

# Proposição de um Método de Análise de Qualidade de Vídeo sem Referência Completa

Alessandro Marchetto<sup>1</sup>, Ricardo Becker<sup>2,3</sup>, Ricardo Balbinot<sup>4</sup>

<sup>1,2</sup>Universidade de Caxias do Sul  
Caixa Postal 32 – 95700000 – Bento Gonçalves – RS – Brasil

<sup>3</sup>Faculdade SENAI de Tecnologia  
91140000 – Porto Alegre – RS – Brasil

<sup>4</sup>IFRS - Instituto Federal do Rio Grande do Sul  
92412-240 - Canoas - RS - Brasil

{amarchetto90, ricardobecker.eng}@gmail.com, ricardo.balbinot@canoas.ifrs.edu.br

**Abstract.** *This paper proposes a method of objective video quality assessment, to analyze the impact of packet loss on measured quality, without full reference. First, for comparison purposes, tests were conducted using full reference. The method consists of inserting an extra frame, every defined interval, which is subsequently used to assist in measuring the quality. The extra frames are transmitted along with the other frames of the video. After comparing the result with the proposed method and full reference, the error is calculated. From the error, a graphic adjustment constant was calculated, so that the results from the method were closer than the results using full reference. Subsequently, the error was recalculated, resulting in a significantly lower value than before the adjustment.*

**Resumo.** *Este trabalho propõe um método de análise objetiva de qualidade de vídeo, para avaliar qual o impacto do descarte de pacotes sobre a qualidade medida, sem a utilização de referência completa. Primeiramente, para fins de comparação, foram realizados testes utilizando referência completa. O método consiste na inserção de um frame extra, a cada intervalo definido, que é posteriormente utilizado para auxiliar na medição da qualidade. Os frames extras são transmitidos juntamente com os demais frames do vídeo. Após comparar o resultado do método proposto com o resultado em referência completa, o erro é calculado. A partir do erro, foi calculada uma constante de ajuste dos gráficos, para que o método proposto se aproxime mais dos resultados em referência completa. Posteriormente, o erro foi recalculado, resultando em um valor significativamente menor do que antes do ajuste.*

## 1. Introdução

De acordo com [Tao et. al. 2007], estudos mostram a transmissão de vídeo como um elemento crescente no tráfego da Internet. Provedores de vídeo estão adotando redes IP como seu veículo para entrega, visto IPTV.

A qualidade do vídeo é afetada em conjunto por vários fatores dependentes da rede e de fatores específicos da aplicação. Por exemplo, perda de pacotes e *jitter* são os principais fatores que dependem da rede, enquanto o codec de vídeo, a perda técnica de recuperação, a taxa de bits de codificação, o esquema de empacotamento, e

características de conteúdo são os principais fatores específicos da aplicação que afetam a qualidade de vídeo e sua sensibilidade ao erro de rede [Tao et. al. 2007].

Dado o crescimento da transmissão de vídeos via rede IP, este trabalho visa analisar a influência da perda de pacotes, um fator que depende da rede, na qualidade percebida do vídeo, analisando a mesma de forma objetiva, através da métrica PSNR[Chikkerur et. al. 2011] [Seshadrinathan et. al. 2010].

A análise da qualidade medida é feita sem a utilização de referência completa, onde a ideia é inserir um *frame* extra, para que a qualidade possa ser aferida. O processo de descarte de pacotes é baseado em um modelo de processo de Poisson, onde cada *frame* é processado individualmente, para diferentes taxas de perda.

Após obter os resultados pelo método proposto e através de referência completa, foi calculado o erro entre os resultados, para ajustados valores do método proposto, para que fiquem mais próximos dos medidos com referência completa.

Na sequência deste trabalho são apresentadas na seção dois as principais referências relacionadas. Na seção três é apresentada a metodologia aplicada e nas seções quatro e cinco os resultados e conclusões, respectivamente.

## 2. Referencial Teórico

Nesta seção serão apresentados conteúdos pertinentes ao desenvolvimento do trabalho. Na subseção 2.1 é apresentada a teoria de descarte de pacotes. Nas seções 2.2 e 2.3 são abordados os temas de qualidade de vídeo e de análise de qualidade de vídeo, respectivamente.

### 2.1. Modelo de descarte de pacotes

Segundo [Jain 1991], processos de Poisson são usados em modelos de filas, os quais, por sua vez, são a base para explicar descarte em redes de pacotes. Eles são apropriados se os dados chegam a partir de um grande número de fontes independentes. Esses processos de chegada são chamados processos de Poisson ou fluxos de Poisson. Os fluxos de Poisson são populares nas teorias de filas, pois as chegadas são sem memória, enquanto o tempo entre chegadas é distribuído exponencialmente.

O valor esperado de uma variável aleatória distribuída através de Poisson é igual a média da distribuição de Poisson, e deve ser um número real ( $\lambda > 0$ ) [Jain 1991]. Segundo [Johnson et. al. 1993], o coeficiente de variação é dado por  $\lambda^{-1/2}$ , enquanto o índice de dispersão é igual a 1.

De acordo com [Yates et. al. 2004], uma variável aleatória discreta  $X$  é dita ter uma distribuição de Poisson, se a função de massa de probabilidade de  $X$  é dada pela equação (1), onde  $e$  é o número de Euler,  $\lambda$  é a média da distribuição de Poisson,  $k$  é gama de valores ocorridos e  $k!$  é o fatorial de  $k$ .

$$f(k; \lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (1)$$

De acordo com [Knuth 1997], distribuições exponenciais são números que representam situações de tempo de chegada, e a distribuição é dada pela equação (2), onde  $\mu$  é a média. Pode-se usar a distribuição exponencial para estimar o instante de ocorrência de um evento em um processo de Poisson.

3

$$F(x) = 1 - e^{-\frac{x}{\mu}} \quad (2)$$

O método de [Knuth 1997] é uma forma de implementar a função cumulativa de probabilidade da distribuição exponencial. Pelo método, se a equação (3) é verdadeira, então o mesmo vale para a (4). Assim, conclui-se que a equação (5) tem distribuição exponencial. Como 1-U é uniformemente distribuído quando U também é, pode-se concluir que a equação (6) é distribuída exponencialmente com média  $\mu$ . Esse método requer menor precisão para estimar o instante de ocorrência dos eventos no processo.

$$y = F(x) = 1 - e^{-\frac{x}{\mu}} \quad (3)$$

$$\xi = F^{-1}(y) = -\mu * \ln(1 - y) \quad (4)$$

$$-\mu * \ln(1 - U) \quad (5)$$

$$X = -\mu * \ln(U) \quad (6)$$

## 2.2. Qualidade de Vídeo

Segundo [Winkler 2007], qualidade de vídeo tem muitos aspectos, e a avaliação da qualidade de sistemas de vídeo digital se tornou algo complexo. Isso se deve a:

- *Sistemas de vídeos consistem de vários componentes que afetando a qualidade do vídeo de alguma maneira: hardware de captura e reprodução; codecs e rede.*
- *A percepção visual é ainda mais complexa. Para medir a qualidade de uma forma precisa, como as pessoas percebem o vídeo e sua qualidade.*

De acordo com [Winkler 2007], vários fatores contribuem para o que um observador percebe como qualidade de vídeo. Elas incluem interesses pessoais, expectativas de qualidade, tipo de *display* e suas propriedades, condições de visualização, entre outras. As métricas utilizadas focam na fidelidade visual do vídeo em termos de distorções causadas pelo processamento, ao invés de fatores subjetivos.

## 2.3. Técnicas de avaliação de Qualidade de Vídeo

### 2.3.1 PeakSignal-to-NoiseRatio - PSNR

Para [Huynh-thu et. al. 2008], a relação sinal-ruído de picoé usada em sistemas analógicos como uma métrica de qualidade. No entanto, a tecnologia de vídeo digital expôs algumas limitações desta, mesmo assim, devido à sua baixa complexidade, ainda é utilizada como métrica para avaliar algoritmos de processamento de imagem.

Para cada sequência de vídeo, o PSNR é medido pela equação (7), onde  $\varepsilon(i)$  é o erro médio quadrático da luminância do pixel, correspondente ao quadro  $i$  no vídeo de referência e compactada, e  $N$  é o número de quadros no vídeo degradado.

$$PSNR = \frac{1}{N} \sum_{i=1}^N 10 * \log_{10} \frac{255^2}{\varepsilon(i)} \quad (7)$$

### 3. Metodologia

#### 3.1. Dataset

A *database* utilizada, obtida do IVP Laboratory [IVP 2014], possui vídeos em alta definição (1920x1088), filmados RAW, a 25 *frames/s* sem compressão, que são utilizados como referência. Antes de gerar distorção nos vídeos, eles foram convertidos de RAW para YUV 4:2:0 e após para RGB. Cada vídeo tem 10 segundos, sem áudio. O vídeo utilizado como referência foi a sequência *laser*, escolhido entre os disponíveis.

#### 3.2. Modelo de Descarte de Pacotes

O processo consiste em simular o processo de fila, encontrado em redes IP, para avaliar qual o impacto destas perdas na qualidade do vídeo. Para tanto, segundo [Mansfield et. al. 2010], perdas entre 5% e 10% do total de pacotes enviados tem um impacto significativo na qualidade de serviço. Para [Zennaro et. al. 2015], menos de 1% de perda de pacotes é “bom” para áudio e vídeo *streaming*, e entre 1% e 2.5% são aceitáveis.

Neste trabalho as perdas foram entre 0% e 10%, considerando [Mansfield et. al. 2010] e [Zennaro et. al. 2015]. Os intervalos na faixa de simulação variam de 0,2%.

O modelo de descarte, baseado no processo de Poisson, foi implementado em *Matlab*. Foi utilizada a equação (8), que expressa quantos pacotes após o atual se dará o descarte.  $\lambda$  é a média esperada do número total de descartes para a população. Começando do primeiro pacote, uma variável auxiliar é utilizada para controlar quais pacotes serão perdidos. Os pacotes perdidos são multiplicados por zero, cor preta. As figuras 1 e 2 ilustram o primeiro *frame* da sequência, com perda de 1% e 10%, respectivamente. O processo é repetido para todos os *frames* do vídeo, para todos os níveis de perda estabelecidos.

$$Prox = round\left(\frac{-\ln(1 - rand(1))}{\lambda}\right) \quad (8)$$

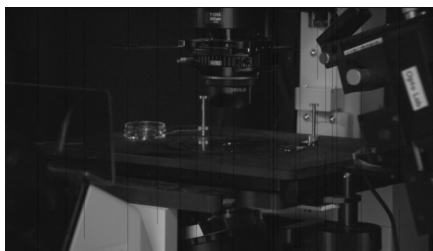


Figura 1. Primeiro *frame* da sequência com perda de 1%, em escalas de cinza.

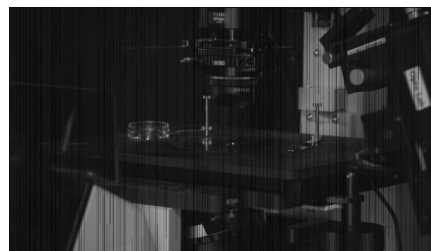


Figura 2. Primeiro *frame* da sequência com perda de 10%, em escalas de cinza.

#### 3.3. Método Proposto

O método proposto consiste em inserir um *frame* extra, que será enviado para o receptor juntamente com os demais *frames* do vídeo. Por exemplo, com o intervalo em 5, um *frame* extra, que é uma cópia do *frame* anterior, é enviado a cada 5 *frames* originais, a fim de

comparação do *frame* extra com o seu original. Como o método utiliza uma referência limitada, primeiramente foram realizados testes utilizando referência completa.

Os intervalos foram definidos como um *frame* extra a cada 2, 4, 6, 8 e 10 *frames*. O *frame* extra também passou pelo modelo de descarte de pacotes, para que seja feita uma medida de qualidade.

## 4. Resultados

### 4.1. Testes utilizando referência completa

A figura 3 representa os valores PSNR medidos para a sequência *laser* em relação aos níveis de descarte de pacotes mencionados. Os valores foram obtidos através de referência completa.

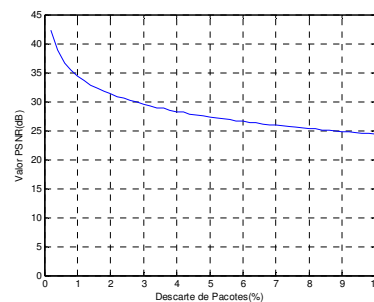


Figura 3. Valores PSNR utilizando referência completa.

### 4.2.2 Métrica PSNR

Os valores PSNR, figura 4, não apresentaram variação significativa entre si, determinando um intervalo de 10 *frames*. O intervalo escolhido gera menor volume de dados.

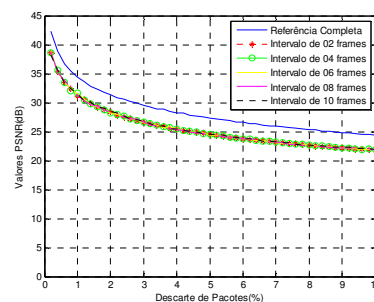


Figura 4. Comparação dos valores PSNR.

O erro RMS, equação (10), e a sua representatividade, equação (11), são mostrados na tabela 1.  $E_{rms}$  é o valor do erro RMS e  $\max(\text{PSNR})$  é o valor máximo medido pela métrica PSNR para determinado intervalo.

$$E_{RMS} = \sqrt{\frac{\sum_{i=1}^n (y(i) - y'(i))^2}{n}} \quad (10)$$

$$R\% = \frac{E_{rms}}{\max(PSNR)} \quad (11)$$

Tabela 1. Erro e representatividade para a métrica PSNR.

Intervalo (frames)	Erro RMS (dB)	Representatividade (%)
2	2.8482	6.7401
4	2.8521	6.7493
6	2.8713	6.7948
8	2.8835	6.8236
10	2.8190	6.6710

### 4.3. Ajuste dos Gráficos

Foi feito um ajuste dos gráficos para aproximar dos valores da referência completa. Para a PSNR, como as curvas se assemelham, ajustando a amplitude, o erro foi extraído através de uma subtração da curva dos resultados do método proposto, no intervalo de 10 frames, da curva em referência completa (figura 5). Para tanto, foi feita uma média pela equação (12), para se obter uma aproximação com menor valor de erro. Assim, aplicando a equação (12) nos dados da figura 5, foi obtido um erro médio de 2.8077 dB. Após aplicar a equação, o resultado foi somado aos valores do método proposto. A figura 6 representa o valor corrigido.

$$E = \frac{\sum_{i=1}^n e(i)}{n} \quad (12)$$

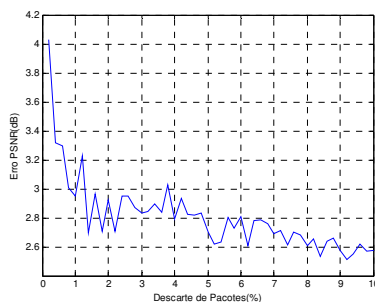


Figura 5. Erro entre os valores PSNR.

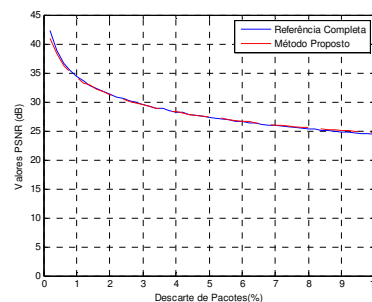


Figura 6. Valores corrigidos para um intervalo de 10 frames.

Após corrigir os valores dos gráficos, para que os mesmos se aproximem mais dos valores de referência completa, o erro RMS e a sua representatividade foram recalculados pelas equações (10) e (11). Os valores obtidos estão expostos na tabela 2.

Para validar o método proposto, o mesmo foi aplicado para as sequências de vídeo *tube* e *tractor*, da mesma *database*. As figuras 7 e 8 representam os resultados.

Ao analisar os gráficos da figura 7 e 8, se comprova o comportamento obtido para a sequência *laser*. Para ajustar ambos os gráficos, foi utilizado o mesmo valor de erro obtido para a sequência *laser*. As figuras 9 e 10 representam os gráficos ajustados.

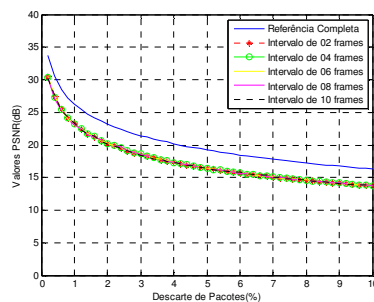
A tabela 3 representa os valores de erro e a representatividade do erro em relação ao máximo valor medido, de acordo com as equações (10) e (11).

**Tabela 2. Representatividade dos valores de erro ajustado para a métrica PSNR, para a sequência *laser*.**

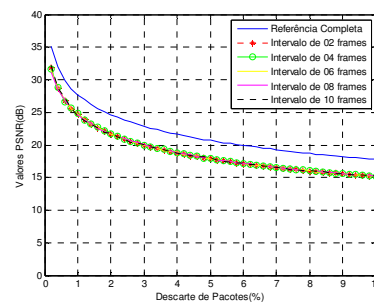
Intervalo	Erro RMS (dB)	Representatividade(%)
2	0.2157	0.5104
4	0.2212	0.5235
6	0.2257	0.5341
8	0.2222	0.5258
10	0.2521	0.5966

**Tabela 3. Erro RMS e Representatividade para os valores corrigidos das sequências *tube* e *tractor*.**

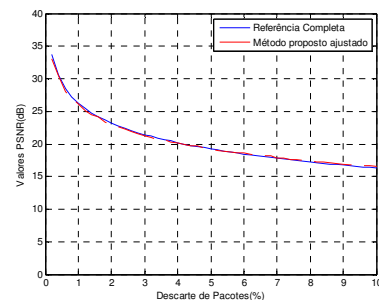
Intervalo	Erro RMS (dB)		Representatividade	
	<i>tube</i>	<i>tractor</i>	<i>tube</i>	<i>tractor</i>
2	0.1589	0.1629	0.4715	0.4628
4	0.1670	0.1931	0.4955	0.5486
6	0.1525	0.1995	0.4523	0.5667
8	0.2010	0.2449	0.5962	0.6959
10	0.1872	0.1751	0.5554	0.4974



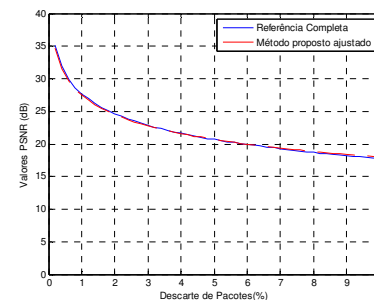
**Figura 7. Valores PSNR para a sequência *tube*.**



**Figura 8. Valores PSNR para a sequência *tractor*.**



**Figura 9. Valores corrigidos para um intervalo de 10 frames, para a sequência *tube*.**



**Figura 10. Valores corrigidos para um intervalo de 10 frames, para a sequência *tractor*.**

## 5. Conclusão

Este trabalho propôs um método para avaliar a qualidade de vídeos enviados por uma rede IP, com a inserção de um *frame* extra, o qual é inserido em intervalos e utilizado para a comparação e medição da qualidade de maneira objetiva. Tal método insere

frames extras durante a transmissão do vídeo. Tais informações podem ser adequadas para gerar a menor sobrecarga possível da rede, porém, ao adequar os valores, também se diminui o número de amostras utilizadas para medir a qualidade do vídeo.

A métrica PSNR se mostrou resultados ajustáveis através de uma constante. Mesmo com a variação na amplitude dos resultados, pode-se considerar o erro “constante”, utilizando o mesmo valor de erro para ajustar as medições. Haja visto que, o valor do erro RMS apresentou-se abaixo de 1% entre os valores de referência completa e os valores do proposto.

A bibliografia utilizada não apresenta implementação como o método proposto, de maneira que são escassos os dados para comparação. Neste sentido, o trabalho mostra potencial para continuidade, vista a necessidade de massificação dos testes com variações diversas dos parâmetros apresentados.

### Referências

- CHIKKERUR, S.; S., V.; R., M.; K., L. J. (2011) “Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison.” *Broadcasting, IEEE Transactions on*: 165 – 182.
- INTERNATIONAL TELECOMMUNICATION UNION.(2012) ITU-R BT.500-13. “Methodology for the subjective assessment of the quality of television pictures.”
- IVP Laboratory database (2014).Disponível em: <<http://ivp.ee.cuhk.edu.hk/research/database/subjective/index.shtml>> Acessado em: 02 outubro 2014.
- JAIN, R. (1991) “Art of Computer Systems Performance Analysis Techniques for Experimental Design Measurements Simulation and Modeling.” New York J. W.
- JOHNSON, N. L.; K., S.; K., A. W. (1993) “Univariate Discrete Distributions.” 2. Ed. New York: John Wiley & Sons.
- KNUTH, D. (1997) “The art of computer programming.” 3. Ed. Addison-Wesley Professional.”
- MANSFIELD, K. C.; A., J. L. (2010) “Computer Networking from LANs to WANs: Hardware, Software, and Security.” Boston: Course Technology, C. L.
- SESHADRINATHAN, K.; S, R.; B., A. C.; C., L. K. (2010) “Study of Subjective and Objective Quality Assessment of Video.” *IEEE Transactions on Image Processing*.
- TAO, S.; A. J.; G., R. (2007) “Real-time monitoring of video quality in IP networks.” *IEEE/ACM Transactions on Networking*.
- WINKLER, S. (2007) “Video quality and beyond.” *15th European Signal Processing Conference, IEEE*: 150 - 153
- YATES, R. D.; G., D. (2004) “Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers.” 2. Ed. New York: J. W.
- ZENNARO, M.; C., E. S., K. R.; R., A. A.; C., K. R. (2006) “Scientific Measure of Africa’s Connectivity” *.Information Technologies and International Dev.*: 55-64