da qualidade da imagem. Por isso é importante incorporar características do HVS nas aplicações de processamento de imagem. Além disso, é conhecido que a resposta do olho às mudanças de intensidade luminosa não é linear. Segundo (WU, RAO, 2006), a sensibilidade dos olhos humanos para discernir entre diferenças de intensidade não dependem apenas da diferença, mas também do nível de intensidade.

Em imagens de 8 bpp - bits por *pixel* (*i.e.* valores de luminância entre 0 e 255), e com valor de luminância de fundo com intensidade entre 0 e 50 (zona de baixo nível de cinza) o HVS é mais sensível quando encontra-se a pequenas distâncias do monitor. Por outro lado, quando os

valores de intensidade para o fundo são superiores a 50, a distância de observação não influencia a diferença percebida. Por influência das condições de luminosidade ambiente, o HVS é mais sensível às áreas de baixo nível de cinza quando a sala está escura e é indiferente para as áreas de alto nível de cinza. Já o brilho do monitor afeta diretamente a sensibilidade de mudanças, ou seja, para um alto brilho do monitor, o HVS é mais sensível e para um baixo brilho é menos sensível às mudanças. Além disso, os erros de quantização são imperceptíveis em uma faixa de baixa luminância e se tornam perceptíveis em uma faixa de média luminância (WU, RAO, 2006).

Outra característica importante é observada quando existem bordas salientes em uma imagem, como é possível observar na Figura 3. Esta é composta por oito barras com tons de cinza na vertical, dentro de cada barra a intensidade de cinza é mantida constante, porém quando se observa na parte mais à direita de uma barra (próximo da fronteira com outra barra) cria-se uma ilusão de que esta parte seja mais escura que o meio da barra, e ao observar a parte mais à esquerda tem-se a ilusão de esta seja mais clara que o meio da barra. A transição de uma faixa para outra é enfatizada pela sensibilidade à intensidade do sistema visual humano (WU, RAO, 2006). Este fenômeno é conhecido na literatura por “*Mach Bands*”.

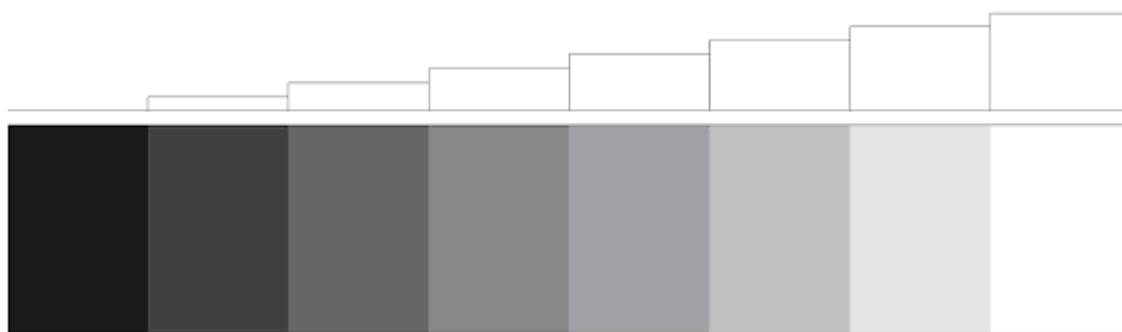


Figura 3. Barras com tons de cinza com a mesma variação entre cada uma (WU, RAO, 2006).

2.2 Formatos de vídeo

Na prática, é comum capturar ou converter um vídeo para um dos ‘formatos intermediários’ antes de comprimir ou transmiti-lo. O Formato Intermediário Comum (CIF –

Common Intermediate Format) é a base de um conjunto popular de formatos listados na Tabela 1.

Tabela 1. Formato dos quadros de vídeo.

Formato	Resolução de Luminância (horizontal × vertical)	Bits por quadro (4:2:0, oito bits por amostra)
Sub-QCIF	128 × 96	147456
Quarter CIF (QCIF)	176 × 144	304128
CIF	352 × 288	1216512
4CIF	704 × 576	4866048

A Figura 4 mostra a componente de luminância de um quadro de vídeo amostrado numa faixa de resoluções, desde 4CIF até Sub-QCIF. A escolha da resolução do quadro depende da aplicação e da capacidade de armazenamento e transmissão disponível. Por exemplo, 4CIF é apropriado para televisão de definição padrão e para vídeo em DVD; CIF e QCIF são populares para aplicações de videoconferência; QCIF e SQCIF são apropriados para aplicações móveis multimídia onde a resolução do display e a taxa de bits (*bitrate*) são limitadas. A Tabela 1 lista o número de bits necessários para representar um quadro descomprimido em cada um dos formatos (assumindo uma amostragem 4:2:0 e 8 bits por amostra de luminância e crominância, estas características serão descritas na Seção 2.3) (RICHARDSON, 2003).

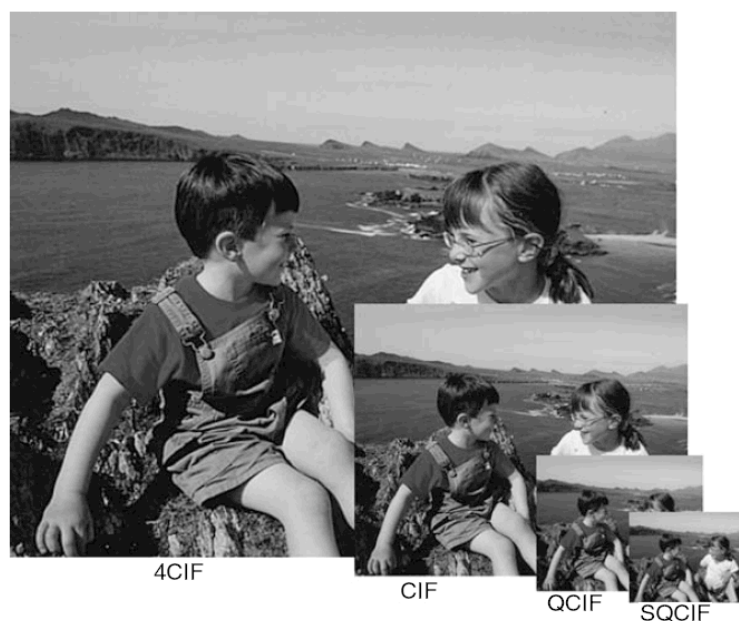


Figura 4. Quadro de um vídeo amostrado no rol das resoluções (RICHARDSON, 2003).

Um formato largamente utilizado para codificar digitalmente sinais de vídeo para produção de televisão é apresentado na recomendação BT.601 da ITU¹ (ITU, 2007). A componente de luminância do sinal de vídeo é amostrada em 13,5 MHz e a crominância em 6,75 MHz para produzir um sinal de componentes Y:Cb:Cr a 4:2:2. Os parâmetros do sinal digital amostrado dependem da taxa de quadros do vídeo (30 Hz para um sinal NTSC e 25 Hz para um sinal PAL/SECAM) e são exibidos na Tabela 2. A maior taxa de quadros de 30 Hz do NTSC é compensada por resolução espacial menor. Assim a taxa de bits será igual (216 Mbps) para ambos os padrões NTSC e PAL/SECAM. A área efetiva mostrada no monitor, área ativa, é menor que o total porque são excluídas as linhas correspondentes aos intervalos de sincronismo horizontal e vertical que existem fora da área visível do quadro (RICHARDSON, 2003).

Tabela 2. Parâmetros da ITU-R BT.601-5 (RICHARDSON, 2003).

	Taxa de quadros de 30 Hz	Taxa de quadros de 25 Hz
Campos por segundo	60	50
Linhas por quadro completo	525	625
Amostras de luminância por linha	858	864
Amostras de crominância por linha	429	432
Bits por amostra	8	8
Taxa de bits total	216 Mbps	216 Mbps
Linhas ativas por quadro	480	576
Amostras ativas por linha (Y)	720	720
Amostras ativas por linha (Cr, Cb)	360	360

Cada amostra tem uma faixa possível de valores que variam de 0 a 255 (8 bits). Os níveis 0 e 255 são reservados para sincronização e o sinal de luminância ativa acaba sendo restrito a uma faixa de 16 (preto) a 235 (branco) (RICHARDSON, 2003).

2.3 Formatos de amostragem YCbCr

Os *pixels* podem ser representados de diversas maneiras, uma delas utiliza as componentes R, G e B, que significam a quantidade de cada uma das cores primárias do *pixel*, sendo elas, respectivamente, vermelho (R – *Red*), verde (G – *Green*) e azul (B – *Blue*). Outra

¹ O termo ‘codificar’ no título da recomendação significa conversão para o formato digital e não implica em compressão.

maneira amplamente utilizada em vídeo digital é a representação Y, Cb e Cr, onde Y representa a luminância (brilho) enquanto Cb e Cr representam a crominância (cor).

A Figura 5 mostra três padrões de amostragem para Y, Cb e Cr que são suportados pelo MPEG-4 Visual e H.264. A amostragem 4:4:4 significa que as três componentes (Y, Cb e Cr) são amostradas com a mesma taxa e existe uma amostra de cada componente para a posição de cada *pixel* na imagem. Os números indicam um múltiplo da taxa de amostragem relativa de cada componente na direção horizontal, *i.e.* cada componente é amostrada com 4 (quatro) vezes a taxa básica do padrão. A amostragem 4:4:4 preserva a fidelidade total das componentes de crominância. Na amostragem 4:2:2 (algumas vezes referida como YUY2), as componentes de crominância possuem a mesma resolução vertical que a luminância, mas a metade da resolução horizontal (os números 4:2:2 significam que para cada quatro amostras de luminância na direção horizontal existem duas amostras Cb e duas Cr). Vídeos no formato 4:2:2 são utilizados para uma reprodução de alta qualidade (RICHARDSON, 2003).

No formato de amostragem popular 4:2:0 ('YV12'), Cb e Cr possuem, cada um, a metade da resolução vertical e horizontal da componente Y. O Termo '4:2:0' é um pouco confuso porque os números não possuem uma interpretação lógica e parecem ter sido escolhidos, historicamente, como um 'código' para identificar este padrão de amostragem, em particular para diferenciá-lo do 4:4:4 e 4:2:2. A amostragem 4:2:0 é largamente utilizada para aplicações como vídeo conferência, televisão digital e armazenamento em DVDs. Devido ao fato de cada componente de crominância possuir um quarto do número de amostras da componente Y, o vídeo 4:2:0 requer exatamente a metade das amostras que o vídeo no formato 4:4:4 precisa (RICHARDSON, 2003).

A amostragem 4:2:0 algumas vezes é descrita como '12 bits por *pixel*'. A razão para isso pode ser observada examinando um grupo de quatro *pixels* (vide grupo delimitado pelas linhas pontilhadas na Figura 5). Usando a amostragem 4:4:4, um total de 12 amostras são requeridas, quatro de cada Y, Cb e Cr. Para uma amostragem com 8 bits/amostra tem-se um total de $12 \times 8 = 96$ bits, ou seja, uma média de $96/4 = 24$ bits por *pixel*. Usando a amostragem 4:2:0, são necessárias quatro amostras para a componente Y e uma para cada componente Cb e Cr, sendo necessário, portanto, um total de $6 \times 8 = 48$ bits. Ou seja, uma média de $48/4 = 12$ bits por *pixel* (RICHARDSON, 2003).

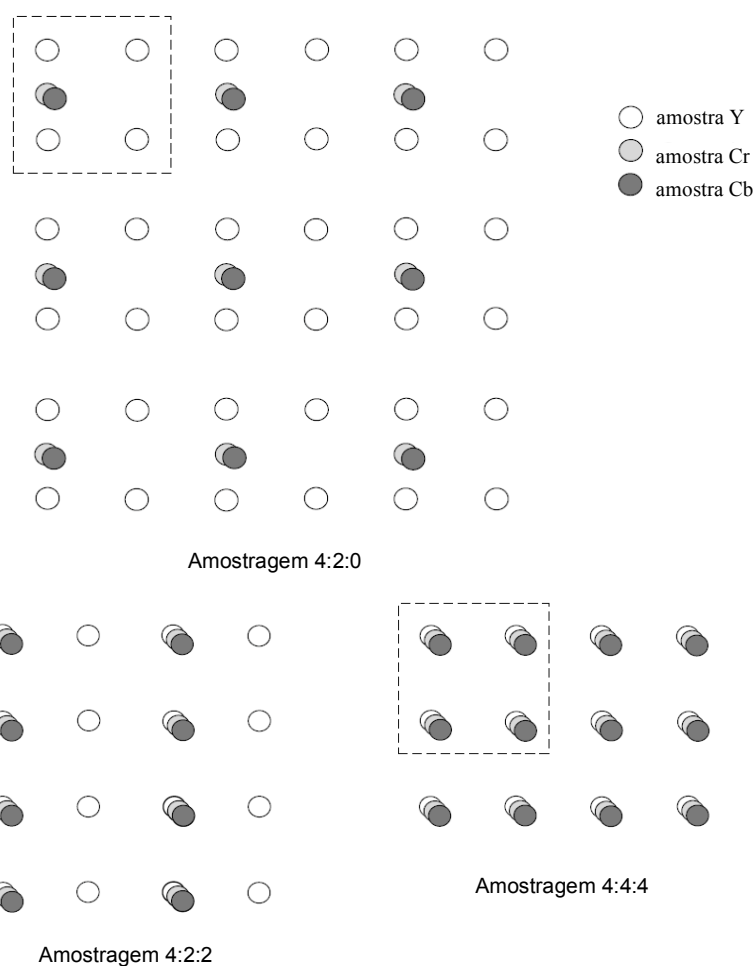


Figura 5. Padrões de amostragem (progressivo) 4:2:0, 4:2:2 e 4:4:4 (RICHARDSON, 2003).

Além disso, o vídeo pode ser entrelaçado ou progressivo. No vídeo entrelaçado, as informações de um quadro são separadas em dois campos, superior e inferior, onde o campo superior possui as linhas ímpares e o inferior possui as linhas pares de luminância, sendo estas linhas intercaladas com e sem a componente de croma. Ou seja, a primeira linha ímpar começa recebendo a componente de croma enquanto a primeira linha par não a contém. Na sequência a segunda linha ímpar deixa de possuir a componente de croma ao passo que a segunda linha par a recebe, e assim segue alternadamente. Isto pode ser visualizado na Figura 6. Por outro lado o quadro progressivo é representado continuamente (progressivamente), *pixel* após *pixel*, sem separação alguma.

Numa sequência de vídeo entrelaçada 4:2:0, as amostras Y, Cb e Cr, correspondentes a um quadro completo do vídeo, são alocadas em dois campos. A Figura 6 mostra o método de

alocação das amostras Y, Cb e Cr em um par de campos entrelaçados adotados no MPEG-4 Visual e H.264. Fica claro, a partir da figura, que o número total de amostras num par de campos é o mesmo que o número de amostras em um quadro progressivo equivalente (RICHARDSON, 2003).

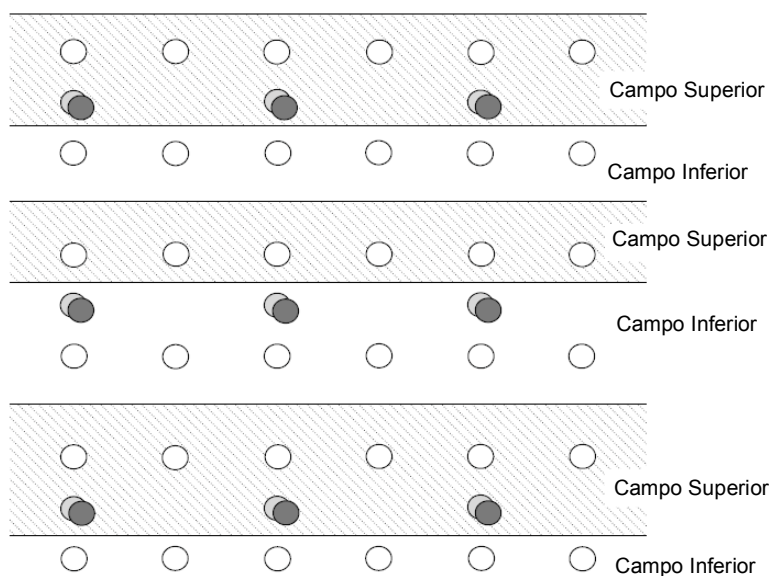


Figura 6. Alocação dos campos superior e inferior da amostragem 4:2:0 (RICHARDSON, 2003).

2.4 Artefatos

Quando uma imagem ou vídeo é degradado, várias características peculiares podem ser encontradas na área de observação e que são chamadas de artefatos. Alguns exemplos de artefatos são: o ruído branco gaussiano e o ruído sal e pimenta, que se apresentam como um “chuvisco”. Há também o borramento da imagem, a blocagem, e outros que serão discutidos a seguir.

Os “chuviscos” existentes em sinais de vídeo analógico, exemplificados na Figura 7, são gerados pelo ruído branco gaussiano (WGN – *White Gaussian Noise*), que provoca queda na qualidade do sinal recebido. Esta queda é gradativa, pois à medida que a relação sinal/ruído

diminui, uma maior degradação da imagem é percebida. O ruído é chamado branco por estar presente em todo o espectro de frequência e não poder ser evitado (SCHWARTZ e SHAW, 1975), (MENDES e FASOLO, 2002). Ele é chamado gaussiano, devido à amplitude da interferência possuir uma distribuição gaussiana.



Figura 7. Exemplo do “chuveiro” encontrado na imagem da televisão analógica.

O ruído sal e pimenta ocorre quando se altera o valor de luminância de alguns *pixels* para o seu valor máximo e outros para o seu valor mínimo. Esses valores se apresentam como pontos brancos e pretos, respectivamente, e por isso este ruído recebe o nome de sal e pimenta.

A blocagem (*blocking*) e o borramento (*blurring*) existentes apenas no vídeo digital são artefatos que ocorrem principalmente devido à compressão da imagem utilizando-se a Transformada Discreta de Co-senos (DCT – *Discrete Cosine Transform*), mas podem também ocorrer devido à perda de informações da imagem (LIU e BOVIK, 2002), (MONTEZ e BECKER, 2004), ou até mesmo devido à interpolação ou efeito zoom (BOVIK, 2000). Como os vídeos digitais necessitam de uma banda muito grande para serem transmitidos sem compressão, é necessário o uso de alguma ferramenta para a redução da quantidade de bits a serem transmitidos (redução da redundância) o que geralmente implica na perda de qualidade visual (STOLFI, 2007).

Quando submetidos à compressão, os vídeos acabam sofrendo degradações devido ao aparecimento de artefatos que não existiam na sua forma original. A blocagem é, talvez, a

distorção mais estudada na codificação baseada em macro blocos. Visualmente, este efeito é observado como uma descontinuidade entre blocos adjacentes numa imagem. É possível observar uma variedade de exemplos deste efeito na Figura 8, sendo mais evidente em torno da face e do veículo no plano de fundo. Os cantos acentuados induzidos pela descontinuidade da vizinhança dos blocos, e pelo espaçamento dos blocos, resultam no efeito de blocagem (WU e RAO, 2006).



Figura 8. Ilustração do artefato blocagem (WU e RAO, 2006).

A perda dos detalhes espaciais em regiões da imagem com atividade espacial entre moderada e alta, e dos contornos dos objetos geram o defeito de borramento. Nos macro-blocos de um quadro codificado, o borramento está diretamente relacionado com a supressão dos coeficientes de ordem elevada da DCT (WU e RAO, 2006). Estes coeficientes de ordem elevada caracterizam as bordas e detalhes finos de uma imagem (SANTOS, 2002). Visualmente o borramento é observado como uma imagem fora de foco. A Figura 9(C) mostra um exemplo deste artefato.

O travamento instantâneo observado em um vídeo pode, por exemplo, ser decorrente da ausência de informações no receptor. Na televisão digital, isto ocorre quando há perda de pacotes na transmissão ou processamento. Como uma forma de amenização de erros, o último quadro recebido é repetido pelo decodificador ou *player* (KUROSE e ROSS, 2007).

No decorrer do desenvolvimento desta dissertação foi encontrada apenas uma referência (FARIAS, 2004) que trata da geração de artefatos em vídeo digital. Naquela tese foram propostos

geradores de artefatos para o ruído branco, a blocagem e o borramento. Além desses, a autora também trata da geração do artefato *ringing*, que não será discutido nesta dissertação. Todavia, além do ruído branco, blocagem e borramento serão tratados o ruído branco gaussiano, o ruído sal e pimenta e o travamento. Outra diferença em relação à tese de (FARIAS, 2004) é a utilização de várias máscaras para a geração do borramento, enquanto a referência utilizou apenas uma.

Para a geração da blocagem, (FARIAS, 2004) calcula a média aritmética de um bloco de *pixels* 8x8 e de um bloco 24x24 onde o bloco 8x8 é o centro deste bloco maior. O próximo passo é calcular a diferença entre essas duas médias. A seguir este valor de diferença é adicionado em cada bloco menor. Além disso, o valor dessa diferença é limitado para evitar uma saturação nos blocos. Na inserção do artefato borramento a autora utilizou-se apenas da máscara matricial 5x5 que aqui também foi utilizada. Ao gerar o ruído branco, a autora, utilizou-se apenas da taxa de 10% de ocorrência do artefato, e utilizou-se de duas variáveis aleatórias que ditaram a posição horizontal e vertical do *pixel* a ser alterado, para um valor aleatório segundo uma distribuição não descrita na sua tese. E por último para a geração do artefato *ringing* foram utilizados dois filtros um passa - alta e outro passa - baixa combinados.

2.5 Métricas de Qualidade de Vídeo Baseadas em *Pixel* com Referência Total

2.5.1 MSE

O MSE (*Mean Square Error*) ou 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), segundo a equação abaixo:

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [(f(i,j) - F(i,j))]^2}{M \times N} \quad (1)$$

Onde $f(i,j)$ é o valor do *pixel* na imagem original, $F(i,j)$ é o valor do *pixel* na imagem avaliada, M é o número de linhas e N o número de colunas do quadro.

Quando uma imagem é avaliada através desta métrica, 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.

2.5.2 PSNR

O PSNR (*Peak Signal to Noise Ratio*) é medido em uma escala logarítmica e depende do erro quadrático médio (MSE) entre duas imagens, uma original e uma processada, relativo ao valor $(2^n - 1)^2$, que representa o quadrado do maior valor possível de sinal na imagem, onde n é o número de bits de cada imagem. O PSNR é expresso por:

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

O PSNR 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. A Figura 9 mostra um comparativo de três imagens: A imagem (A) é a original, as imagens (B) e (C) são versões degradadas (borradas) da imagem original. A imagem (B) tem um PSNR medido de 30,6 dB enquanto a imagem (C) tem um de 28,3 dB (refletindo uma qualidade pior).

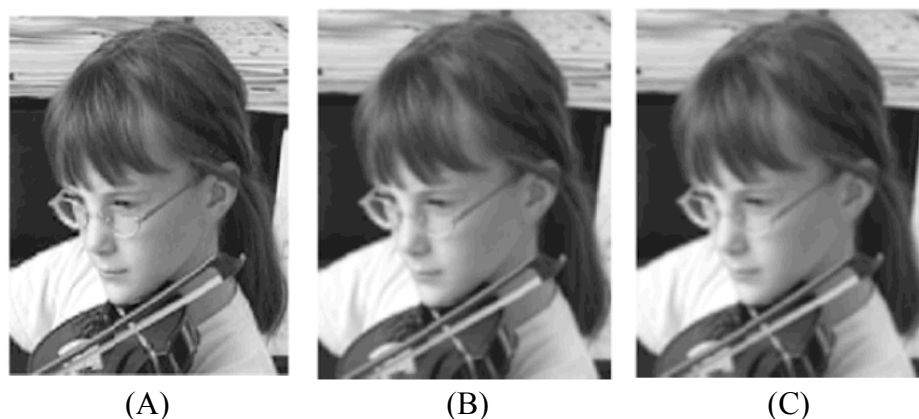


Figura 9. Exemplo de PSNR: (A) Original, (B) 30,6 dB e (C) 28,3 dB (RICHARDSON, 2003).

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, já mencionado, diz respeito a imagem “original”, que pode não apresentar uma boa qualidade. O PSNR não possui uma correlação muito boa com as medidas subjetivas de qualidade de vídeo, aquelas definidas na ITU-R BT.500 (ITU, 2000). 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. Por exemplo, a Figura 10 mostra uma versão distorcida da imagem original da Figura 9(A) onde apenas o fundo da imagem está borrado. Esta imagem possui um PSNR de 27,7 dB relativo à original. A maioria dos observadores avaliaria esta imagem como significativamente melhor que a imagem (C) na Figura 9 porque o rosto está mais nítido, contradizendo o valor PSNR. Este exemplo mostra que o valor PSNR não necessariamente se correlaciona com a qualidade subjetiva “real”. Neste caso, um observador humano dá mais importância à região da face, que é particularmente mais sensível a distorções nesta área.



Figura 10. Imagem com o fundo borrado (PSNR = 27,7 dB) (RICHARDSON, 2003).

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. Muitas aproximações foram propostas, mas nenhuma delas emergiu, ainda, como uma alternativa melhor ou comparável aos testes subjetivos (WU e RAO, 2006), (RICHARDSON, 2003). O Grupo VQEG (*Video Quality Experts Group*) do ITU-T incentiva o desenvolvimento de padrões para avaliação objetiva da qualidade de vídeo (VQEG, 2000). Em março de 2000, o VQEG, na primeira bateria de testes, onde dez sistemas propostos foram testados sobre condições idênticas, reportou que nenhuma delas foi considerada aplicável para padronização (WU e RAO, 2006), (RICHARDSON, 2003). Além disso, o VQEG realizou uma segunda bateria de avaliações em 2003, embora os relatórios e resultados destes testes ainda não estejam publicamente disponíveis (WU e RAO, 2006).

3 METODOLOGIA

O propósito do presente trabalho é medir a qualidade do vídeo a partir da geração controlada de artefatos. Esta geração é importante para o desenvolvimento de novas métricas e de formas de detecção de artefatos. Além disso, ao gerar os artefatos de forma sintética têm-se o controle da intensidade de cada artefato e a certeza de que não estão sendo causadas outras características no vídeo além do desejado, ao contrário do que se obtém por métodos reais de geração de artefatos, como perda de sinal e interferência ou ocorrência de ruído.

3.1 Materiais

Durante o desenvolvimento das atividades foram utilizados alguns software importantes para o controle, observação e análise dos resultados obtidos, sendo estes: YUVviewer, MatLab, Origin, GLYUV, GNUPlot, Gcc. Foram ainda utilizadas fontes de vídeo 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 sofrer nenhuma interferência externa, além de serem utilizadas comumente na literatura e terem características importantes, como a facilidade na leitura e manipulação dos pixels (AUWERA, DAVID, REISSLEIN, 2008). Além disso estão disponíveis em (MMLAB, 2009).

O YUVViewer (para plataformas Windows®) e o GLYUV (para plataformas MAC OS X®) são programas (*players*) úteis na observação e monitoração da imagem obtida após a geração de algum artefato. O funcionamento dessas ferramentas consiste em ler o arquivo de vídeo *byte a byte* e exibir uma imagem com as suas componentes de luminância e croma para cada quadro, seguido da exibição do quadro seguinte em forma de uma sobreposição de imagens (quadros), obedecendo à taxa de atualização do vídeo *e.g.* 30 quadros por segundo. Isto cria a ilusão de movimentação para o HVS. O Origin e o GNUPlot foram utilizados para desenhar os gráficos dos resultados obtidos.

O software MatLab foi utilizado para a geração de histograma do gerador de variáveis aleatórias e o Gcc é um compilador da linguagem ANSI C, que foi utilizado para compilar os algoritmos que serão aqui descritos.

Além disso, foram eleitos alguns traços de vídeo visando abranger a maior diversidade de características, tais como: movimentação, textura e coloração. Os traços escolhidos foram: “*Mobile*”, “*Akiyo*”, “*Stefan*” e “*Coastguard*”, os quais pertencem ao rol dos vídeos comumente utilizados em testes de ferramentas que envolvem este tipo de informação (Video Traces Research Groups, 2008). Todos estes vídeos possuem duração de 10 segundos, 300 quadros e uma taxa de 30 quadros por segundo. A escolha de vídeos que possuem estas características se deu devido à utilização destes na avaliação subjetiva, a qual, segundo (ITU, 2000), não deve exceder 30 minutos.

Ainda foram utilizados alguns computadores para realização dos experimentos, sendo eles: um PowerBook G4 1.5GHz, com 1.25GB de memória RAM, com Mac OS X 10.4.11. e um Intel Pentium Core 2 Quad 2.4GHz com 4 GB de RAM e disco rígido de 350 GB, com uma placa de vídeo GeForce 7300SE com 256MB de memória, com Windows XP. O primeiro dos computadores foi utilizado para fazer a geração de todos os artefatos e o cálculo do PSNR e MSE dos vídeos gerados com os algoritmos desenvolvidos. E o segundo foi utilizado para realização das validações e para exibição dos vídeos aos espectadores dos testes subjetivos. Além disso foi utilizada uma televisão LG Full-HD (1080p) de 42 polegadas para a exibição dos vídeos na avaliação subjetiva. O uso dos dois computadores se fez necessário, devido ao *player*, utilizado na execução dos vídeos durante os testes subjetivos, ser desenvolvido para plataforma Windows® e pelo fato de se poder utilizar uma entrada da TV com interface DVI-D (com resolução de alta definição com 1080 linhas). Na Figura 11 é possível ver a disposição dos equipamentos e como eles estão interligados. Em um primeiro momento, os vídeos foram gerados no computador PowerBook, em seguida eles foram copiados para o computador Intel Pentium Core 2 Quad, o qual exibiu os vídeos na TV LG de 42”, utilizando o *player* YUVViewer.

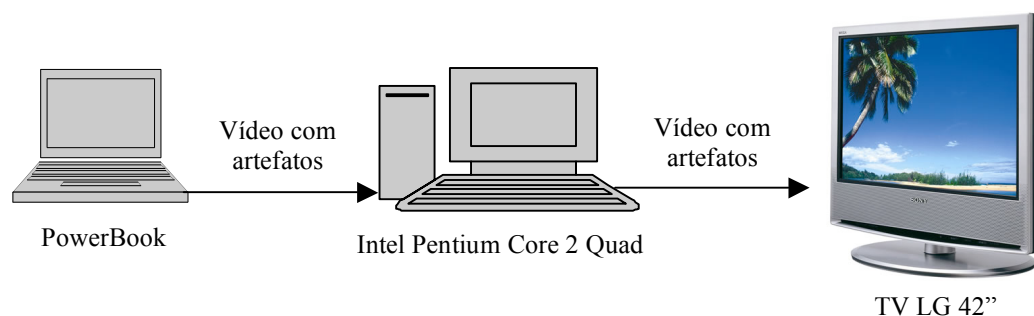


Figura 11. Desenho esquemático da disposição dos equipamentos.

3.2 Formato dos Arquivos YUV

No formato YUV cada quadro é composto por suas componentes Y, U e V. As componentes são representadas pelo valor dos bytes do quadro convertidos em decimal (cada *pixel* corresponde a 1 byte). No caso da representação de um *pixel* com 8 bits, seu valor estará entre 0 e 255. Para um vídeo com resolução de 352x288 *pixels*, o número total de bytes para representação da componente de luminância (Y) é igual a 101376 bytes. As componentes U e V dos quadros de um vídeo com taxa de amostragem 4:2:0 são representados com a metade dos bytes da componente Y, ou seja, 50688 bytes para a componente U e 50688 bytes para a componente V.

Os primeiros bytes representam os valores de luminância (componente Y) para o primeiro quadro, os bytes seguintes são responsáveis pelas componentes de croma (U e V). Na sequência são encontrados os bytes referentes à luminância para o segundo quadro, seguidos dos valores de croma (U e V) para o segundo quadro, e, assim, sucessivamente. Na Figura 12 encontra-se uma representação esquemática ilustrando a estrutura dos arquivos YUV utilizados.

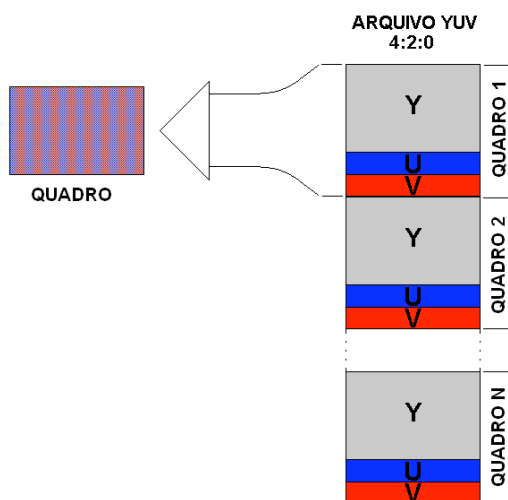


Figura 12. Estrutura dos arquivos YUV.

3.3 Gerador de Artefatos

Os algoritmos foram desenvolvidos utilizando a linguagem de programação ANSI C e tendo como meta modelar os ruídos e artefatos conforme o que é observado pelo HVS na televisão analógica e digital atuais.

Para todos os artefatos, o funcionamento dos algoritmos de geração segue os passos abaixo:

- 1- Abre o arquivo YUV original;
- 2- Calcula o número de quadros baseado no tamanho do arquivo em bytes e na resolução (para verificar a integridade dos dados);
- 3- Lê um quadro inteiro;
- 4- Percorre os *pixels* alterando-os em função do artefato a ser inserido;
- 5- Cria o arquivo YUV, se ele não existe, e grava o quadro alterado no novo arquivo;
- 6- Repete os passos 3, 4 e 5 até chegar ao fim do arquivo;
- 7- Fecha o arquivo YUV original e libera as variáveis da memória.

A Figura 13 apresenta o fluxograma do algoritmo descrito.

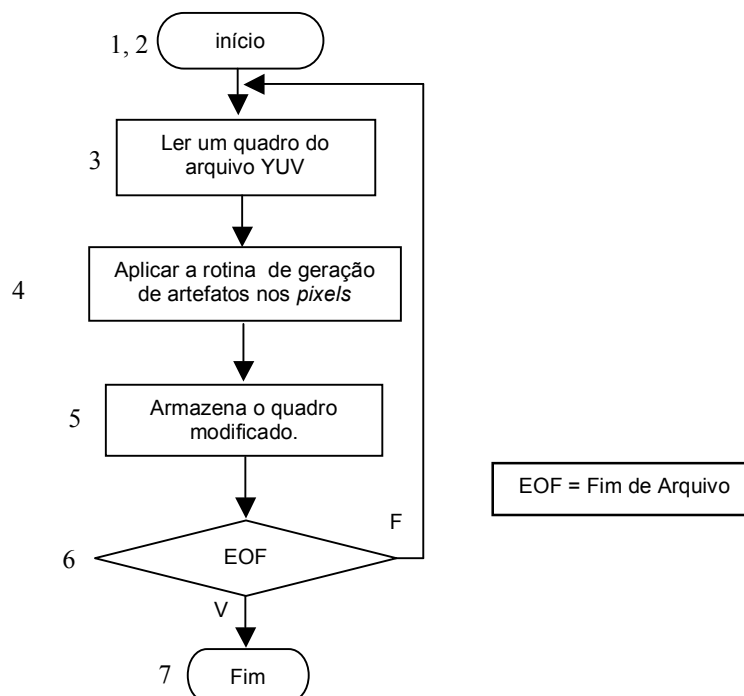


Figura 13. Fluxograma do funcionamento básico dos algoritmos.

O gerador de artefatos aqui descrito manipula apenas os valores de luminância dos *pixels*, visto que estes possuem o maior impacto sobre a imagem e o PSNR e MSE calculados levam em conta apenas essa componente. O que irá diferenciar qual artefato está sendo embutido no vídeo é a alteração a que serão submetidos os *pixels*. Esta mudança é definida no passo 4 (quatro) e descrita pelas equações para geração dos respectivos artefatos, que serão descritas a seguir.

3.3.1 Ruído Branco Gaussiano

O ruído branco gaussiano possui uma distribuição uniforme quanto ao local de ocorrência, ou seja, a probabilidade de um *pixel* ser afetado pelo ruído branco gaussiano independe de sua posição espacial, e apresenta uma distribuição gaussiana quanto ao valor da amplitude do ruído que é adicionado aos *pixels*. A geração deste artefato utilizou um gerador de variáveis aleatórias com distribuição uniforme, o qual seleciona aleatoriamente os *pixels* que sofrerão alterações. A

taxa de probabilidade de alteração é um parâmetro controlável. Quando um *pixel* é selecionado, este possui seu valor de luminância alterado, segundo uma variável aleatória de distribuição gaussiana com média e desvio padrão alteráveis no algoritmo. Além disso, o valor de luminância é truncado em 255 (codificação com 8 bits) se o valor do *pixel* mais o valor da variável aleatória for maior que 255, e em 0, caso o valor dessa soma seja negativo. Isso se repete para todos os *pixels* de cada quadro, e para todos os quadros do arquivo.

O algoritmo gerador do artefato de ruído branco gaussiano é descrito segundo a equação.

$$\begin{cases} Y[i, j] = O[i, j] + R & \text{para } X < P \\ Y[i, j] = O[i, j] & X \geq P \end{cases} \quad (3)$$

Onde Y representa o quadro modificado, O representa o quadro original, i e j representam a posição espacial do *pixel* no quadro. P é a probabilidade de um *pixel* ser escolhido para sofrer a adição do ruído Gaussiano, X é a variável aleatória (com distribuição uniforme) que determina se um *pixel* será afetado e R é o ruído branco gaussiano gerado por uma variável aleatória com distribuição Gaussiana, média zero e desvio padrão quarenta, valores estes também utilizados por outros trabalhos (TANG e CAHILL, 1991).

Na linguagem de programação ANSI C a função geradora de variáveis aleatórias (*rand()*) possui uma distribuição uniforme. Como durante a geração do ruído branco gaussiano fez-se necessário um gerador de variáveis aleatórias com distribuição gaussiana, foi desenvolvida uma forma para a criação de um gerador de variáveis aleatórias com distribuição gaussiana a partir de um gerador de variáveis aleatórias com distribuição uniforme. Com isto, em (ABDURAHIMAN, 2009) foi encontrada a seguinte equação:

$$VAG = \left(\sum_{1}^{12} VAU - 6 \right) * \sigma + \mu \quad (4)$$

Onde VAG é a variável aleatória com distribuição gaussiana, VAU são variáveis aleatórias com distribuição uniforme, σ é o desvio padrão e μ a média da distribuição gaussiana.

Foram efetuados testes para validação desta equação e comprovar sua natureza na representação da distribuição. Para validar a equação foram gerados cem mil (100.000) valores aleatórios, utilizando média zero (0) e desvio padrão quarenta (40) e foi montado um histograma, utilizando o programa MatLab. O resultado do histograma é exibido na Figura 14. Além disso, esta validação foi executada várias vezes com valores de média e desvio padrão diferentes para comprovar o seu comportamento.

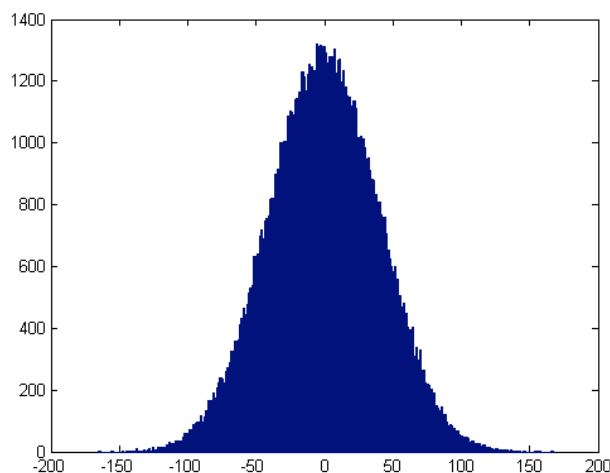


Figura 14. Histograma do gerador de variáveis aleatórias com distribuição gaussiana, com média 0 e desvio padrão 40.

Como é possível observar na Figura 14, o gerador de variáveis aleatórias reflete uma distribuição gaussiana e pode ser utilizado no gerador do artefato proposto.

Através da equação (3) é possível obter o valor dos *pixels* de um vídeo degradado pelo gerador do artefato WGN. Um exemplo da aplicação desse algoritmo é exibido na Figura 15, com 2% (A) e 15% (B) dos *pixels* afetados por este artefato. Com 15% nota-se uma maior incidência deste artefato na imagem.



Figura 15. Quadro do vídeo “Akiyo” onde (A) os *pixels* possuem 2% de probabilidade de ocorrência e (B) os *pixels* possuem 15% de probabilidade de ocorrência.

3.3.2 Sal e Pimenta

Para a geração do artefato Sal e Pimenta foram utilizados dois geradores de variáveis aleatórias com distribuição uniforme, um para selecionar os *pixels* que serão alterados e outro para escolher se serão alterados para o valor máximo (255) ou mínimo de luminância (0). Este procedimento se repete para todos os *pixels* de cada quadro e para todos os quadros do vídeo, sendo que a probabilidade de um *pixel* ser sorteado é controlável.

O procedimento para gerar o artefato é representado pela equação.

$$\begin{cases} Y[i,j] = 255 * (X.MOD.2) & \text{para } R < P \\ Y[i,j] = O[i,j] & R \geq P \end{cases} \quad (5)$$

Y é o quadro gerado, i e j representam a posição espacial do *pixel*, X e R são variáveis aleatórias com distribuição uniforme, P é a porcentagem de *pixels* que serão afetados pelo artefato e O é o quadro original. O operador “*MOD.*” é o operador que retorna o resto da divisão. O dividendo é a variável X e o divisor é o valor 2 (*i.e.* caso X seja um valor ímpar, a operação retorna o valor 1, caso contrário retorna 0). Um exemplo do artefato gerado através desse procedimento é mostrado na Figura 16.



Figura 16. Quadro do vídeo “*Mobile*” com artefato sal e pimenta, onde em (A) a probabilidade de cada *pixel* ser selecionado é 1% e em (B) a probabilidade de cada *pixel* ser selecionado é 5%.

Da Figura 16 nota-se que, qualitativamente, o artefato se assemelha ao gerado pelo WGN, com a ressalva de que pontos claros e escuros são notados na imagem.

3.3.3 Borramento

Para se entender o funcionamento do borramento é necessário um conceito importante que diz respeito ao filtro da média, o qual é utilizado em processamento digital de imagens. Este filtro é utilizado para ocasionar a interferência entre os *pixels*, simulando assim o borramento. Seu funcionamento consiste em uma matriz (máscara) que define quais serão os *pixels* envolvidos no cálculo da média. Este valor de média é atribuído como luminância ao *pixel* central. Assim, é realizada uma convolução dessa matriz no quadro inteiro (MARQUES FILHO e VIEIRA NETO, 1999). Esta convolução se dá da seguinte forma: para uma máscara de 3 *pixels* linear, primeiro são tomados os *pixels* 1, 2 e 3 da imagem, para o cálculo do valor do *pixel* 2, em seguida são pegos os *pixels* 2, 3 e 4, para o cálculo do valor do *pixel* 3, e assim sucessivamente. As equações apresentadas na sequência explicitam as máscaras utilizadas, onde o elemento central da matriz é o elemento a receber a média, e os outros elementos com o valor 1 são os *pixels* considerados para o cálculo da média.

Quanto maior a máscara, maior o borramento obtido. Isto se deve ao fato de que mais *pixels* estarão interferindo entre si. Na geração do borramento, aqui, são utilizadas máscaras lineares e matriciais, cada uma delas com tamanhos distintos. Foram empregadas três abordagens para as máscaras lineares e três para as matriciais. Neste trabalho foram utilizadas apenas máscaras lineares horizontais, porém é possível obter outras formas de borramento com outros tipos de máscaras, como por exemplo máscaras lineares verticais.

Para a primeira abordagem com máscara linear, foi utilizada uma máscara de média com três *pixels*, ou seja, cada *pixel* sofre a interferência dos dois *pixels* adjacentes, sendo estes, o da direita e o da esquerda.

$$M = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad (6)$$

Na outra abordagem linear, foi utilizada a máscara média com cinco *pixels*, ou seja, cada *pixel* sofre a interferência dos dois *pixels* da direita e dos dois *pixels* da esquerda.

$$M = \frac{1}{5} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (7)$$

Em uma terceira aplicação foi utilizada a máscara linear média com sete *pixels*, onde o *pixel* central recebe o valor da interferência dos três *pixels* anteriores e três posteriores.

$$M = \frac{1}{7} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (8)$$

O valor resultante é a média aritmética desses *pixels*. A descrição matemática que rege a alteração sofrida nos *pixels* ao executar o algoritmo gerador desse artefato, utilizando as máscaras lineares, obedece à equação

$$Y[i, j] = \frac{1}{n} \left(\sum_{k=i-\frac{(n-1)}{2}}^{i+\frac{(n-1)}{2}} O[k, j] \right) \quad (9)$$

Onde Y é o quadro gerado com o artefato borramento, n é o tamanho da máscara linear utilizada (e.g. para máscara média de 5 *pixels*, $n = 5$). O é o quadro original, i e j denotam a posição espacial do *pixel* no quadro. Com esta equação pode-se estimar o valor de cada *pixel* (gerado pela máscara linear). Um exemplo da aplicação da máscara com filtro de média linear com 5 *pixels* pode ser vista na Figura 17(A).

Com relação à abordagem com matrizes, a primeira técnica utiliza a máscara matricial 3x3 (10).

$$M = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (10)$$

Ou seja, cada *pixel* sofre a interferência de todos os *pixels* adjacentes a ele. Na sequência, foi utilizada a máscara matricial 5x5 (11) seguindo o raciocínio da média da matriz 3x3.

$$M = \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (11)$$

O último caso utiliza a máscara matricial 7x7 (12).

$$M = \frac{1}{49} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (12)$$

Assim, pode-se gerar o artefato utilizando as máscaras matriciais através da equação

$$Y[i,j] = \frac{1}{n^2} \left(\sum_{k=i-\frac{(n-1)}{2}}^{i+\frac{(n-1)}{2}} \sum_{l=j-\frac{(n-1)}{2}}^{j+\frac{(n-1)}{2}} O[k,l] \right) \quad (13)$$

Onde Y é o quadro gerado com o artefato borramento, n é o tamanho da máscara matricial utilizada (e.g. para máscara 5x5, $n = 5$). O é o quadro original e i e j denotam a posição espacial do *pixel* no quadro. Um exemplo da utilização da máscara 7x7 é exibido na Figura 17 (B).

3.3.4 Blocagem

Para a geração do artefato blocagem foi utilizado um gerador de números aleatórios com distribuição uniforme para a seleção dos *pixels* que serão alterados, outro para selecionar a duração do artefato e um último que delimita em qual quadro se realizará o início do artefato. Quando um *pixel* é selecionado este tem o seu valor copiado aos *pixels* adjacentes numa matriz 8x8. A probabilidade de um *pixel* ser selecionado e a duração máxima do artefato (em quadros) são parâmetros de entrada controláveis do algoritmo. O tamanho 8x8 foi escolhido devido ao tamanho dos macro-blocos encontrados nos codificadores MPEG-2 e H.264 (HANZO, CHERRIMAN e STREIT, 2007), (LIU, BOVIK, 2002), (RICHARDSON, 2003) e (STOLFI, 2007). Um exemplo de artefato gerado pela aplicação deste procedimento é visualizado na Figura 18.

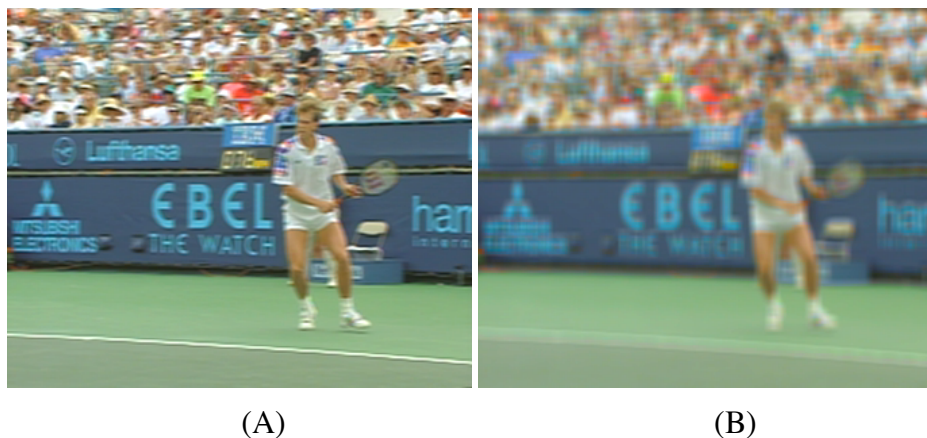


Figura 17. Quadro do vídeo “*Stefan*” com a aplicação do artefato borramento em (A) utilizando uma máscara linear de 5 *pixels* e (B) utilizando uma máscara matricial 7x7.

3.3.5 Travamento

Na geração do travamento foram utilizados dois geradores de variáveis aleatórias com distribuição uniforme. Um é responsável por selecionar o quadro inicial que receberá o travamento e o outro pela escolha da duração do travamento que pode ser de até trinta (30) quadros, equivalente à duração de 1 segundo.



Figura 18. Quadro do vídeo “*Coastguard*” com a aplicação do algoritmo gerador de *blocking*, com ocorrência de 5% por *pixel*.

Quando um vídeo é submetido ao efeito de travamento, o quadro inicial do travamento (que é escolhido através da aplicação da variável aleatória) é copiado para os quadros subsequentes até a duração (que também é selecionada aleatoriamente) ser atingida.

A equação abaixo descreve o procedimento adotado.

$$\begin{cases} Y[n] = Y[n-1] & \text{para } R \leq n \leq R + P \\ Y[n] = O[n] & \text{caso contrário} \end{cases} \quad (14)$$

$Y[n]$ é o quadro do vídeo gerado (com travamento), $Y[n-1]$ é o quadro anterior do vídeo gerado, $O[n]$ é um quadro do vídeo original, n é o número do quadro, R é a variável aleatória para o início do travamento e P é a variável aleatória da duração do travamento (*i.e.* até 30 quadros).

3.4 MSU *Video Quality Measurement Tool*

O MSU *Video Quality Measurement Tool* [Graphics and Media Lab, 2009] é uma ferramenta livre (gratuita), porém de código fonte fechado, que implementa várias métricas com referência total em vídeos e imagens, sendo elas: PSNR, MSAD, Delta, *Blurring Measure*, *Blocking Measure*, SSIM index, VQM e MSE. A métrica MSAD, em linhas gerais, efetua a diferença absoluta entre os *pixels* coloridos de uma imagem e é utilizada para testar codificadores e filtros. O valor da métrica Delta é a diferença média entre as componentes de cor dos *pixels* correspondentes da imagem. A *Blurring Measure* permite comparar a quantidade de borramento existente entre duas imagens. Se o valor dessa métrica para a primeira imagem for maior que o valor para a segunda, isso significa que a segunda imagem é mais borrada que a primeira. Já a *Blocking Metric* foi criada para medir o efeito subjetivo do artefato blocagem numa sequência de vídeo, utilizando informações dos quadros anteriores para ter uma maior precisão. E por último a VQM utiliza-se da DCT para corresponder à percepção visual humana (XIAO, 2000), (MSU Quality Measurement Tool, 2009).

Esta ferramenta (MSU *Video Quality Measurement Tool*) foi utilizada como gabarito para realizar a validação do algoritmo desenvolvido para o cálculo do MSE e PSNR.

3.5 PSNR

Durante o desenvolvimento do algoritmo para cálculo do PSNR e MSE foi utilizada a linguagem de programação ANSI C. As razões da utilização desta linguagem são: a facilidade da adaptação deste algoritmo para sistemas embarcados e uma maior velocidade na sua execução, ao contrário do que acontece com linguagens que necessitam de máquinas virtuais para a sua execução, como o JAVA e o .NET.

Para o cálculo do PSNR é necessário antes calcular o MSE. Portanto, foi desenvolvido o algoritmo para cálculo do MSE em conjunto com o PSNR. Após a sua implementação, foram realizados testes comparativos com o software proprietário MSU, para validação da ferramenta desenvolvida.

3.5.1 Validação do PSNR

Para realizar a validação do algoritmo desenvolvido que calcula o PSNR e o MSE para vídeos, foi utilizada a ferramenta MSU. Foram gerados oito (8) vídeos degradados utilizando o gerador de artefatos desenvolvido. Em seguida, foram calculados o MSE e o PSNR desses vídeos, com base em seus originais, utilizando o MSU e o algoritmo do PSNR e MSE desenvolvido. Com os resultados foram realizadas comparações dos valores calculados pelos dois programas, na Tabela 3 (onde são exibidos apenas os valores dos 20 primeiros quadros de um dos vídeos utilizados para a validação). Observa-se a inexistência de diferença entre os resultados. Na tabela são exibidos apenas os valores de PSNR, consequentemente os valores de MSE também são válidos, por estes fazerem parte do cálculo do PSNR.

Tabela 3.Comparação de resultados para o valor de PSNR entre o algoritmo desenvolvido e o software MSU.

Quadro	MSU (dB)	Algoritmo desenvolvido (dB)	Erro
0	33,03	33,03	0
1	33,03	33,03	0
2	33,02	33,02	0
3	33,01	33,01	0
4	33,05	33,05	0
5	33,10	33,10	0
6	33,10	33,10	0
7	33,09	33,09	0
8	33,03	33,03	0
9	33,01	33,01	0
10	33,01	33,01	0
11	33,01	33,01	0
12	33,00	33,00	0
13	33,00	33,00	0
14	32,98	32,98	0
15	32,95	32,95	0
16	32,95	32,95	0
17	32,98	32,98	0
18	32,98	32,98	0
19	32,95	32,95	0
20	32,94	32,94	0

3.6 Comparação com Medidas Subjetivas

Em (SILVA, 2009) foram utilizados vídeos cuja imagem original era de boa qualidade e foram inseridos os artefatos aqui propostos para degradação do vídeo a ser avaliado. Assim, foi possível observar que existe uma correlação da MOS, obtida pela avaliação subjetiva, e dos valores de PSNR e MSE, obtidos através das métricas objetivas aqui utilizadas. Com isso é possível notar que existe um intervalo de validade para as métricas *full-reference* que foram descritas.

Para a avaliação subjetiva foi utilizado o método de exibição dos vídeos chamado SDSCE (*Single Double Stimulus for Continuous Evaluation*) – duplo estímulo simultâneo para avaliação contínua no qual, foram exibidos, para todos os vídeos, o original (sem degradação) à esquerda e

o degradado à direita. A seguir, os avaliadores assinalavam as notas em uma escala de 1 (um) a 5 (cinco), sendo 1 atribuído a uma imagem com qualidade excelente, 2 para uma imagem de boa qualidade, 3 para regular, 4 para pobre e 5 para uma qualidade ruim.

Foram efetuadas avaliações subjetivas utilizando 6 (seis) vídeos com o artefato borramento, 2 (dois) com blocagem, 5 (cinco) com sal e pimenta, 10 (dez) com ruído branco gaussiano e 1 com o artefato travamento. Todos estes artefatos foram gerados a partir dos algoritmos descritos aqui, sendo que cada um dos vídeos do borramento foi gerado com uma das máscaras apresentadas (um vídeo com cada máscara). Para a blocagem foram gerados vídeos com 1 e 5% de ocorrência e para o sal e pimenta foram gerados vídeos com 1, 2, 3, 4 e 5% de ocorrência. O travamento foi gerado como descrito anteriormente, aleatoriamente, e para o ruído branco gaussiano foram gerados com 2, 4, 6, 8, 10, 15, 20, 30, 50 e 100% de ocorrência.

Durante os testes 33 observadores submeteram-se aos testes. Cada sessão de exibição dos vídeos para a avaliação era composta por 3 (três) observadores. Os vídeos eram exibidos de forma aleatória, porém sempre o original à esquerda e o avaliado à direita (os observadores eram avisados quanto a isso).

A partir da análise dos resultados foram traçados gráficos de comparação entre as métricas objetivas e subjetivas. Estes gráficos podem ser observados na Figura 19 e Figura 20. Na Figura 19 foram traçados os gráficos de correlação da MOS (subjetiva) pelo PSNR (objetivo) e uma curva “*fit*” para cada tipo de artefato (intervalo de confiança de 95%). Neste gráfico, através da curva “*fit*” é possível observar que para os artefatos borramento, ruído sal e pimenta e ruído branco gaussiano existe uma correlação notável entre as duas métricas, com relação ao artefato travamento esta correlação pode ser efetuada, apesar do número de amostras utilizadas ter sido muito pequeno.

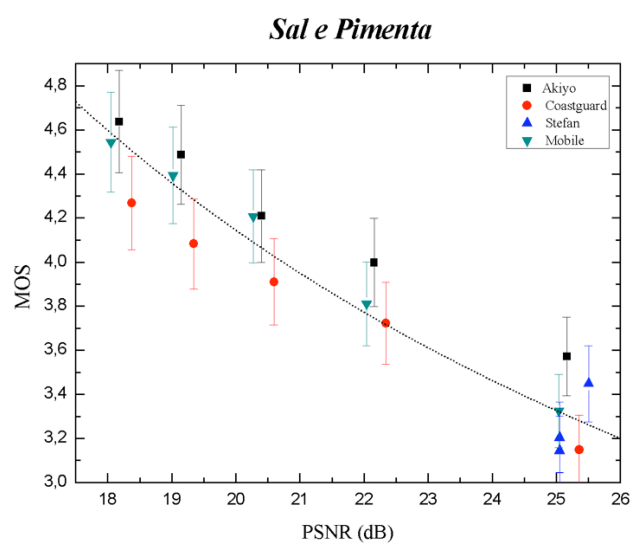
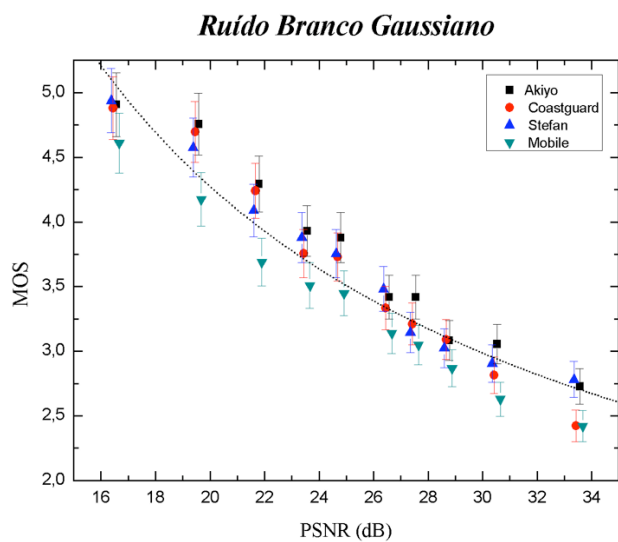
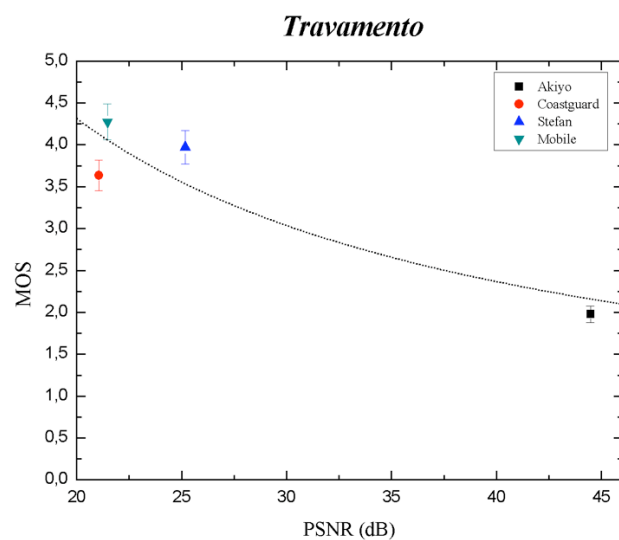
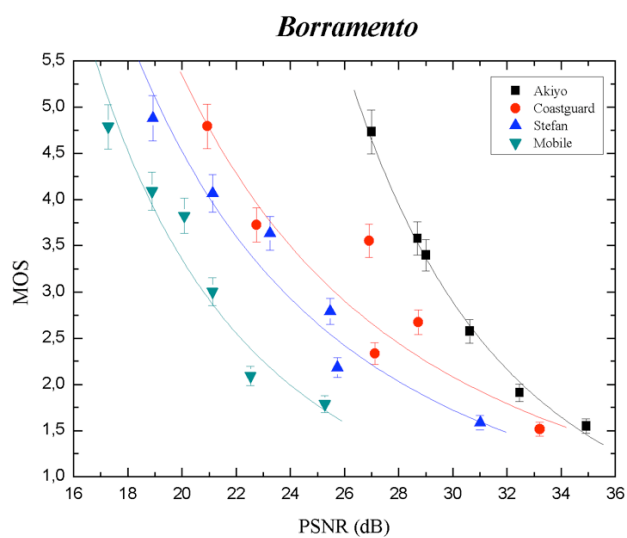


Figura 19. Gráficos de correlação da MOS x PSNR para alguns dos artefatos gerados e avaliados (SILVA, 2009).

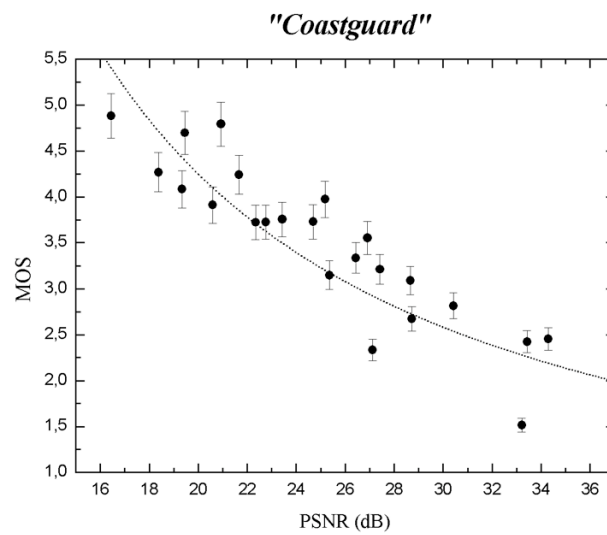


Figura 20. Gráfico da comparação da métrica subjetiva (MOS) com a métrica objetiva (PSNR) para todos os artefatos de um traço de vídeo (SILVA, 2009).

4 RESULTADOS

Com a execução do algoritmo gerador de artefatos foi possível obter degradações no vídeo próximas às observadas em situações reais (*i.e.* provocadas por mecanismos de compressão ou transmissão). Nesta seção, são apresentados os resultados obtidos utilizando-se apenas o vídeo “*coastguard*”. Porém, como dito anteriormente, foram utilizados adicionalmente três outros vídeos com características diferentes. Para fins de facilidade na interpretação dos resultados, são exibidos para todos os artefatos o gráfico do MSE, o gráfico do PSNR e a imagem de um dos quadros que apresenta o artefato gerado.

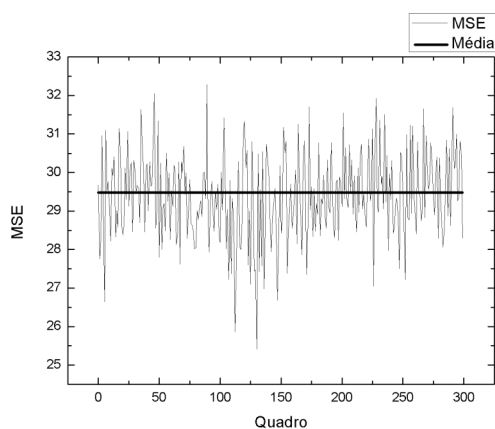
4.1 Ruído Branco Gaussiano

Com a execução desse experimento é esperado que com o aumento da incidência de ruído, o MSE aumente e o PSNR diminua.

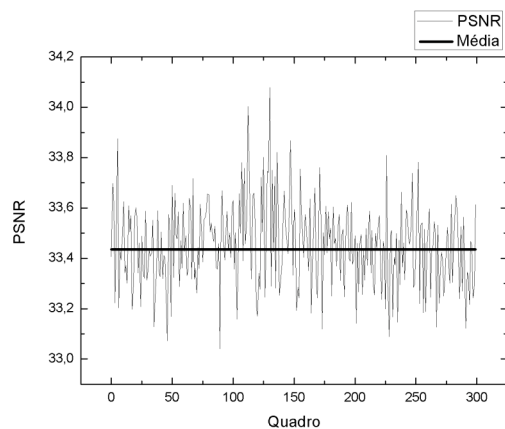
Para 2%, 50% e 100% de taxa de ocorrência de ruído branco gaussiano, a Figura 21 (A) e (B), Figura 22 (A) e (B) e Figura 23 (A) e (B) representam o comportamento do MSE e PSNR, respectivamente, em função do número de quadros do vídeo. Observa-se que quanto maior o MSE, maior a diferença entre o quadro original e o quadro avaliado. Ao contrário do MSE, um PSNR maior significa uma maior fidelidade com o quadro original.

Ao realizar a análise dos gráficos do ruído branco gaussiano é possível perceber um comportamento aleatório do MSE e do PSNR (no decorrer dos quadros), o que é naturalmente explicado através da natureza deste artefato.

Analisando a média do MSE para cada uma das quantidades de ruído inserido, nota-se que um aumento de 2% na incidência de ruído no vídeo, implica em um aumento de 29,49 no valor do MSE e acarreta também uma diminuição no valor do PSNR (pois este é calculado levando o MSE em consideração). Este resultado comportou-se conforme o esperado para este artefato.

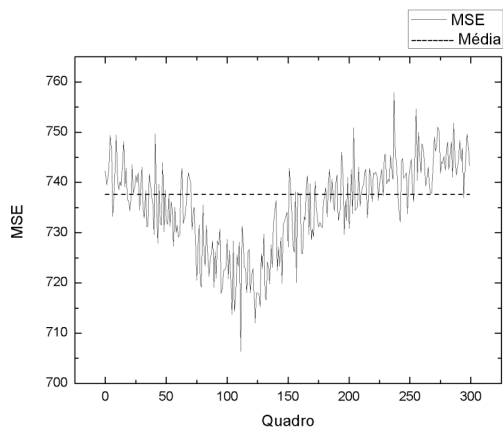


(A)

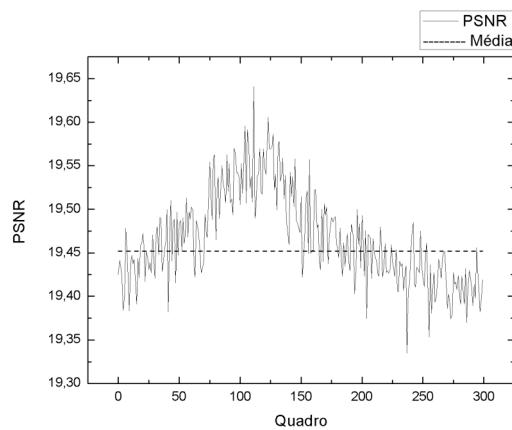


(B)

Figura 21. MSE (A) e PSNR (B) para 2% de ruído branco gaussiano.



(A)



(B)

Figura 22. MSE (A) e PSNR (B) para 50% de ruído branco gaussiano.

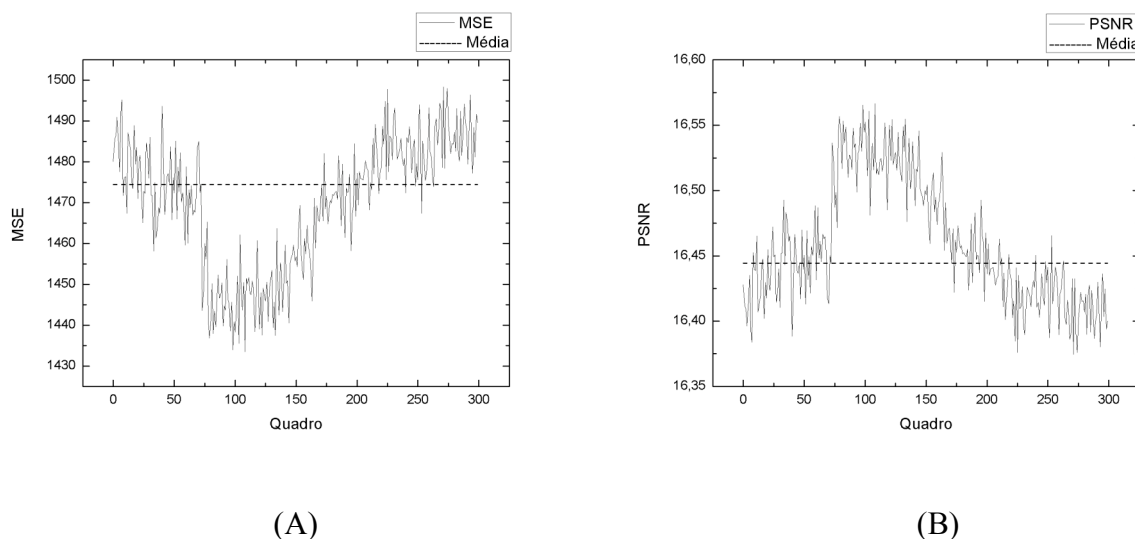
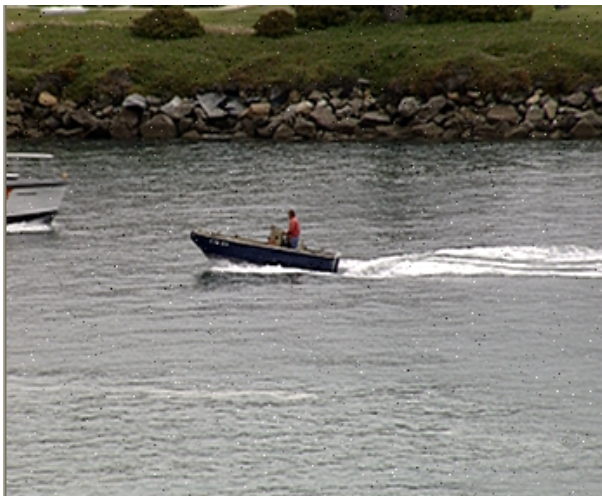


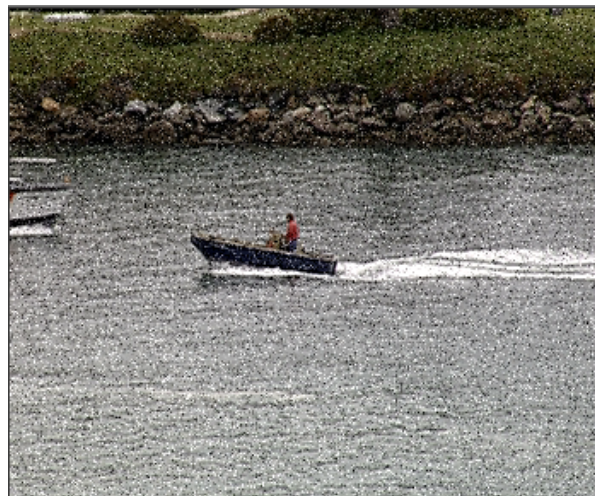
Figura 23. MSE (A) e PSNR (B) para 100% de ruído branco gaussiano.

A forma com que se manifesta visualmente o ruído branco gaussiano pode ser observada na Figura 24 (A), (B) e (C). Nota-se que a imagem mesmo com 100% dos *pixels* afetados pelo ruído branco gaussiano ainda é visualizável, porém ela torna-se extremamente irritante ao espectador devido à movimentação dos pontos claros e escuros no decorrer do tempo.

Ao observar a Figura 24 (A), pode-se fazer um paralelo com o resultado obtido em (SILVA, 2009) exibido na Figura 19 (C), onde o ruído branco gaussiano em pequenas proporções é tolerável pelo espectador, porém conforme a sua incidência aumenta, este torna-se extremamente incômodo.



(A)



(B)



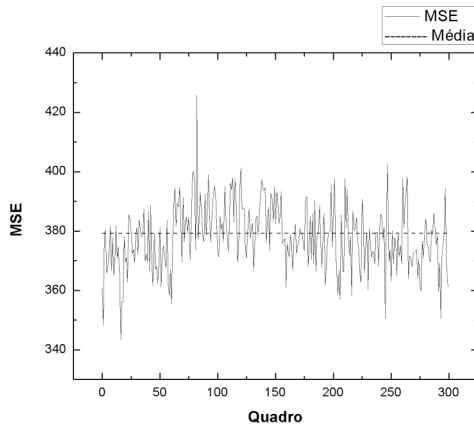
(C)

Figura 24. Quadro do vídeo “*coastguard*” apresentando o ruído branco gaussiano, com ocorrência de 2% (A), 50% (B) e 100% (C).

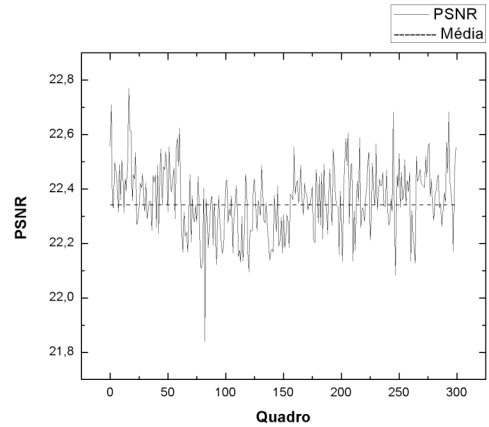
4.2 Sal e pimenta

Ao executar os experimentos envolvendo o ruído sal e pimenta, espera-se que conforme a incidência desse ruído seja maior, o erro quadrático médio (MSE) também seja maior, dado a diferença entre os quadros do vídeo original e do degradado.

Para 2 e 5% de taxa de ocorrência do artefato sal e pimenta, a Figura 25 (A) e (B) e a Figura 26 (A) e (B) representam o comportamento do MSE e PSNR, respectivamente, em função do número de quadros do vídeo.

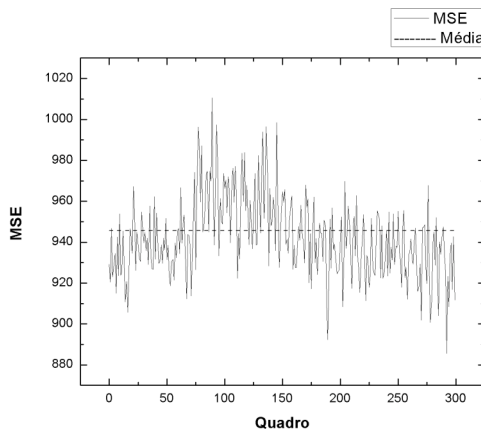


(A)

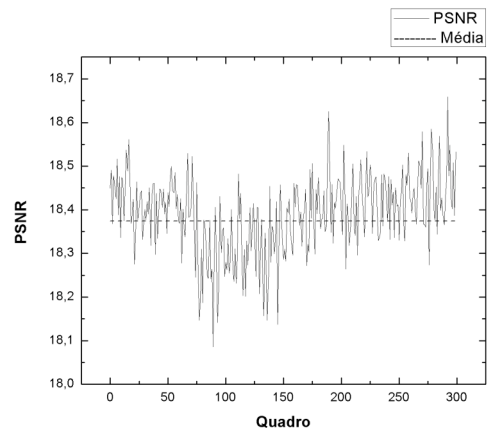


(B)

Figura 25. MSE (A) e PSNR (B) para 2% de ruído sal e pimenta.



(A)

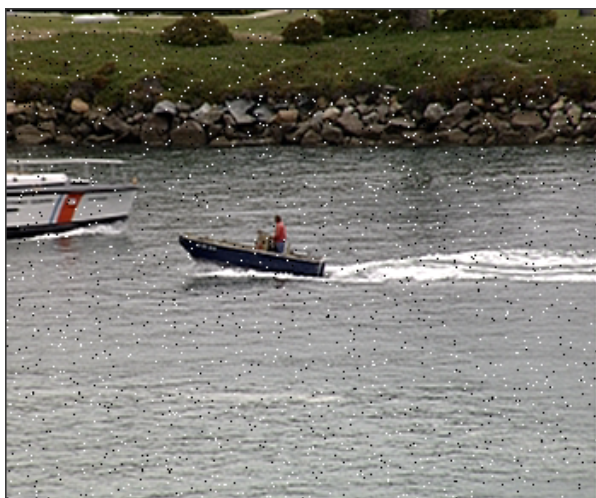


(B)

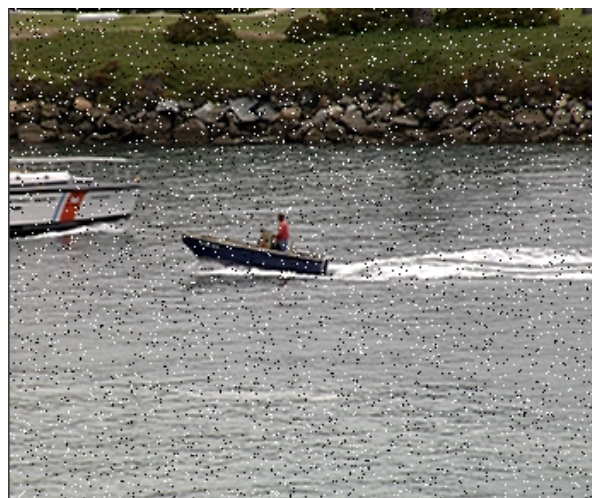
Figura 26. MSE (A) e PSNR (B) para 5% de ruído sal e pimenta.

Analisando os gráficos traçados com o artefato sal e pimenta é perceptível a semelhança com aqueles obtidos para o ruído branco gaussiano. Esta semelhança está relacionada com a natureza aleatória de ambos. Além disso, o impacto visual causado pelos dois artefatos ao observador é análogo, isto pode ser comprovado através da Figura 27 (A) e (B).

Ao observar a média do MSE para cada uma das incidências geradas (1, 2, 3, 4 e 5%), percebe-se que o aumento de 1% na incidência deste artefato, acarreta um aumento de $189,05 \pm 0,68$ no valor médio do MSE. Isto era esperado, dado o aumento na diferença entre os quadros original e degradado.



(A)



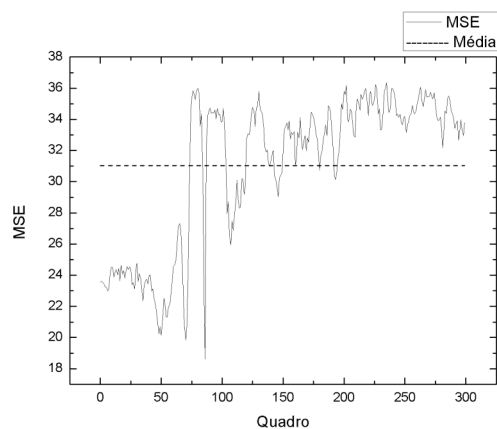
(B)

Figura 27. Quadro do vídeo “*coastguard*” apresentando ruído sal e pimenta, com ocorrência de 2% (A) e com 5% (B).

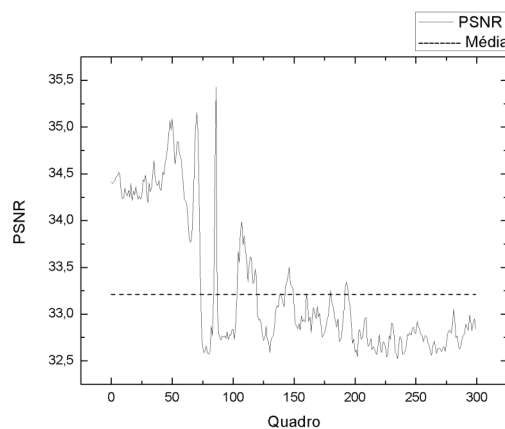
4.3 Borramento

Antes de efetuar os testes para o borramento espera-se que quanto maior o número de *pixels* envolvidos no cálculo da média, maior seja o MSE e menor seja o PSNR, acarretando em uma pior qualidade visual.

Os gráficos para o MSE e PSNR por número de quadros obtidos para o artefato borramento com as máscaras lineares de 3, 5 e 7 *pixels* são exibidos na Figura 28 (A) e (B), Figura 29 (A) e (B) e Figura 30 (A) e (B).

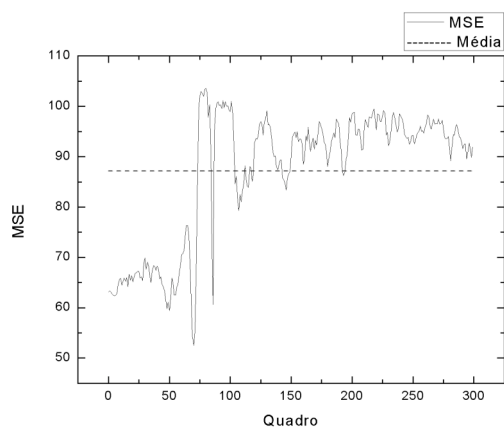


(A)

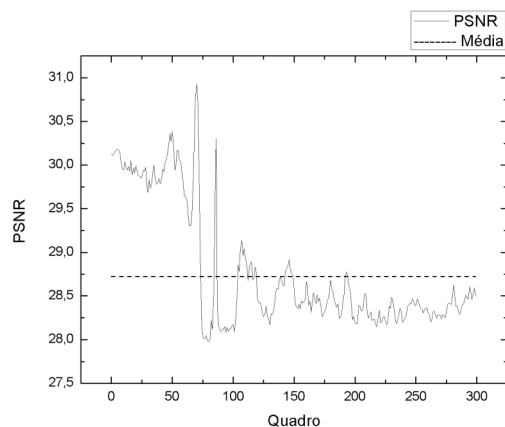


(B)

Figura 28. MSE (A) e PSNR (B) para o borramento utilizando a máscara linear de 3 *pixels*.



(A)



(B)

Figura 29. MSE (A) e PSNR (B) para o borramento utilizando a máscara linear de 5 *pixels*.

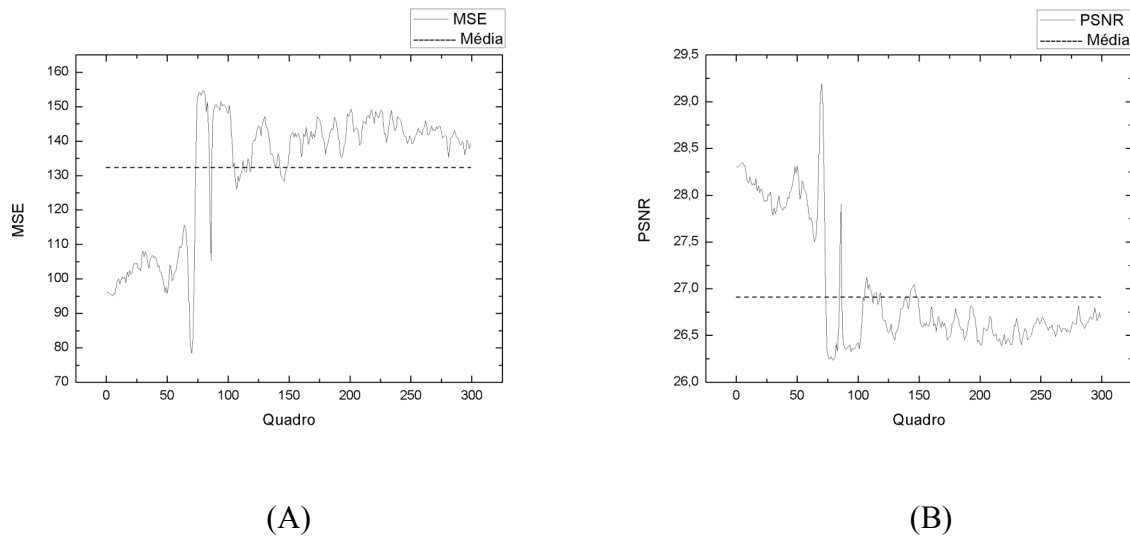
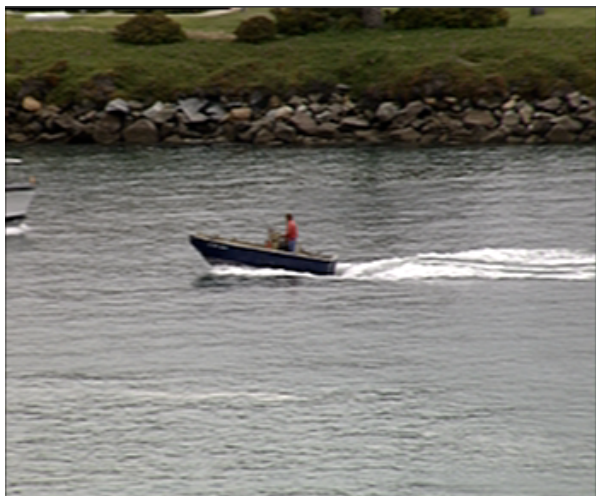


Figura 30. MSE (A) e PSNR (B) para o borramento utilizando a máscara linear de 7 *pixels*.

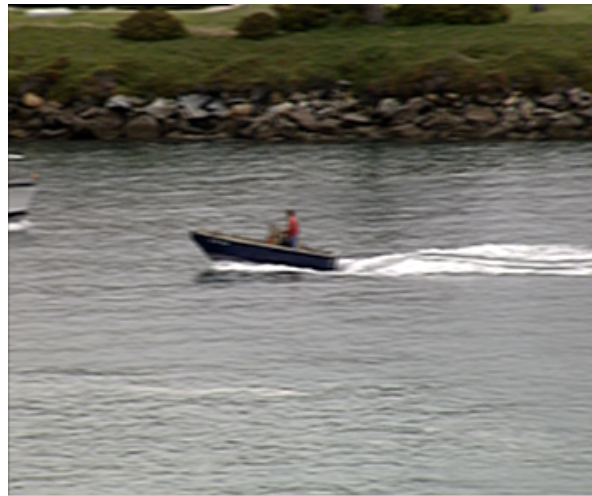
Observando os gráficos obtidos para o artefato borramento, com as máscaras lineares, é observável uma semelhança muito grande entre eles. O quesito que diferencia cada uma das máscaras é a média obtida para cada uma delas, que comportou-se conforme o esperado, ou seja, de forma diretamente proporcional ao tamanho da máscara utilizada. Porém, destacam-se dois quadros em especial, os quadros de número 70 e 86 (os “picos” e “vales” acentuados existentes nos gráficos de PSNR e MSE, respectivamente). Ao efetuar uma análise mais minuciosa no conteúdo desses quadros, nota-se que no quadro de número 70 existe um movimento vertical na câmera, o que ocasiona um borramento no vídeo original, ou seja, como o quadro de referência das métricas utilizadas (MSE e PSNR) já é borrado, o impacto da aplicação da máscara do borramento linear é menor nos *pixels* deste quadro do que nos outros. Por outro lado, no quadro de número 86 o movimento da câmera se dá na direção horizontal, gerando também um borramento no vídeo original. Por isso é observável que conforme aumenta-se o tamanho da máscara linear utilizada, a degradação neste quadro aumenta também, ocasionando um aumento no erro quadrático médio de 86,6 e uma diminuição de 5,1 db no valor de PSNR da máscara de 3 *pixels* para a máscara de 7 *pixels* (para observar tal fenômeno, vide diferença entre os picos dos gráficos da Figura 28, Figura 29 e Figura 30).

Ao notar essas características descritas anteriormente, nota-se a possibilidade de se utilizar as máscaras do borramento para detecção da movimentação da câmera no vídeo, visto que

máscaras horizontais são sensíveis a movimentos verticais, espera-se que máscaras verticais sejam sensíveis a movimentos horizontais da câmera.



(A)



(B)

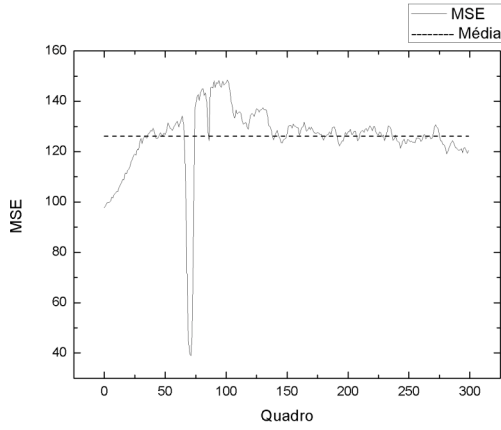


(C)

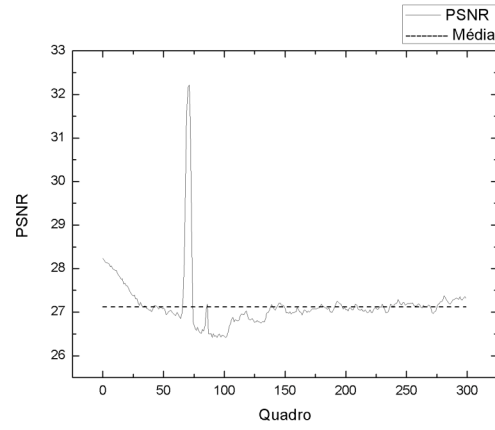
Figura 31. Quadro do vídeo “*coastguard*” com borramento gerado pelas máscaras lineares de 3 (A), 5 (B) e 7 (C) *pixels*.

Conforme o que pode ser observado na Figura 31, o impacto visual gerado pelas máscaras lineares é diretamente proporcional ao tamanho da máscara utilizada, ou seja quanto maior a máscara, maior é a degradação do vídeo.

Para o borramento utilizando as máscaras matriciais de 3×3 , 5×5 e 7×7 *pixels* foram obtidos os gráficos de MSE e PSNR da Figura 32 (A) e (B), Figura 33 (A) e (B) e Figura 34 (A) e (B).

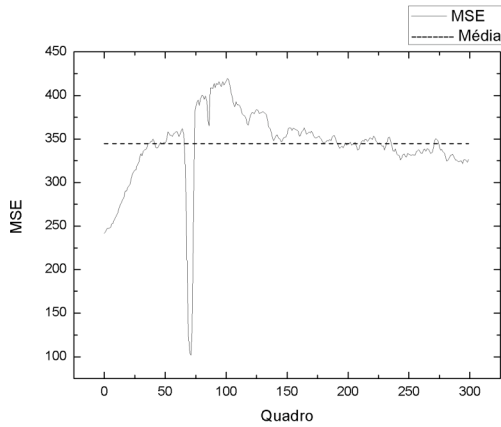


(A)

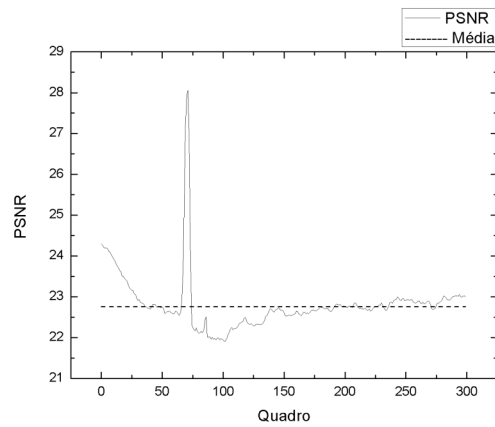


(B)

Figura 32. MSE (A) e PSNR (B) para o borramento utilizando a máscara matricial de 3×3 *pixels*.



(A)



(B)

Figura 33. MSE (A) e PSNR (B) para o borramento utilizando a máscara matricial de 5×5 *pixels*.

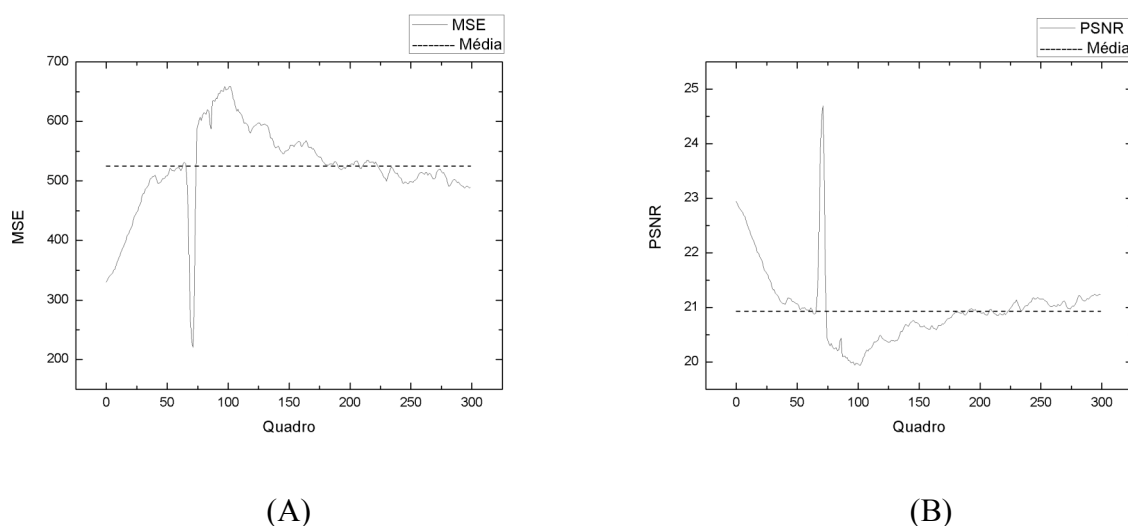


Figura 34. MSE (A) e PSNR (B) para o borramento utilizando a máscara matricial de 7x7 *pixels*.

Pode-se perceber, no artefato borramento matricial, que apenas um quadro se destaca com valor muito diferente dos outros para o MSE e o PSNR. Este quadro, na verdade, é parte integrante da movimentação da câmera, um pouco mais acentuada, que ocorre durante a execução do vídeo, o que gera um borramento na imagem do vídeo original (como já descrito anteriormente). Este quadro corresponde ao de número 71, quadro este que também se destacou na aplicação das máscaras lineares (porém, não foi este o quadro “pico” ou “vale” observado nos gráficos do borramento linear). Isto se deve ao fato de, no quadro 70, haver apenas borramento na vertical, e no quadro 71 ocorrer um borramento vertical e horizontal, dado à movimentação da câmera em diagonal (para cima e ligeiramente à direita). Ainda com relação ao borramento, é interessante notar a semelhança entre os gráficos de todos os tipos deste artefato. Esta semelhança se deve à idéia de interferência entre os *pixels* que foi utilizada para a geração do borramento e pela movimentação da câmera ser sempre a mesma (pois é o mesmo vídeo original utilizado). Além disso, é interessante observar o que ocorreu com o quadro de número 86 para o borramento com as máscaras matriciais. Este quadro (86) deixa de se destacar devido às máscaras matriciais gerarem uma maior degradação neste quadro. Isto ocorre, pois, como o borramento do vídeo original era horizontal e agora o borramento gerado é de forma vertical, horizontal e diagonal (ou seja, para a máscara 3x3 matricial, os *pixels* envolvidos na média são os que se encontram acima, abaixo, à esquerda, à direita, à esquerda e acima, à direita e acima, abaixo e à esquerda e abaixo e

à direita), nota-se que o seu destaque foi muito inferior à aplicação das máscaras lineares, isso se deu pelo fato do borramento causado pela movimentação da câmera neste quadro ser estritamente horizontal, fato este que não atenuou a influência das máscaras bidimensionais no borramento.



(A)



(B)



(C)

Figura 35. Quadro do vídeo “*coastguard*” com borramento gerado pelas máscaras matriciais de 3x3 *pixels* (A), 5x5 *pixels* (B), 7x7 *pixels* (C).

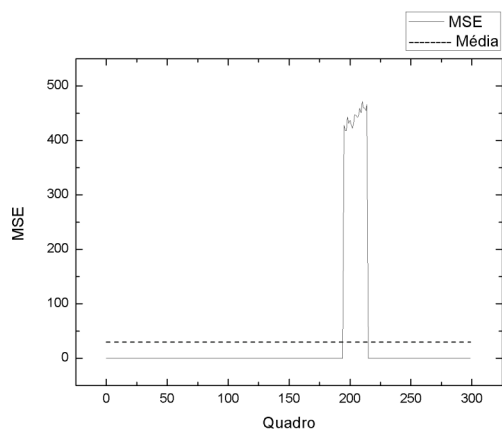
Conforme o que pode ser observado na Figura 35, o impacto visual gerado pelas máscaras matriciais é diretamente proporcional ao tamanho da máscara utilizada, assim como ocorreu nas

máscaras lineares, ou seja, quanto maior a máscara, maior é a degradação do vídeo. Entretanto, ao analisar os gráficos para todas as máscaras utilizadas, pode-se confirmar a previsão de que quanto maior for o número de *pixels* envolvidos no cálculo da média maior a degradação gerada. Isso é percebido ao comparar as médias de todos os borramentos, ou seja, para os borramentos matriciais a média do MSE foi maior do que para os borramentos lineares.

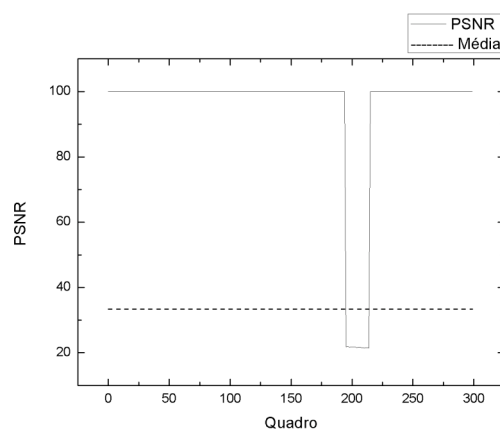
4.4 Blocagem

Com a geração do artefato blocagem é esperado que surjam blocos na imagem, de forma a simular a blocagem observada nos vídeos digitais.

Foram traçados, para o artefato blocagem com ocorrência de 1 e 5%, os gráficos da Figura 36 (A) e (B) e Figura 37 (A) e (B) para o MSE e PSNR. Uma ocorrência de 1%, significa que cada *pixel* possui uma probabilidade de 1% de ser transformado em um bloco, e a ocorrência de 5% significa que essa probabilidade é 5 vezes maior, ou seja, aparece uma maior quantidade de blocos nos quadros que sofrem a blocagem.



(A)



(B)

Figura 36. MSE (A) e PSNR (B) para 1% de ocorrência de blocagem, 20 quadros foram afetados.

Ao observar os gráficos de blocagem percebe-se uma característica importante em seu comportamento, devido ao vídeo permanecer idêntico ao original a maior parte do tempo. Quando ocorre a blocagem o gráfico do PSNR acaba sofrendo um “degrau” abrupto até que a blocagem cesse. Para 1% dos *pixels* afetados, a blocagem teve início no quadro de número 195 e teve a duração de 20 quadros, ou seja, até o quadro de número 214. Já para 5% dos *pixels* afetados, o início ocorreu no quadro de número 171 e teve a duração de 6 quadros, ou seja, até o quadro de número 176. Entretanto, observa-se que a quantidade de *pixels* afetados no quadro implica em um MSE maior, por existir maior diferença entre o quadro degradado e o original, pois mais blocos são inseridos. Por isso, o valor da média do MSE depende também da duração do artefato. Isto é visto na Figura 36 e Figura 37, onde para 1% dos *pixels* afetados a média do MSE foi maior do que para 5% dos *pixels* afetados. Isso é devido à duração da blocagem a 1% ter sido cerca de 3 vezes maior do que para 5%. Além disso, a diferença observada entre os valores médios de PSNR foi de apenas 2,5 db, onde o valor menor (33,4 db) foi obtido com a ocorrência de 1% e o maior (35,9 db) com a ocorrência de 5% deste artefato, isso também se deve à duração do artefato ter sido maior para 1% de ocorrência.

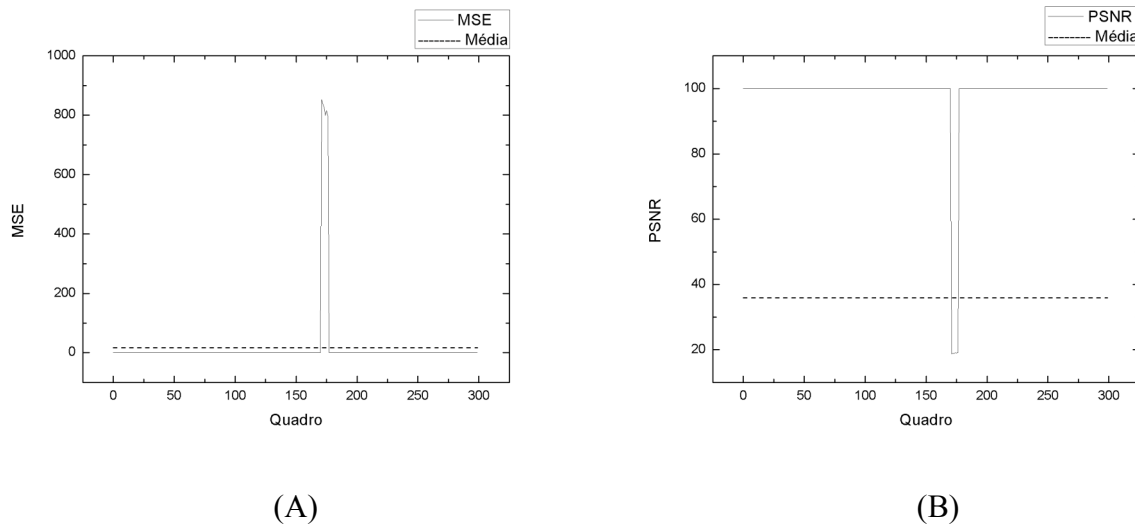


Figura 37. MSE (A) e PSNR (B) para 5% de ocorrência de blocagem, 6 quadros foram afetados.

Ao observar o quadro estático da Figura 38, observa-se uma semelhança com os blocos observados na transmissão de televisão digital, porém ao assistir o vídeo relacionado, percebe-se

que os blocos ficam movendo-se por toda a imagem, o que não havia sido previsto. Este movimento dos blocos ocasiona uma baixa semelhança com o que realmente é observado nas transmissões de televisão digital. Isso ocorre porque para cada quadro gerado com este artefato são selecionados aleatoriamente os *pixels* que sofrerão a blocagem, isso leva à seleção de *pixels* distintos a cada quadro, fazendo com que os blocos sejam gerados em posições diferentes a cada quadro.



(A)



(B)

Figura 38. Quadro do vídeo “*coastguard*” apresentando o artefato blocagem com ocorrência de 1% (A) e 5% (B).

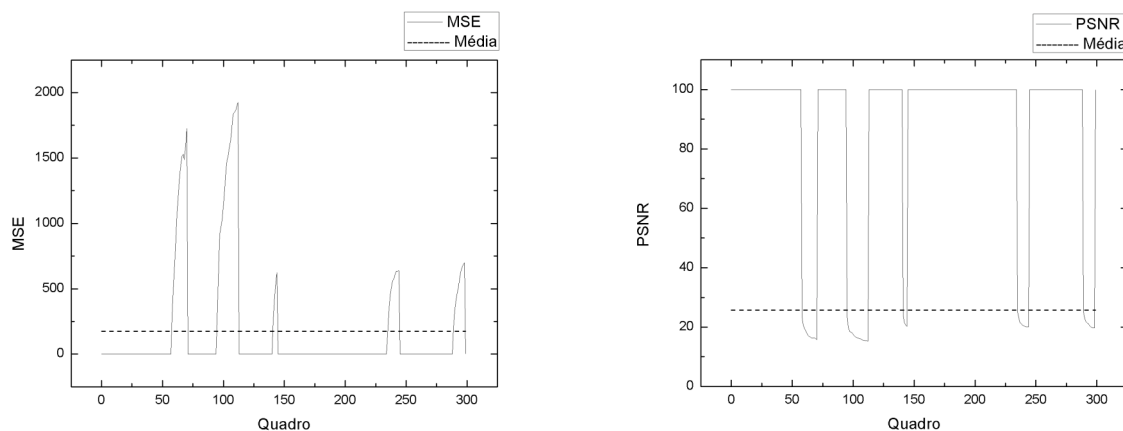
4.5 Travamento

Ao serem realizados os testes com o algoritmo para a geração do travamento, a expectativa é que durante sua ocorrência, a imagem fique estática, até a duração do artefato ser atingida, quando então o vídeo volta a ser exibido.

Para o travamento foram obtidos os gráficos de MSE e PSNR da Figura 39 (A) e (B).

O travamento possui uma característica muito própria, pois o PSNR cai abruptamente quando se inicia o travamento e em seguida possui um comportamento mais suave até o seu final (isso pode ser observado no gráfico (B) da Figura 39), quando sofre um ganho brusco no valor de

PSNR. Isto se deve à diferença gritante que existe entre um quadro e o seu precedente (travado). Logo em seguida, quando o quadro está travado, ele possui diferenças apenas ligeiramente maiores do que os outros quadros que pertencem ao travamento, provocando o efeito de suavização de valor visto nas curvas. Quando o vídeo destrava, o PSNR volta ao seu valor máximo, pois os quadros voltam a ser iguais.



(A)

(B)

Figura 39. MSE (A) e PSNR (B) para o travamento.

A partir das análises dos resultados foi observado que cada artefato gera um padrão próprio de comportamento para o PSNR e MSE.

5 CONCLUSÃO E TRABALHOS FUTUROS

Com a realização deste trabalho foi possível observar, diante dos resultados obtidos, que a geração dos artefatos, em sua maioria, aproximou-se muito do observado nas transmissões de televisão atuais, como foi possível observar pelas figuras exibidas no capítulo anterior. O artefato de blocagem, quando observado em um quadro isolado se parece muito com o real experimentado na televisão digital. Porém, quando observada uma sequência de quadros que apresentam este artefato, devido aos blocos serem gerados aleatoriamente, eles acabam por se movimentar em toda a área da imagem, o que não é comumente observado na prática. Todavia, para solucionar este problema é sugerido que o algoritmo seja aprimorado, detectando as áreas mais propícias ao aparecimento da blocagem (WU, RAO, 2006) e submetendo estas áreas ao algoritmo da blocagem.

Além disso, os resultados obtidos através dos gráficos de PSNR e MSE mostram-se animadores quanto à detecção de características próprias de cada artefato, pois isso poderá ser utilizado futuramente para a detecção automática de cada um deles. Foi possível também notar, através dos resultados, que, valores de PSNR abaixo de 20 dB implicam em uma qualidade intolerável para o vídeo, como pode ser observado nas figuras do capítulo anterior.

Nos resultados de (SILVA, 2009) onde as métricas totalmente referenciadas (MSE e PSNR) aqui implementadas foram aplicadas em vídeos cuja imagem original era de boa qualidade e foram inseridos os artefatos aqui propostos, foi possível observar que existe uma correlação da MOS, obtida pela avaliação subjetiva, e dos valores de PSNR e MSE, obtidos através das métricas objetivas. Com isso é possível notar que existe um intervalo de validade para as métricas aqui descritas, visto que, para os artefatos propostos, o comportamento das métricas objetivas e subjetiva pôde ser correlacionado.

Como trabalhos futuros sugere-se que sejam criados outros artefatos, como o efeito mosquito, erro de quantização, efeito escada, *ringing*, contornos falsos, *jerkiness*, entre outros. E dando continuidade a este trabalho tentar desvendar características próprias para cada um desses artefatos gerados. Pode-se criar vídeos combinando vários artefatos e fazer a avaliação. Propõe-se ainda a realização do estudo comparativo das métricas totalmente referenciadas com a avaliação subjetiva, para mais amostras do travamento e para os outros artefatos a serem gerados. Além

disso, acredita-se na possibilidade de serem criados algoritmos utilizando-se de fundamentos da inteligência artificial, para que estes descubram padrões nos ruídos e consigam detectá-los automaticamente. Outra alternativa pode ser utilizar-se de métricas sem referência ou com referência reduzida para buscar outros padrões interessantes nestes artefatos. E ainda, utilizando-se das características apresentadas para cada um dos artefatos desenvolvidos, seria interessante criar métricas sem referência que visassem detectar os artefatos e medir a qualidade da imagem baseada neles. Mais uma possibilidade é utilizar as máscaras do borramento para detecção da movimentação da câmera no vídeo, visto que máscaras horizontais são sensíveis a movimentos verticais, espera-se que máscaras verticais sejam sensíveis a movimentos horizontais da câmera, além da utilização do filtro gaussiano para o borramento. Mais uma opção promissora, seria o desenvolvimento de rotinas para tentar suavizar os artefatos aqui apresentados, pois uma vez detectados poderiam ser reduzidos ou atenuados na imagem final, como o filtro da média faz, por exemplo, no caso do ruído branco gaussiano. E a partir do desenvolvimento de novas métricas, estas poderiam ser embarcadas para que, por exemplo, o próprio decodificador detectasse os artefatos e pudesse aplicar a rotina correta para corrigir o artefato detectado, melhorando assim a qualidade da imagem e o impacto visual que aquele “defeito” teria sobre o telespectador.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDURAHIMAN, V. **Geração de Números e Variável Aleatórias**, Disponível em: <http://www.comp.ita.br/~rahiman/CC236/Geracao_RNeRV.pdf>, Acesso em: 4 de abril de 2009.
- ALBIN, S. ROUGERON, G., PÉROCHE, B. TRÉMEAU, A. **Quality Image Metrics for Synthetic Images Based on Perceptual Color Differences**. IEEE Transactions on Image Processing, Vol. 11, No. 9, Setembro, 2002, p. 961-971.
- AUWERA, G. Van der, David, P. T., REISSLEIN, M., **Traffic Characteristics of H.264/AVC Variable Bit Rate Video**. IEEE Communications Magazine. Novembro, 2008, P.164-174
- BOVIK, A. **Handbook of Image & Video Processing**, Academic Press, Canada, 2000.
- DTV, **Site Oficial da TV Digital Brasileira**, Disponível em: <<http://www.dtv.org.br/>> Acesso em: 5 de Dezembro de 2008.
- DURANT, L. **Comparing Objective and Subjective Picture Quality Measurements**. Tektronix Technical Brief, 1998. Disponível em: <http://www.tek.com/Masurement/App_Notes/25_12866/eng/25W_12866_0.pdf> Acesso em: 21 de Janeiro de 2009.
- EDEN, A. **No-Reference Image Quality Analysis for Compressed Video Sequences**. IEEE Transactions on Broadcasting, Vol. 54, No. 3, Setembro 2008, p. 691-697.
- FARIAS, M. C. Q. de, **No-Reference and Reduced Reference Video Quality Metrics: New Contributions**, Tese de doutorado do departamento de Engenharia Elétrica e Computacional da Universidade da Califórnia, Santa Barbara, USA, Setembro de 2004.
- FARIAS, M. C. Q. de, FOLEY, J. M., MITRA, S. K. **Detectability and Annoyance of Synthetic Blocky, Blurry, Noisy, and Ringing Artifacts**. IEEE Transactions on Signal Processing, Vol. 55, No. 6 part 2, Junho, 2007 p. 2954-2964

FERNANDES, J., LEMOS, G., SILVEIRA, G. **Introdução à Televisão Digital Interativa: Arquitetura, Protocolos, Padrões e Práticas.** Disponível em: <http://www.cic.unb.br/~jhcf/MyBooks/itvdi/> Acesso em: 17 de Dezembro de 2008.

Graphics and Media Lab, The Laboratory of Computer Graphics at Department of Computer Science of the Moscow State University. Disponível em: <http://research.graphicon.ru>. Acesso em: 7 de Fevereiro de 2009.

GUNAWAN, I. P., GHANBARI, M. **An Efficient Reduced-Reference Video Quality Metric.** Proceedings of the 26th Picture Coding Symposium (PCS 2007), Lisboa, Portugal, Novembro, 2007, p. 07-09.

GUNAWAN, I. P., GHANBARI, M. **Reduced-reference Picture Quality Estimation by Using Local Harmonic Amplitude Information.** Proc. London Communications Symposium 2003, University College London, UK, 8-9, September, 2003, p. 137-140.

HANZO, L., CHERRIMAN, P., STREIT, J. **Video Compression and Communications**, John Wiley & Sons Ltd., England, 2007.

HOFFMANN, H. ITAGAKI, T. WOOD, D. HINZ, T. WIEGAND, T. **A Novel Method for Subjective Picture Quality Assessment and Further Studies of HDTV Formats.** IEEE Transactions on Broadcasting, Vol. 54, No. 1, Março, 2008, p. 1-13.

ITU. Recommendation ITU-R BT.500: **Methodology for the Subjective Assessment of the Quality of Television Pictures.** International Telecommunication Union, 2000.

ITU. Recommendation ITU-R BT.601: **Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide Screen 16:9 Aspect Ratios.** International Telecommunication Union, 2007.

ITU. Recommendation ITU-T P.910: **Subjective Video Quality Assessment Methods for Multimedia Applications.** International Telecommunication Union, 2008.

KLAUE, J., RATHKE, B., WOLISZ, A. **EvalVid – A Framework for Video Transmission and Quality Evaluation.** Proc. of the 13th International Conference on Modelling Techniques

- and Tools for Computer Performance Evaluation, Urbana, Illinois, USA, Setembro, 2003, p. 255-272.
- KOUMARAS, H., PLIAKAS, T., KOURTIS, A. **A Novel Method for Pre-Encoding Video Quality Prediction**. Mobile and Wireless Communications Summit, 2007. 16th IST, 2007, p. 1- 4.
- KUROSE, J. F., ROSS, K. W. **Redes de Computadores e a Internet**, Addison Wesley, São Paulo, SP, 2007.
- LAMBRECHT, C. J. van den B.; VERSCHEURE, O. **Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System**. IS&T/SPIE, San Jose, CA, USA, 1996, Vol. 2668, p. 450-461.
- LIU, S., BOVIK, A. C. **Efficient DCT-Domain Blind Measurement and Reduction of Blocking Artifacts**. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 12, Dezembro, 2002 p. 1139-1149.
- LU, L., LU, X. **Quality Assessing of Video Over a Packet Network**. Second Workshop on Digital Media and its Application in Museum & Heritage, 2007, p. 365-369.
- MARQUES FILHO, O., VIERA NETO, H. **Processamento Digital de Imagens**, Brasport, Rio de Janeiro, RJ, 1999.
- MENDES, L. L., FASOLO, S. A. **Introdução a Televisão Digital**. Anais da Semana Internacional das Telecomunicações, Santa Rita do Sapucaí, 2002, p. 253-258.
- MMLAB, **Multimedia Signal Processing and Understanding LAB**, DISI – Dipartimento di Ingegneria e Scienze dell'Informazione, Università degli Studi di Trento. Disponível em: <<http://mmlab.science.unitn.it/downloads/material>> Acesso em: 9 de Fevereiro de 2009
- MONTEZ, C., BECKER, V. **TV Digital Interativa: Conceitos e Tecnologias**. Anais do WebMídia & LA-Web, Ribeirão Preto – SP, 2004, p. 39-77.

MSU Quality Measurement Tool, Disponível em:
http://compression.ru/video/quality_measure/info_en.html. Acesso em: 7 de Fevereiro de 2009

QIU, S., RUI, H., ZHANG, L. **No-reference Perceptual Quality Assessment for Streaming Video Based on Simple End-to-End Network Measures**. International Conference on Networking and Services (ICNS'06), Setembro, 2006, p. 53 – 58.

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

SANTOS, J. de J. **Um Estudo Sobre os Métodos de Avaliação Qualitativa e de Recuperação de Regiões Afetadas pelo Efeito de Blocação em Imagens Codificadas por Transformada**. Tese de Mestrado do departamento de Comunicações da Universidade de Campinas, Campinas, SP, 2002.

SCHWARTZ, M., SHAW, L., **Signal Processing**: Discrete Spectral Analysis, Detection, and Estimation. McGRAW-HILL KOGAKUSHA, LTD, Tokyo, Japão, 1975.

SILVA, E. **Avaliação Subjetiva de Videos**, Relatório Técnico, CPGEI, UTFPR, 2009.

SLANINA, M., RÍCNY, V. **A Comparison of Full-Reference Image Quality Assessment Methods**. Electronic Communication System and New Generation Technology (ELKOM), 2006.

STOLFI, G. **Compressão de Imagens em Movimento, Padrões MPEG-2 e H.264**, 2007. Disponível em: <http://www.lcs.poli.usp.br/~gstolfi/>. Acesso em: 5 de Dezembro de 2008.

TANG, H., CAHILL, L. W. **Improved Median Filter for Restoration of Images Corrupted with Gaussian Noise**. Electronics Letters, Vol. 27, No. 12 de 6 de junho 1991, p.12-14.

TRIANTAFYLLIDIS, G. A., TZOVARAS, D., STRINTZIS, M. G. **Blocking Artifact Detection and Reduction in Compressed Data**. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 10, Outubro, 2002, p. 877-890.

Video Traces Research Group, Disponível em: <<http://trace.eas.asu.edu/index.html>>. Acesso em 17 de Dezembro de 2008.

VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Abril, 2000. Disponível em: <<http://www.vqeg.org/>>. Acesso em: 21 de março de 2009.

WANG, Y. Survey of Objective Video Quality Measurements. Technical Report, EMC Corporation Hopkinton, 2006.

WANG, Z., BOVIK, A. C. Modern Image Quality Assessment. Morgan & Claypool, USA, 2006

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

WHITAKER, J., BENSON, B., Standard Handbook of Video and Television Engineering. McGraw-Hill, New York, 2003.

WOLF, S., PINSON, M. Video Quality Measurement Techniques. NTIA Report 02-392. U.S. Department of Commerce – National Telecommunications and Information Administration. Junho, 2002. Disponível em: <<http://www.its.bldrdoc.gov/pub/ntia-rpt/02-392/>>. Acesso em: 21 de março de 2009.

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

XIAO, F. DCT-based Video Quality Evaluation, student final project Digital Video Processing (EE392J), Stanford University, winter, 2000. Disponível em: <http://compression.ru/video/quality_measure/vqm.pdf>. Acesso em: 21 de março de 2009.

RESUMO:

Este trabalho teve como objetivo o estudo e a criação controlada de artefatos em vídeo digital e a sua mensuração com o emprego de métricas objetivas. Foram criados artefatos do tipo ruído branco gaussiano, sal e pimenta, borramento, blocagem e travamento. Para avaliação de seu impacto no vídeo digital foram utilizadas as métricas MSE (*Mean Squared Error*) e PSNR (*Peak Signal-to-Noise Ratio*)

PALAVRAS-CHAVE:

Qualidade de vídeo. Métricas com referência total. PSNR. Geração de artefatos em vídeo. Vídeo digital. Avaliação objetiva da qualidade de vídeo.

ÁREA/SUB-ÁREA DE CONHECIMENTO:

3.04.06.03-0 Sistemas de Telecomunicações

2009

Nº: 498

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](#)