

UNIVERSIDADE FEDERAL DO PARANÁ

AVALIAÇÃO DE MÉTODO PARA MELHORIA DA QUALIDADE DE VÍDEO  
*STREAMING* SOBRE IP UTILIZANDO ENVIO REDUNDANTE DE  
INFORMAÇÕES

CURITIBA

2014

ALEXANDRE AUGUSTO FIEDLER FONÇATTI

AMANDA BIACCHI ELOY

AVALIAÇÃO DE MÉTODO PARA MELHORIA DA QUALIDADE DE VÍDEO  
*STREAMING* SOBRE IP UTILIZANDO ENVIO REDUNDANTE DE  
INFORMAÇÕES

Dissertação apresentada como requisito parcial à obtenção do grau de Bacharel em Engenharia Elétrica, no Curso de Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Prof. Dr. Carlos Marcelo Pedroso

CURITIBA

2014

## TERMO DE APROVAÇÃO

ALEXANDRE AUGUSTO FIEDLER FONÇATTI

AMANDA BIACCHI ELOY

AVALIAÇÃO DE MÉTODO PARA MELHORIA DA QUALIDADE DE VÍDEO  
*STREAMING* SOBRE IP UTILIZANDO ENVIO REDUNDANTE DE  
INFORMAÇÕES

Monografia aprovada como requisito parcial para a obtenção do grau de Bacharel em Engenharia Elétrica, o Curso de Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná, pela seguinte banca examinadora:

---

Prof. Dr. Carlos Marcelo Pedroso

Orientador – Departamento de Engenharia Elétrica, UFPR

---

Prof. Dr. Evelio Martín García Fernández

Departamento de Engenharia Elétrica, UFPR

---

Prof. Dr. Luis Henrique Assumpção Lolis

Departamento de Engenharia Elétrica, UFPR

Curitiba, 30 de Dezembro de 2014

## **RESUMO**

As aplicações de vídeo são atualmente responsáveis por grande parte do tráfego da Internet e este número só tende a crescer nos próximos anos. Este crescimento de demanda requer o desenvolvimento de novas técnicas para que a qualidade de experiência do usuário seja a melhor com a maior eficiência possível. Neste trabalho foi proposta a utilização do protocolo SCTP cuja propriedade de transmissão multicaminhos permite o uso de um caminho secundário para transmissão de pacotes importantes em redundância. A porcentagem de redundâncias enviadas é ponderada de acordo com os tempos de atraso fim-a-fim medidos em redes de transmissão reais. Os testes foram realizados por simulações computacionais e a qualidade final do vídeo é avaliada pelas métricas objetivas PSNR e SSIM. Os resultados demonstram que o envio de redundâncias é efetivo, diminuindo a degradação do vídeo causada por atrasos da rede.

## **ABSTRACT**

Video applications are nowadays responsible for most of the Internet traffic and this number tends to grow in the next years. This growth demands the development of new techniques to improve the user's quality of experience with the best efficiency possible. In this work, it was proposed the use of the SCTP protocol, whose multistreaming property allows the use of a second path for transmission of important packets in redundancy. The amount of redundant information to be sent is weighted according to the measured end-to-end delay. The tests were made using computational simulations and end-to-end delay data sampled in real networks. The final video quality is evaluated by the objective metrics PSNR and SSIM. The results show that the method is effective and reduces the impairments on perceived video quality caused by network delays.

## Lista de Figuras

Figura 1 - (a) Divisão de um quadro em macroblocos (b) Amostras croma e luma (MANOEL, 2007). .....	16
Figura 2 - Artefatos de vídeo. (a) erro de fatia (b) erro de fatia com partes perdidas realçadas (c) blocking (GREENGRASS; EVANS; BEGEN, 2009).....	21
Figura 3 - Processo de avaliação do SSIM (WANG et. al., 2004) .....	28
Figura 4 - Porcentagem de redundância utilizada de acordo com limites superior e inferior. ....	33
Figura 5 - Fluxograma do método utilizado .....	35
Figura 6 - Resultado dos testes para rede 3G com alto atraso, valores de PSNR a esquerda, valores para redundância a direita .....	37
Figura 7 - Resultado dos testes para rede 3G com baixo atraso, valores de PSNR a esquerda, valores para redundância a direita .....	38
Figura 8 - Resultado dos testes para rede 3G com médio atraso, valores de PSNR a esquerda, valores para redundância a direita .....	39
Figura 9 - Resultados de qualidade para segundo Teste, valores de PSNR a esquerda, valores para redundância a direita .....	41
Figura 10 - Resultados de qualidade para diferentes estruturas de GoP, valores de PSNR a esquerda, valores para redundância a direita .....	42
Figura 11 - Resultados dos testes de qualidade para diferentes valores de $\alpha$ . 43	
Figura 12 - PSNR após aplicação do método. No eixo esquerdo estão indicados os valores de PSNR e no eixo direito a porcentagem de redundância utilizada. ....	44
Figura 13 - SSIM após aplicação do método. No eixo esquerdo estão indicados os valores de SSIM e no eixo direito a porcentagem de redundância utilizada. ....	45
Figura 14 - Resultados do PSNR para os testes utilizando a rede Wi-Fi como caminho primário.....	47

Figura 15 - Resultados do SSIM para os testes utilizando a rede Wi-Fi como caminho primário..... 48

## **Lista de Tabelas**

Tabela 1 – Tipos de chunk de dados (STEWART, 2007).....	23
Tabela 2- Comparação entre PSNR, SSIM e MOS (ZINNER et. al., 2010) .....	29
Tabela 3 - Limites para primeiro teste .....	36
Tabela 4 - Resultado dos testes para rede 3G com alto atraso .....	37
Tabela 5 - Resultado dos testes para rede 3G com baixo atraso.....	38
Tabela 6 - Resultado dos testes para rede 3G com médio atraso .....	39
Tabela 7 - Limites para segundo teste .....	40
Tabela 8 - Resultados de qualidade para segundo Teste .....	40
Tabela 9 - Resultados de qualidade para diferentes estruturas de GoP .....	41
Tabela 10 - Resultados dos testes de qualidade para diferentes valores de $\alpha$	42
Tabela 11 - Resultados dos testes de qualidade com os vídeos.....	44
Tabela 12 - Valores médios dos resultados obtidos.....	48



## LISTA DE SIGLAS

3G - Third Generation

4G - Fourth Generation

ADSL - Asymmetric Digital Subscriber Line

B - Bidirectionally Predictive Coded Picture

CIF - Common Intermediate Format

FPS - Frames per Second

GOP - Group of Pictures

GVT - Global Village Telecom

HD - High Definition

HSPA - High Speed Packet Access

I - Intra Coded Picture

ICMP - Internet Control Message Protocol

IETF - Internet Engineering Task Force

IP - Internet Protocol

ITU - International Telecommunication Union

LTE - Long Term Evolution

MOS - Mean Opinion Score

SD - Standard Definition

MPEG - Moving Picture Experts Group

MSE - Mean Squared Error

MTU - Maximum Transmission Unit

NTSC - National Television System Committee

P - Predictive Coded Picture

PAL - Phase Alternating Line

PMR - Path Maximum Retransmissions

PR-SCTP - Partially Reliable Stream Control Transport Protocol

PSNR - Peak signal-to-noise ratio

QoE - Quality of Experience

QoS - Quality of Service

RFC- Request for Comments

RGB - Red Green Blue

RTO - Retransmission Timeout

RTP - Real-Time Transport Protocol  
RTT - Round Trip Time  
SCTP - Stream Control Transmission Protocol  
SI - Switching I  
SIGTRAN - Signalling Transport  
SP - Switching P  
SRTT - Smoothed Round Trip Time  
SSIM - Structural Similarity Index  
TCP - Transmission Control Protocol  
UDP - User Datagram Protocol  
UMTS - Universal Mobile Telecommunication System  
VoIP - Voice over Internet Protocol  
WCDMA - Wide-Band Code-Division Multiple Access  
WiMax - Worldwide Interoperability for Microwave Access

# SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 Contexto.....	11
1.2 Objetivos .....	12
1.3 Estratégia Adotada .....	12
1.3 Estrutura do trabalho .....	13
2 CONCEITOS FUNDAMENTAIS .....	14
2.1 MPEG .....	14
2.1.1 Padrão de cores .....	14
2.1.2 Quadros MPEG .....	15
2.1.3 Compressão .....	17
2.1.3.1 Subamostragem .....	17
2.1.3.2 Redundância espacial .....	18
2.1.3.3 Redundância temporal .....	18
2.1.4 Distorções .....	19
2.2 Protocolo SCTP .....	21
2.2.1 Seleção de Caminho no SCTP.....	24
2.2.2 Método Delay-Centric.....	25
2.2.3 Extensão PR-SCTP.....	25
2.3. Métricas de avaliação de qualidade de vídeo .....	26
2.3.1 PSNR .....	27
2.3.2 SSIM.....	28
2.3.3 MOS .....	29
3 METODOLOGIA .....	30
3.1 Obtenção de dados.....	30
3.2 Ferramentas utilizadas.....	30

3.3 Funcionamento do Programa.....	31
4 RESULTADOS.....	36
4.1 Ajuste do Método .....	36
4.1.1 Obtenção dos limites <i>LI e LS</i> .....	36
4.1.2 Primeiro teste .....	36
4.1.3 Segundo teste .....	40
4.1.4 Verificação do melhor valor de GoP.....	41
4.1.5 Verificação do $\alpha$ para SRTT .....	42
4.3 Resultados Obtidos .....	43
4.3.1 Coastguard.....	45
4.3.2 Akyio.....	45
4.3.3 Football.....	46
4.3.4 Bridge close.....	46
4.3.5 Big Buck Bunny .....	46
4.3.6 Rede Wi-Fi como primária.....	47
5 CONCLUSÃO .....	49
REFERENCIAS .....	50

# 1 INTRODUÇÃO

## 1.1 Contexto

O desenvolvimento de dispositivos como *smartphones e tablets* com sua capacidade de suportar aplicações multimídia em tempo real tornou aplicativos com suporte a vídeo uma das ferramentas mais utilizadas atualmente. O estudo "*Visual Networking Index*" conduzido pela Cisco apontou que 63% do tráfego da internet no Brasil em 2013 foi causado por aplicações de vídeo e este tráfego deverá triplicar até 2018 sendo responsável por 82% de tráfego online (ZAIA, 2014).

Tais aplicativos requerem maiores capacidades das redes de telecomunicações. Redes Wi-Fi padrão 802.11 tem uso limitado porque não podem fornecer recursos necessários para manter a qualidade na transmissão de vídeo, principalmente em função de sua relativamente pequena área de cobertura e por sua limitação de banda a medida que o número de usuários aumenta. Redes 3G, como WCDMA (*Wide-Band Code-Division Multiple Access*) e UMTS (*Universal Mobile Telecommunications System*), fornecem uma melhor conexão com usuário e uma maior cobertura de rede, porém a taxa de transferência desta tecnologia não é adequada para transmissão de vídeos em alta definição. Já as redes 4G, como LTE (*Long Term Evolution*) e WiMax, provem confiabilidade de conexão semelhante ao 3G e maior taxa de transmissão, podendo chegar a taxas de 100 MB/s.

Além disso, o alto estresse imposto a estas redes atualmente acaba criando tempos de atraso que torna inconveniente a visualização de vídeos em *streaming*, ou seja, durante processos onde o vídeo esteja sendo entregue em tempo real ao reprodutor de mídia do usuário (HARYADI, 2012). Estes tempos de atraso podem causar danos a qualidade da imagem pela perda de pacotes, cuja severidade dependerá do tipo da informação que este pacote carrega. A perda de quadros-chave para reconstrução do vídeo pode causar até o congelamento do vídeo, enquanto a perda de quadros que não são usados como referência a nenhum outro causam distorções menores.

## 1.2 Objetivos

O objetivo geral do trabalho é investigar os benefícios obtidos com o uso de transmissão redundante com o protocolo SCTP (*Stream Control Transmission Protocol*) para melhorar a qualidade de experiência de reprodução de vídeo *streaming* em situações onde estão disponíveis redes Wi-Fi ou ADSL (*Asymmetric Digital Subscriber Line*) e redes móveis.

Os objetivos específicos são: (I) estabelecer uma estratégia para transmissão redundante, (II) estabelecer parâmetros adequados de configuração para aplicação planejada, (III) analisar o desempenho do sistema proposto utilizando simulações computacionais.

## 1.3 Estratégia Adotada

Para melhoria da qualidade de experiência de usuário na visualização de vídeo *streaming*, propomos o envio de pacotes de vídeo redundantes por um segundo meio de transmissão.

O método para transmissão de informações redundantes será proposto através do estudo de características de codificação utilizado pelo MPEG-4 (*Moving Pictures Experts Group*) parte 10 (H.264). O método será especificado através de algoritmos. A proposta origina-se de um projeto de mestrado em execução no Programa de Pós-Graduação em Engenharia Elétrica da UFPR, pelo mestrando Carlos Gouvêa.

Especificamente, serão analisados os tempos de atraso das redes em locais de alto tráfego e serão estudados quais segmentos de informação devem ser enviados em redundância para uma melhoria na qualidade percebida pelo usuário. Os tempos de atraso serão coletados de redes Wi-Fi, ADSL e redes móveis pelo teste de *ping*, que mede o tempo que uma mensagem leva para ir e voltar de um servidor específico. Serão utilizados nos testes vídeos de diferentes resoluções e diferentes situações em cena (muito e pouco movimento), disponibilizados pelo *Video Quality Experts Group*, comumente utilizados em pesquisas na área para os testes.

Os testes serão então implementados por uma rotina de programação desenvolvida na plataforma *shell* do Linux. Os vídeos originais serão comprimidos no padrão MPEG-4 parte 10 e será simulada sua

transmissão pelas redes, considerando os valores de atraso anteriormente obtidos. A transmissão será feita em duas situações: (I) envio normal em cada uma das redes e (II) envio normal por uma rede e envio de pacotes importantes em redundância por uma segunda rede. A quantidade de pacotes redundantes enviados será ponderada de acordo com a qualidade da rede principal. Os vídeos serão remontados e recodificados para o formato original após a transmissão em cada caso. A análise de qualidade será feita métricas objetivas de qualidade PSNR (*Peak Signal-to-Noise Ratio*) e SSIM (*Structural Similarity Index*), que analisam a relação sinal-ruído de pico e a diferença estrutural dos vídeos, respectivamente. Com os resultados obtidos será avaliada a influência do método proposto na qualidade final do vídeo e finalmente serão apresentadas as conclusões obtidas.

### **1.3 Estrutura do trabalho**

Este trabalho está organizado da seguinte forma. O capítulo 2 apresenta os conceitos fundamentais, como o padrão MPEG (*Moving Picture Experts Group*) e suas principais características, os efeitos que perdas ocasionam à qualidade de experiência do usuário, as características do protocolo SCTP, seu método de seleção de caminhos e sua extensão PR-SCTP (*Partially Reliable Stream Control Transmission Protocol*) e as métricas de avaliação de qualidade de vídeo utilizadas. O capítulo 3 apresenta a metodologia usada, incluindo a descrição da amostragem de dados e a base de vídeos, bem como os procedimentos de simulação. O capítulo 4 apresenta os resultados obtidos e a interpretação dos mesmos. Finalmente, o capítulo 5 apresenta as conclusões e trabalhos futuros.

## 2 CONCEITOS FUNDAMENTAIS

### 2.1 MPEG

Um quadro digital de uma imagem em escala cinza de 640 x 480 pixels de definição padrão (SD – *standart definition*) que usa 8 bits para representar cada pixel, consome 2,45 Mbits de memória. Se esta imagem for atualizada com uma razão de 30 quadros por segundo, o requerimento de taxa de transmissão deste vídeo seria de aproximadamente 70Mbps, valor que cresce significativamente quando adicionada informação de cor. A taxa de transferência de bits de um vídeo não comprimido é de tipicamente 270Mbps para SD e 1,485 Gbps para recursos de alta definição (HD) (GREENGRASS; EVANS; BEGEN, 2009). Deste modo, o espaço necessário para o armazenamento e os recursos de rede necessários para a transmissão de um vídeo sem compactação são grandes, surgindo a necessidade do desenvolvimento de padrões de compressão de vídeos digitais.

O MPEG, “*Motion Picture Experts Group*”, é um grupo de especialistas criado em 1988 para criar padrões de compactação de vídeo e áudio. Dos padrões estabelecidos pelo grupo surgiram métodos de compressão que fazem uso de diversas técnicas que resultam em impressionantes taxas de compressão com uma boa qualidade de imagem final. Atualmente existem 5 padrões para este método, sendo eles: MPEG-1, MPEG-2, MPEG-4, MPEG-7 e o MPEG-21 (WATKINSON, 2004). O MPEG-4 parte 10, um dos padrões comumente utilizados atualmente, oferece maior eficiência na compressão de vídeos do que as versões anteriores, reduzindo a taxa de transmissão do vídeo anteriormente mencionado para 3 Mbps, para SD, e 9 Mbps para HD (GREENGRASS; EVANS; BEGEN, 2009).

#### 2.1.1 Padrão de cores

Existem várias formas de representar cores digitalmente e a escolha do sistema utilizado influencia diretamente na eficiência de compactação de um vídeo.

Um dos métodos mais comuns de representação é o sistema RGB (*Red-Green-Blue*), utilizado em diversos tipos de telas e monitores, onde as



imagens são formadas a partir da combinação de três cores primárias: vermelho, verde e azul. Outro sistema existente é o Y'CbCr (também denominado YUV), em que as componentes utilizadas para formação de cores nas imagens são a luminância (Y), que define a intensidade luminosa ou o brilho, a croma azul (Cb) e a croma vermelha (Cr).

As componentes R, G e B possuem uma alta correlação entre si, resultando em redundâncias em sua representação. Além disso, o olho humano tem uma limitação física que o torna mais sensível à luminosidade do que a cores. Por estes motivos, a compressão de vídeos utiliza sistemas do tipo luminância e croma como o Y'CbCr.

Com transformações lineares e não lineares é possível converter a representação do sistema de cores do RGB para o Y'CbCr, o que permite uma primeira compactação ao reduzir as informações de cor, com redução praticamente imperceptível na qualidade subjetiva percebida pelo usuário.

O padrão MPEG-4 H.264/AVC considera que os vídeos de entrada são codificados com formato Y'CbCr.

### 2.1.2 Quadros MPEG

Um vídeo é constituído por quadros (*frames*) apresentados de maneira sequencial ao longo do tempo. Um quadro é subdividido em macroblocos, de 16x16 pixels, que contem 16x16 amostras luma (Y') e as associadas amostras croma (8x8 amostras Cb e 8x8 amostras Cr). Os macroblocos são dispostos em fatias (*slices*) com o mesmo tipo de quadro, ou seja, uma fatia I contém apenas macroblocos do tipo I, por exemplo (RICHARDSON, 2003).

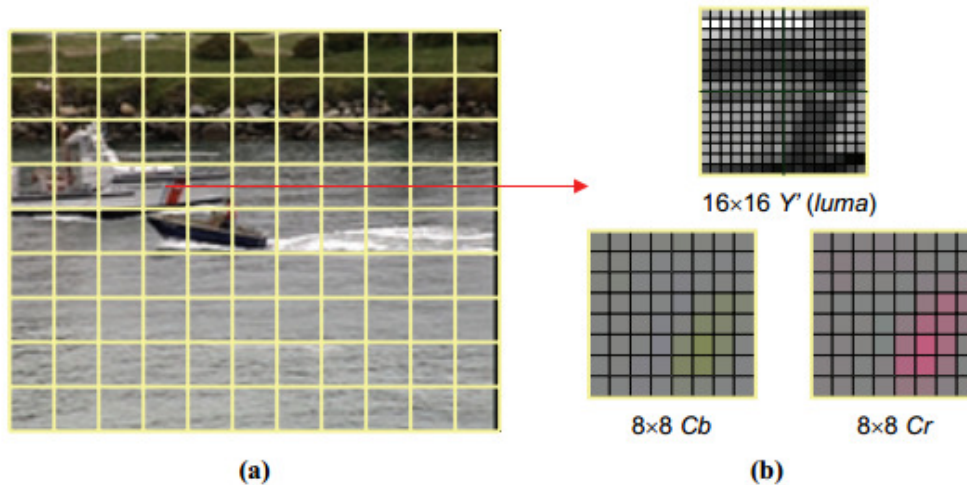


Figura 1 - (a) Divisão de um quadro em macroblocos (b) Amostras cor e luma (MANOEL, 2007).

A primeira imagem de uma sequência é codificada sem referência à nenhuma outra imagem e é chamada de quadro I (*intra-frame*) e nela é aplicada a redundância espacial, diminuindo o tamanho do quadro. Os quadros I são a base para a decodificação/descompressão de outros quadros. O quadro P (*predictive frame*) usa um quadro de referência, que pode ser um quadro I ou quadro P anterior, para prever o próximo quadro usando redundância temporal. O quadro B (*bi-predictive frame*) usa como referência o quadro I ou P anterior e posterior para compensação de movimento.

O MPEG-4 parte 10 tem como particularidade em relação aos padrões anteriores a inserção de dois novos quadros, o SP (*switching P*) e o SI (*switching I*), como solução a resiliência de erros, acesso aleatório e transferência entre fluxos binários codificados (SETTON; GIROD, 2005).

Os quadros são reunidos em um grupo de imagens (GoP – *Group of Pictures*). Um GoP inclui um quadro I e todos os quadros seguintes até o próximo quadro I. Para suportar os padrões NTSC (*National Television System Committee*) e PAL (*Phase Alternating Line*), o GoP tem tipicamente 12 e 15 quadros, respectivamente, mas também é comum o uso de GoP de 30 quadros (GREENGRASS; EVANS; BEGEN, 2009).

Existem vários tipos de estrutura de GoP, que devem ser escolhidos de acordo com o vídeo a ser codificado e com as condições do meio de transmissão, já que um vídeo com um maior número de quadros P terá uma menor compressão. Tipicamente, um GoP com 15 quadros contém um quadro

I, quatro quadros P e dez quadros B. A estrutura do GoP é descrita pelo número de quadros e pelo número de quadros B entre dois quadros P, por exemplo, a representação "15:2" significa que o GoP tem 15 quadros e dois quadros B separando cada quadro P.

### 2.1.3 Compressão

A compressão é baseada em três técnicas básicas:

- Subamostragem, que reduz as informações de cores que não são percebidas pelo olho humano;
- Redundância espacial, que baseado na propriedade em que pixels de um mesmo quadro são relacionados, remove as informações redundantes dentro do próprio quadro;
- Redundância temporal, que remove a informação redundante entre quadros diferentes.

#### 2.1.3.1 Subamostragem

A subamostragem é a redução da taxa de amostragem dos componentes de crominância em relação aos de luminância, eliminando parte da informação sem impacto à percepção humana.

Os formatos mais utilizados para codificação de vídeos são os seguintes:

- 4:4:4 – para cada 4 amostras de luminância (Y), existem 4 amostras de crominância azul (Cb) e 4 amostras de crominância vermelha (Cr), ou seja, não há subamostragem;
- 4:2:2 – para cada 4 amostras de Y na direção horizontal, existem 2 amostras de Cb e 2 amostras de Cr - verticalmente as amostras Y e Cb e Cr tem a mesma resolução;
- 4:2:0 – para cada 4 amostras de Y, existe uma amostra de Cb e uma amostra de Cr - as amostras de crominância possuem metade da resolução horizontal e vertical das amostras de luminância.

O formato 4:2:0 é o mais utilizado para o padrão MPEG-4 H.264/AVC (SILVA, 2007).

### 2.1.3.2 Redundância espacial

A compressão espacial se baseia na redundância espacial ou intra-quadros, ou seja, na semelhança existente entre pixels espacialmente vizinhos em um mesmo quadro. A redundância pode ser reduzida através de uma codificação intra-quadros. Aplica-se a transformada discreta do cosseno aos blocos de *pixels* da imagem, obtendo-se uma matriz de coeficientes em frequência. Como as frequências mais importantes para o olho humano são as mais baixas, é realizada uma divisão da matriz de coeficientes por uma matriz de quantização que possui valores maiores para frequências mais altas. Os resultados são truncados para o valor inteiro mais próximo. O coeficiente que resultar no valor 0 não será transmitido, levando a uma compactação com perdas (SILVA, 2007). O MPEG recomenda o uso de uma série de matrizes de quantização para obter uma boa relação entre qualidade da imagem e compactação. No entanto, o usuário pode especificar a matriz de quantização durante o processo de codificação.

### 2.1.3.3 Redundância temporal

A compressão temporal se baseia na redundância temporal ou inter-quadros, isto é, na similaridade que um quadro apresenta temporalmente com seu vizinho. Muitos *pixels* de um quadro para outro não mudam de valor, como no caso de um mesmo fundo de cena, ou tem uma pequena variação causada por uma mudança de iluminação, ou o *pixel* apenas se desloca de um quadro para o outro, em uma situação de movimento na imagem (SILVA, 2007).

Com base nestas características, um quadro pode buscar em quadros previamente codificados e reconstruídos informações que permitam sua reconstrução, eliminando redundâncias e reduzindo o número de bits utilizados. São chamados de quadros de referência aqueles usados para geração do quadro predito.

Uma das técnicas utilizadas para predição de quadros é a compensação do movimento existente na cena, que consiste em subtrair os valores das amostras de um quadro em relação a um quadro anterior ou posterior. Se os blocos forem semelhantes, o valor do bloco resultante será zero, reduzindo o número de bits utilizados.

Portanto, primeiramente e necessário que se determine o movimento do quadro, operação feita pela Estimção de Movimento, que busca o deslocamento de uma determinada região do quadro atual em relação a um quadro de referência. Este deslocamento pode ser causado pelo movimento da câmera ou do próprio objeto, ou pode ser apenas uma mudança de percepção do usuário causada por uma mudança de iluminação, por exemplo. A diferença entre a posição do objeto do quadro atual e do quadro de referência é chamada de vetor de movimento.

Como a estimção de movimento considerando objetos inteiros seria complexa e a análise considerando cada *pixel* individualmente seria de alto custo computacional, decidiu-se a utilização de blocos de *pixels* chamados macroblocos para a comparação. Assim, a estimção de movimento é feita pela busca do bloco de referência mais similar ao bloco atual.

Estimado o movimento e determinado o vetor de movimento é feita a compensação de movimento e é construído o quadro atual predito pela combinação de blocos provindos de quadros de referência.

No padrão H.264 existem dois tipos de predição temporal. A predição para quadros P utiliza apenas um quadro de referência e a predição para quadros B pode utilizar até dois quadros, que podem ser passados ou futuros na ordem de apresentação (FERREIRA, 2009).

Outra técnica é a interpolação da compensação de movimento, em que um sinal é codificado com uma resolução temporal baixa (tipicamente 1/2 ou 1/3 da taxa de quadros) e o sinal com resolução total é obtido pela interpolação do sinal de baixa resolução e a adição de um termo de correção. O sinal a ser reconstruído por interpolação é obtido pela adição do termo de correção a uma combinação de uma referência passada e futura (GALL, 1991).

Uma das vantagens da interpolação temporal é que mesmo que uma área não esteja presente na predição passada, ela será considerada pois o algoritmo usa também a predição futura, e que a codificação e a predição são separadas, portanto o erros não são propagados.

#### 2.1.4 Distorções

A qualidade final de um vídeo recebido pela Internet depende de vários parâmetros, entre eles a qualidade da rede. O estudo "*Effects of Internet*

*Path Selection on Video-QoE: Analysis and Improvement*" (VENKATARAMAN; CHATTERJEE, 2014) aponta que a ocorrência de degradações causadas por falhas da rede de transmissão, como um atraso fim-a-fim maior que 500ms, podem ser melhoradas pelo uso de caminhos alternativos de envio pela Internet.

Artefatos de vídeos são distorções visíveis em vídeos em relação ao original, causadas pela perda de quadros de um GoP, erros de codificação ou compressão excessiva. Nem todos os artefatos prejudicam o usuário da mesma maneira, a qualidade é só considerada como inaceitável em casos que os quadros-chave são danificados. Quando estes quadros-chave são rapidamente recuperados a aceitação do usuário muda. Em vídeos com baixa movimentação apenas quadros I são considerados como inaceitáveis. Já em vídeos com mais movimento, quadros I e P são considerados como chave.

Segundo o artigo "*Not All Packets Are Equal part 2*" (GREENGRASS; EVANS; BEGEN, 2009), que analisa o efeito visual da perda de pacotes IP na transmissão de vídeos MPEG, existem quatro tipos de artefatos:

- Erro de fatia, que ocorre quando pelo menos um pacote IP contendo quadros I, P ou B é perdido, como demonstrado pelas Figuras 2a e 2b. Se um pacote de referência é perdido, o erro se propaga pelo vídeo até o próximo GoP.
- *Blocking* ou pixelamento, ocorre quando um quadro de referência é perdido, ou seja quadro P ou I. Se esta perda ocorre durante um trecho de intenso movimento no vídeo, os quadros subsequentes não terão a informação necessária para decodificação, pixelando o vídeo, conforme na Figura 2c. O erro é normalizado quando um quadro I completo é decodificado.
- *Ghosting*, é a perda de pacotes I próximo a uma mudança de cenário, que resulta em um "fantasma" na tela. Isso se deve ao fato de, com a perda do quadro I, a imagem continuar aparecendo com referência ao quadro I anterior. A medida que os quadros P e B do novo GoP chegam, novas informações são inseridas no cenário antigo, formando tal efeito. A

integridade da imagem volta quando um novo quadro I sem perdas é recebido.

- Quadro congelado, é a perda de vários quadros em sequência que causa o congelamento da imagem no ultimo quadro recebido. Este estado permanece que novos quadros sejam recebidos.



Figura 2 - Artefatos de vídeo. (a) erro de fatia (b) erro de fatia com partes perdidas realçadas (c) blocking (GREENGRASS; EVANS; BEGEN, 2009)

O artigo define também o impacto da perda de cada quadro na qualidade do vídeo.

Quadros I, que não fazem referência à nenhum outro quadro, são portanto os de maior impacto. Como os dois últimos quadros B de um GoP também fazem referência ao próximo quadro I, em um GoP de 15:2 por exemplo a perda de um quadro I impacta 17 outros quadros do vídeo.

A importância dos quadros P é dependente de sua posição no GoP, já que os quadros do final do GoP são baseados em quadros do início (YOON; ZHANG; BANERJEE, 2014). Desta forma, um dano a um quadro P inicial só será recuperado até o envio do próximo quadro I, tornando-o quase tão significativo quanto um quadro I. Adicionalmente, o impacto na experiência do usuário será pior quanto maior o movimento na cena do vídeo.

Como quadros B não são referência a nenhum outro, a perda deste só afetará o quadro atual, causando uma pequena distorção que será vista pelo usuário por aproximadamente 1/30 de segundo.

## 2.2 Protocolo SCTP

O protocolo *Stream Control Transport Protocol* (SCTP) é um protocolo orientado a conexão, padronizado pela RFC 4960. O protocolo foi

criado pelo grupo de pesquisadores do transporte de sinalização (SIGTRAN) do IETF (*Internet Engineering Task Force*), e foi originalmente projetado para superar as limitações do protocolo TCP (*Transmission Control Protocol*) para transporte de sinalização da telefonia.

O SCTP é um protocolo de camada de transporte, como o TCP e o UDP (*User Datagram Protocol*). O protocolo implementa controle de congestionamento, utiliza o conceito de orientação à mensagens, permite a transmissão ordenada ou não-ordenada de dados e também oferece serviço confiável ou não confiável (KIM; KOH, 2007). Adicionalmente, o protocolo SCTP tem suporte à serviços de múltiplos fluxos, denominado multistreaming, que permite utilizar uma para transmissão de dados de diversos fluxos independentes.

O *multihoming* permite que o ponto de destino possua endereços IP alternativos que podem ser usados em caso de falhas. Esta característica é usada para aumentar a resiliência da rede e fornecer qualidade de serviço (QoS) para aplicações em tempo real, como vídeo *streaming* ou voz sobre IP (*VoIP*) (KELLY; MUNTEAN; PERRY, 2004).

Um pacote SCTP é constituído por um cabeçalho (*header*) comum e diversas porções (*chunks*). Um pacote IP pode conter vários *chunks de dados do SCTP*.

O *header* comum tem 12 bytes e contém a porta de origem e destino, um *tag* de verificação, que é um número inteiro gerado para que o receptor identifique que o pacote pertence à associação atual e não é resquício de uma associação passada, e um *checksum*, para detecção de erros de transmissão.

Os *chunks* são iniciados pela identificação do tipo do *chunk*, e contem também as *flags*, utilizadas para funcionalidades dependentes do tipo, o tamanho, e o *payload* do *chunk* (*user data*), também dependente do tipo.

Atualmente são padronizados 13 tipos de *chunk* pela RFC 4960, conforme descrito na Tabela 1.



Tabela 1 – Tipos de chunk de dados (STEWART, 2007)

Identificador	Tipo do <i>chunk</i>
0	DATA
1	INIT
2	INIT-ACK
3	SACK
4	HEARTBEAT
5	HEARTBEAT-ACK
6	ABORT
7	SHUTDOWN
8	SHUTDOWN-ACK
9	ERROR
10	COOKIE-ECHO
11	COOKIE-ACK
12	ECNE (Reservado para <i>Explicit Congestion Notification Echo</i> )
13	CCWR (Reservado para <i>Congestion Window Reduced</i> )
14	SHUTDOWN-COMPLETE
15 - 62	Reservado pelo IETF
63	Extensão de chunk (definido pelo IETF)
64 - 126	Reservado pelo IETF
127	Extensão de chunk (definido pelo IETF)
128 - 190	Reservado pelo IETF
191	Extensão de chunk (definido pelo IETF)
192 - 254	Reservado pelo IETF
255	Extensão de chunk (definido pelo IETF)

O *chunk* DATA é usado para o envio de dados em uma associação.

O INIT é usado para iniciar uma associação com o receptor, sendo o primeiro *chunk* a ser enviado na conexão. O INIT ACK é a resposta recebida em seguida, confirmando que a associação foi estabelecida.

O SACK informa ao transmissor dos *chunks* DATA e quais *chunks* foram recebidos.

O HEARTBEAT é enviado periodicamente para os diferentes endereços IPs negociados durante a inicialização para identificar se estes caminhos estão disponíveis. O HEARTBEAT ACK é a mensagem de confirmação recebida, indicando que a conexão funciona apropriadamente. Por padrão, as mensagens de HEARTBEAT são enviadas com período de 30 segundos.

O ABORT *chunk* é usado para abortar uma associação, em casos de erros irrecuperáveis.

O SHUTDOWN é usado quando um dos destinatários deseja encerrar a associação atual. O SHUTDOWN ACK é a confirmação que o

SHUTDOWN foi recebido. Antes de ser enviado, todos os dados de *buffer* devem ser enviados.

O SHUTDOWN COMPLETE é enviado pelo emissor do SHUTDOWN, confirmando que a associação foi finalmente encerrada.

O ERROR é um *chunk* enviado para informar o outro ponto da conexão que existem problemas na transmissão atual.

O COOKIE ACK é utilizado durante a inicialização da associação e deve preceder todos os *chunks* DATA e SACK. O COOKIE ECHO é usado pelo ponto que iniciou a conexão como resposta (WELTE, 2003).

### 2.2.1 Seleção de Caminho no SCTP

A conexão entre dois endereços IP no protocolo SCTP é chamada de associação. Durante uma operação normal, quando uma associação é estabelecida um dos endereços IP é escolhido como primário e todos os outros endereços são considerados como caminhos secundários, que são monitorados periodicamente, pelo envio de uma mensagem periódica de pulsação (*heartbeat*).

A detecção de falhas do protocolo SCTP é feita pela função *Retransmission Timeout* (RTO), que determina o tempo máximo que o SCTP aguarda pela resposta do pacote antes de retransmiti-lo, por padrão, determinado no início de uma associação com 3 segundos. Este valor é atualizado a cada nova medida de latência (*Round Trip Time - RTT*), ou seja, o tempo entre o envio de um pacote - de dados ou *heartbeat* - e o recebimento de sua confirmação (ACK). O cálculo do RTO é feito pela seguinte fórmula:

$$RTO = SRTT + 4 \cdot RTTVAR \quad (2.1)$$

sendo SRTT é a medida suavizada do RTT dada por:

$$SRTT_i = (1 - \alpha) \cdot SRTT_{i-1} + \alpha \cdot RTT_i \quad (2.2)$$

onde tipicamente  $\alpha = \frac{1}{8}$  (STEWART, 2007).

RTTVAR é a variação nas medidas do RTT dado por:

$$RTTVAR_i = (1 - \beta) \cdot RTTVAR_{i-1} + \beta \cdot |SRTT_{i-1} - RTT_i| \quad (2.3)$$

onde tipicamente  $\beta = \frac{1}{4}$  (STEWART, 2007).

O temporizador RTO é iniciado a cada novo segmento enviado. Por padrão o RTO mínimo é configurado em 1 segundo. Se a transmissão é confiável, caso seja expirado o tempo limite deste temporizador para o recebimento da confirmação de recebimento do segmento, uma retransmissão é realizada utilizando um caminho secundário e um contador de retransmissões é incrementado. Por padrão, valor mínimo do RTO seguinte é dobrado. Caso o número de retransmissões ultrapasse o limite, dado pelo parâmetro PMR (*Path Maximum Retransmissions*), um novo caminho primário é escolhido. Desta forma, por padrão, devem ser detectadas 6 falhas antes de alterar o caminho primário. O tempo mínimo para detecção desta falhas é de 63 segundos (WALLACE; SHAMI, 2012).

### 2.2.2 Método Delay-Centric

O algoritmo *delay-centric* foi desenvolvido por (KELLY et al., 2004) para melhorar a qualidade de transmissão de dados sensíveis a atrasos na rede. Ele monitora os tempos de atraso dos caminhos disponíveis e faz a transição automática para aqueles com menor latência. O monitoramento é feito através do atraso suavizado entre o envio e recebimento de ACK de um pacote de dados ou *heartbeat* (*SRTT - Smoothed Round Trip Time*).

### 2.2.3 Extensão PR-SCTP

O protocolo STCP quando utilizado no modo confiável envia todos os pacotes perdidos novamente. No entanto, em aplicações de tempo real existe um limite de tempo para o recebimento da informação. Desta forma, retransmitir todos os pacotes pode ser um desperdício de recursos, inclusive colaborando para o aumento do atraso e aumentando as perdas.

O PR-SCTP (*Partially Reliable Stream Control Transport Protocol*) é uma extensão do SCTP usada para fornecer uma transporte parcialmente confiável sobre uma conexão SCTP. Esta permite que o usuário especifique o quanto o serviço de transporte deve ser persistente na tentativa de enviar a mensagem ao receptor. Um exemplo de um serviço parcialmente confiável é especificado pela RFC 3758 (STEWART; RAMALHO, 2014) como "*timed*

*reability*". Este serviço permite que o usuário indique um limite de tempo para que o transmissor tente transmitir ou retransmitir a mensagem.

O PR-SCTP distingue a confiabilidade das mensagens de acordo com o seu "tempo de vida". Antes do tempo de vida de uma mensagem expirar, ela é considerada como confiável o protocolo fará tentativas de retransmissão, mas quando este expira, a mensagem é descartada. Quanto maior o tempo de vida, maior a confiabilidade da mensagem (WANG; JIN, 2003).

O tempo de vida é calculado por:

$$Tempodevida_N = R_{app} + \frac{L_N}{B_{midia} \cdot 1024} \quad (2.4)$$

onde  $R_{app}$  é o atraso de acordo com a taxa de controle da camada superior em milissegundos,  $L_N$  é o comprimento da mensagem da aplicação em bytes e  $B_{midia}$  é a taxa de transferência média da mídia em bits por segundo (KIM; KOH, 2007). O valor de  $R_{app}$  é ajustado de acordo com o estado do *buffer* de aplicação, atraso de rede e *jitter*. Não havendo um controle destas taxas, este valor será definido como zero.

Em vídeo *streaming*, pacotes com atraso maior que sua ordem de reprodução são inúteis, prejudicando a qualidade do vídeo e desperdiçando recursos de banda. Assim, a confiabilidade parcial é de grande auxílio neste tipo de aplicação, onde é possível configurar o tempo de vida para um limite de tempo em que o pacote ainda seja aplicável para a remontagem do vídeo. Neste trabalho não utilizaremos a política de retransmissão, pacotes atrasados serão descartados pois os mesmos já estão sendo enviados por uma segunda via em redundância. Entretanto este parâmetro pode ser utilizado para aprimoramento dos resultados.

A configuração de políticas diferentes de retransmissão para diferentes pacotes será o suporte para a priorização de pacotes contendo quadros I e P.

### 2.3. Métricas de avaliação de qualidade de vídeo

A avaliação da qualidade de vídeo pode ser feita através de métricas subjetivas ou objetivas. A avaliação subjetiva mede a percepção de usuários sobre a qualidade de reprodução do vídeo em um ambiente controlado. A

avaliação objetiva é realizada através de métodos numéricos. Os métodos objetivos podem ser classificados métodos sem referência, com referência completa ou com referência parcial (SERRAL-GRACI; CERQUEIRA, 2010). Os métodos com referência completa comparam o vídeo recebido com o vídeo transmitido. Os métodos sem referência utilizam apenas o vídeo recebido e parâmetros de desempenho de rede. Os métodos com referência parcial utilizam o vídeo transmitido e parte do vídeo recebido. Neste trabalho foram utilizadas duas métricas objetivas com referência completa, a relação sinal ruído de pico (PSNR) e o índice de similaridade estrutural (SSIM).

### 2.3.1 PSNR

O PSNR (*Peak-to-NoiseRatio*) é uma métrica utilizada para a avaliação da qualidade de uma imagem degradada, onde são numericamente comparados os estados inicial e final.

O PSNR é a razão entre a maior potência possível de um sinal e a potência do ruído que afeta a qualidade da sua representação. Por causa da variação dinâmica que muitos sinais tem, o PSNR é geralmente expressado em decibel.

O cálculo do PSNR é realizado através da seguinte relação:

$$PSNR = 20 \cdot \log \left( \frac{MAX_f}{\sqrt{MSE}} \right) \quad (2.5)$$

onde MSE (*Mean Squared Error*) é o erro quadrático médio, e  $MAX_f$  é o maior valor de sinal que existe na imagem original. Por exemplo, em uma imagem que utiliza oito bits por pixel o valor de  $MAX_f$  será 255. O erro quadrático médio indica a média quadrática da diferença pixel por pixel entre o vídeo original e o vídeo degradado, calculado da seguinte forma:

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [f(i,j) - g(i,j)]^2 \quad (2.6)$$

onde 'm' e 'n' representam respectivamente a resolução vertical e horizontal em pixels, a matriz ' $f(i,j)$ ' representa o vídeo original e a matriz ' $g(i,j)$ ', o vídeo degradado.

A desvantagem desta métrica é que são comparados valores estritamente numéricos, não levando em consideração as limitações humanas e a percepção real do usuário.

### 2.3.2 SSIM

O SSIM (*Structural Similarity Index*) usa medidas de distorção estrutural ao invés do erro, como o PSNR. Ele se baseia no fato que a visão humana é especializada em extrair informações estruturais do campo de visão. Assim, a medida da distorção estrutural deve fornecer uma melhor correlação à impressão subjetiva (WANG, 2006).

Os pixels são interdependentes entre si, principalmente quando estão espacialmente próximos. Estas dependências carregam importantes informações sobre a estrutura dos objetos da cena.

A luminância vista na superfície de um objeto é resultado da iluminação e da refletância, mas a estrutura do objeto independe da iluminação. Assim, para analisar a informação estrutural deste objeto, precisamos eliminar a influência da iluminação.

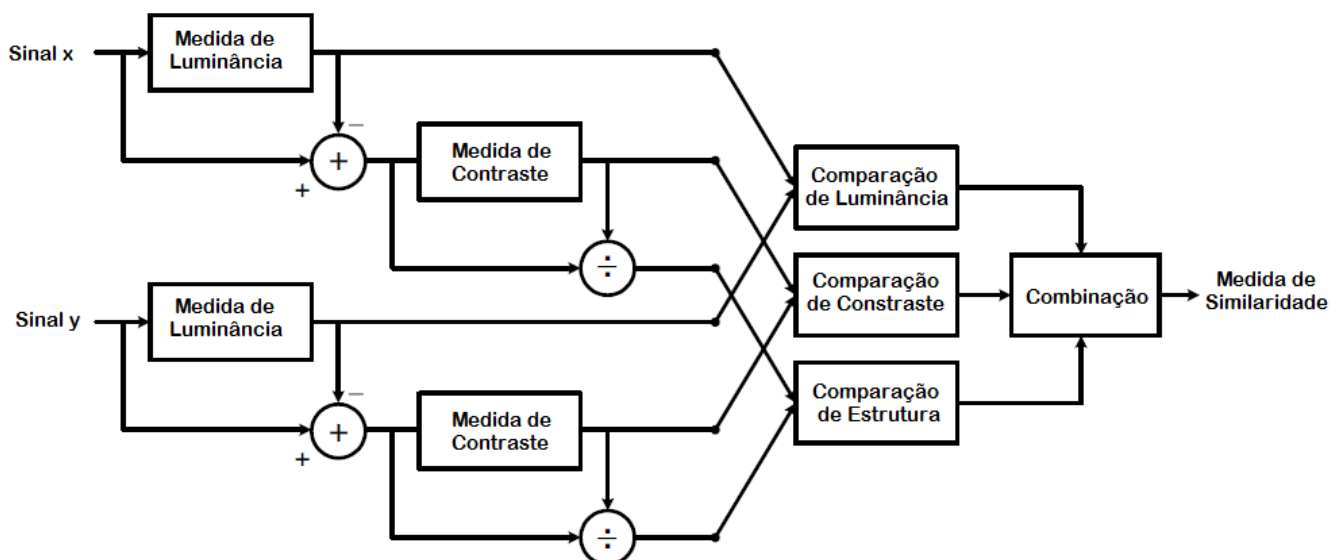


Figura 3 - Processo de avaliação do SSIM (WANG et. al., 2004)

O processo de avaliação do SSIM é ilustrado na Figura 3. Sendo x e y dois sinais da mesma imagem, um com perfeita qualidade e outro degradado, pode-se analisar a qualidade do segundo comparando sua similaridade com o

primeiro sinal. O sistema é dividido em três partes: luminância, contraste e estrutura. Primeiramente é estimada a intensidade média de cada sinal, para comparação da luminância. Em seguida, esta intensidade é removida dos sinais e estima-se e compara-se então o contraste. Então, para fins matemáticos, os sinais são normalizados e é comparada a estrutura. A medida de similaridade geral é finalmente obtida pela combinação da comparação dos três fatores. A escala varia de 0 a 1, indicando 0 nenhuma similaridade e 1 similaridade total (PANDE, 2013).

### 2.3.3 MOS

O MOS (*Mean Opinion Score*) é uma medida de qualidade subjetiva baseada na percepção humana para avaliação da qualidade de experiência (QoE) de um vídeo, padronizada pelo *International Telecommunication Union* (ITU) em (ITU-T, 2007a) e (ITU-T, 2007b).

O conceito de avaliação subjetiva de qualidade foi inicialmente proposto para avaliação de ligações telefônicas, onde indivíduos seriam colocados em um ambiente controlado e determinariam a qualidade de uma ligação segundo sua percepção. O mesmo conceito foi estendido para a avaliação de vídeo, com a metodologia de análise apresentada nas normas supracitadas. O avaliador deve indicar sua percepção de qualidade segundo uma escala que varia de 1 a 5, sendo 1 a pior qualidade e 5 a melhor.

A tabela 2 correlaciona os valores de SSIM, PSNR e MOS, com base em estudos anteriormente realizados (ZINNER et. al., 2010).

Tabela 2- Comparação entre PSNR, SSIM e MOS (ZINNER et. al., 2010)

<b>PSNR</b>	<b>SSIM</b>	<b>MOS</b>	<b>Significado</b>
$PSNR \geq 45$	$SSIM \geq 0,99$	5	Excelente
$33 \leq PSNR < 45$	$0,95 \leq SSIM < 0,99$	4	Bom
$27,4 \leq PSNR < 33$	$0,88 \leq SSIM < 0,95$	3	Regular
$18,7 \leq PSNR < 27,4$	$0,5 \leq SSIM < 0,88$	2	Pobre
$PSNR < 18,7$	$SSIM < 0,5$	1	Ruim

### 3 METODOLOGIA

#### 3.1 Obtenção de dados

Os vídeos de referência utilizados são arquivos disponibilizados pelo *Video Quality Experts Group* e comumente utilizados em pesquisas na área. Originalmente, os vídeos estão no padrão YUV 4:2:0, utilizando o formato 1080HD com resolução de 1920x1080 e 24 fps, e o formato CIF com resolução 352x288 e 25 fps.

Foram coletadas amostras de atraso em redes de acesso tipo ADSL, WIFI e 3G. A rede ADSL tem taxas de 10 Mbps e é operada pela GVT (*Global Village Telecom*). Foi utilizada a tecnologia HSPA+ da Claro na obtenção dos dados da rede 3G, uma evolução da terceira geração de redes móveis utilizada atualmente pela maioria das operadoras, com taxas médias de 3 Mbps. A rede Wi-Fi utilizada no estudo está localizada no Departamento de Engenharia Elétrica da UFPR, que utiliza o padrão IEEE802.11g, com taxas de 54 Mbps, e está conectada a uma rede Ethernet que por sua vez utiliza um enlace de 20 Gbps para conexão com a Internet.

Os dados foram coletados em horário comercial e o atraso fim-a-fim foi obtido pelo monitoramento do envio de pacotes ICMP (*Internet Control Message Protocol*) ECHO REQUEST/REPLY, utilizando o programa *ping* em um computador conectado à rede testada. Foi escolhido o servidor de vídeos do site *NETFLIX* para o teste, cujo endereço é 198.45.48.71. Estes dados foram obtidos através do trabalho de mestrado "Método para Melhoria da Qualidade de Transmissão de Vídeos sobre o Protocolo SCTP" feito por Alex Torres (TORRES, 2014).

#### 3.2 Ferramentas utilizadas

A simulação foi implementada por uma rotina de programação em *shell script*, na plataforma Linux.

O programa utiliza o Evalvid (KLAUE; RATHKE; WOLISZ, 2003), uma ferramenta para a avaliação da qualidade de vídeos transmitidos em redes de comunicação real ou simuladas. Além de medir parâmetros de qualidade de



serviços (QoS), como taxas de perdas, atrasos e *jitter*, também são fornecidas as métricas de avaliação padrão SSIM e PSNR.

Foram utilizadas as seguintes ferramentas, integrantes do Evalvid, cuja funcionalidade será descrita no próximo tópico:

- *ffmpeg*: permite a codificação de vídeos em diversos formatos, incluindo H.264 e YUV
- *MP4Box*: cria o arquivo apropriado para a transmissão
- *mp4trace*: realiza a transmissão do vídeo
- *ffprobe*: analisa vídeos codificados
- *etmp4*: faz a remontagem do vídeo utilizando como referência a captura de pacotes na origem e no destino
- PSNR: calcula o PSNR e o SSIM

### 3.3 Funcionamento do Programa

O funcionamento do programa é demonstrado pelo fluxograma da Figura 5. Primeiramente foi utilizado o conversor de áudio e vídeo *ffmpeg* para codificação do vídeo original em formato YUV420p para o formato m4v, com GoP 12:2. Então o programa *MP4Box* encapsula os bits do vídeo nos pacotes que serão transmitidos pela rede. Sem esta ferramenta, o servidor não saberia que informação colocar em cada pacote e portanto os mesmos não seriam criados corretamente. Foram utilizados pacotes com MTU (*Maximum Transmission Unit*) de 1460 bytes, tendo como saída um vídeo no formato mp4.

O *mp4trace* então recebe o vídeo em MPEG e origina o *trace* necessário para que o mesmo seja enviado pela rede, ou seja, a lista de frames e pacotes a ser enviados. O próprio programa faz o envio do vídeo via protocolo RTP/UDP.

Tendo estes dados em mãos, o próximo passo é a avaliação da quantidade de informação redundante é necessária a cada instante, para que não ocorram perdas de qualidade na recepção.

Um programa realizado pela equipe organiza a lista de quadros e calcula a quantidade total de quadros no vídeo e por GoP. As informações sobre os GoPs e a organização de quadros são extraídas pela ferramenta *ffprobe*, que reúne as informações de fluxos de mídia.

Total do GoP é calculado por:

$$T = \emptyset_I + (N \times \emptyset_P) + (O \times \emptyset_B) \quad (3.1)$$

Sendo:

$\emptyset_I =$  Tamanho médio do quadro I

$\emptyset_P =$  Tamanho médio do quadro P

$\emptyset_B =$  Tamanho médio do quadro B

$N =$  Número de quadros P por GOP

$O =$  Número de quadros B por GOP

Analisando os valores de atraso de rede obtidos, observa-se que, apesar de existir uma tendência do atraso de aumentar ou diminuir ao longo do tempo, existem picos em que os valores mudam abruptamente. Por este motivo, o atraso instantâneo foi suavizado pelo cálculo do SRTT (*Smoothed Round Trip Time*), utilizando uma média móvel exponencial (EMA), que pode ser visto também como um filtro passa baixa, definido da seguinte forma:

$$SRTT_i = (\alpha \times RTT_i) + [(1 - \alpha) \times SRTT_{i-1}] \quad (3.2)$$

Sendo:

$\alpha =$  coeficiente de suavização

RTT=atraso instantâneo

SRTT=Atraso instantâneo suavizado

A quantidade de informações que deve ser enviada em redundância é um importante aspecto pois o envio de informações em excesso desperdiça recursos de banda e a falta de informações prejudica a qualidade. A porcentagem de envio redundante foi obtida como uma função linear entre dois valores limite de atraso suavizado, conforme o gráfico apresentado na figura 4.

Além disso, na maioria dos serviços de *streaming* de vídeo em tempo real é utilizado um tempo limite máximo de  $t_{max} = 200ms$  para obtenção do pacote antes do descarte. Logo, a quantidade de informação redundante a ser enviada foi baseada na diferença do tempo limite ( $D_t$ ) e a SRTT instantânea, ou seja, quão próxima o atraso suavizado está do tempo limite para descarte.

$$D_t = t_{max} - SRTT \quad (3.3)$$

Estabelecido isso, o programa foi ajustado para calcular a quantidade de informações redundantes que deve ser enviada baseado na diferença  $D_t$ . Para isso foram utilizados dois valores para definição dos limites. Acima do limite superior ( $L_s$ ) considera-se que nenhuma redundância é necessária, enquanto abaixo do limite inferior ( $L_l$ ) 100% de redundância deve ser enviada. Entre os limites uma quantidade de redundância proporcional deve ser transmitida. O gráfico a seguir representa a quantidade de informação redundante a ser enviada em função da diferença  $D_t$ .

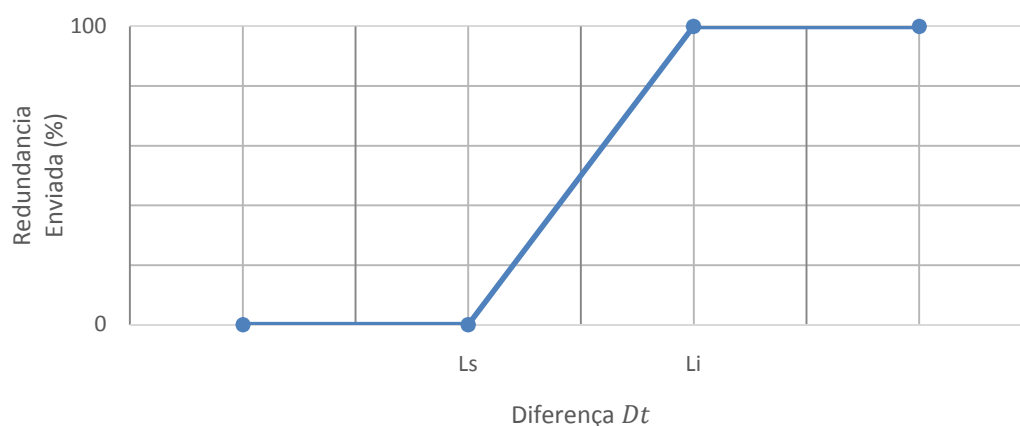


Figura 4 - Porcentagem de redundância utilizada de acordo com limites superior e inferior.

A eficácia do envio de redundâncias também é influenciado pelo tipo de quadro a ser transmitido. Como descrito anteriormente, quadros I e quadros P iniciais tem maior importância na reconstrução do vídeo e, portanto, devem ser priorizados, segundo a seguinte política:

- Caso 1: Informação redundante necessária por GoP menor que o tamanho do frame I
  - Enviar apenas os pacotes iniciais de cada frame I até obter a quantidade necessária.
- Caso 2: Informação redundante necessária por GoP maior que o tamanho do frame I, porém menor que a soma dos frames P do GoP.
  - Enviar o frame I, após isso enviar a quantidade de frames P necessária para obter a quantidade requisitada de

informação em caminho secundário, priorizando o envio dos frames P mais próximos do início do GoP.

- Caso 3: redundância necessária por GoP maior que a soma dos frames I e P.
  - Enviar todos os pacotes dos frames I e P, e após isso enviar pacotes B para completar a quantidade de informação necessária.

Desta forma, obtém-se a quantidade em bytes de cada tipo de quadro a ser enviada em redundância dentro de cada GoP através de:

$$Env_i = p \times \emptyset_I \quad , p < \frac{\emptyset_I}{T} \quad (3.4)$$

$$Env_P = \left( p - \frac{\emptyset_I}{T} \right) \times (N \times \emptyset_P) \quad , \frac{\emptyset_I}{T} < p < \frac{\emptyset_I + (N \times \emptyset_P)}{T} \quad (3.5)$$

$$Env_B = \left[ p - \frac{\emptyset_I + (N \times \emptyset_P)}{T} \right] \times (O \times \emptyset_B) \quad , p > \frac{\emptyset_I + (N \times \emptyset_P)}{T} \quad (3.6)$$

onde  $Env_i$  é a quantidade de informação do frame I por GoP a ser enviada em redundância,  $Env_p$  é a quantidade do total de informação de frame P por GoP,  $Env_B$  é a quantidade total de informação de frame B também por GoP e 'p' é a porcentagem de informação redundante. As variáveis  $\emptyset_I$ ,  $\emptyset_P$ ,  $\emptyset_B$ ,  $O$ ,  $T$ ,  $N$  e  $p$  são definidas anteriormente nesta seção.

Tendo esses valores calculados o programa simula o envio dos pacotes nas redes 3G, WiFi e ADSL com e sem redundâncias, levando em consideração o tempo de envio de cada pacote individual, capacidade do canal, atrasos de cada rede e o tamanho de cada pacote. O programa então remonta o vídeo m4v eliminando os pacotes cujo tempo de recebimento ultrapassou o tempo de espera limite, definido como 200ms, utilizando a ferramenta. Em seguida o vídeo é novamente convertido para o padrão YUV. Finalmente, calcula-se o PSNR e SSIM de cada vídeo enviado, obtendo a análise objetiva de qualidade de vídeo.

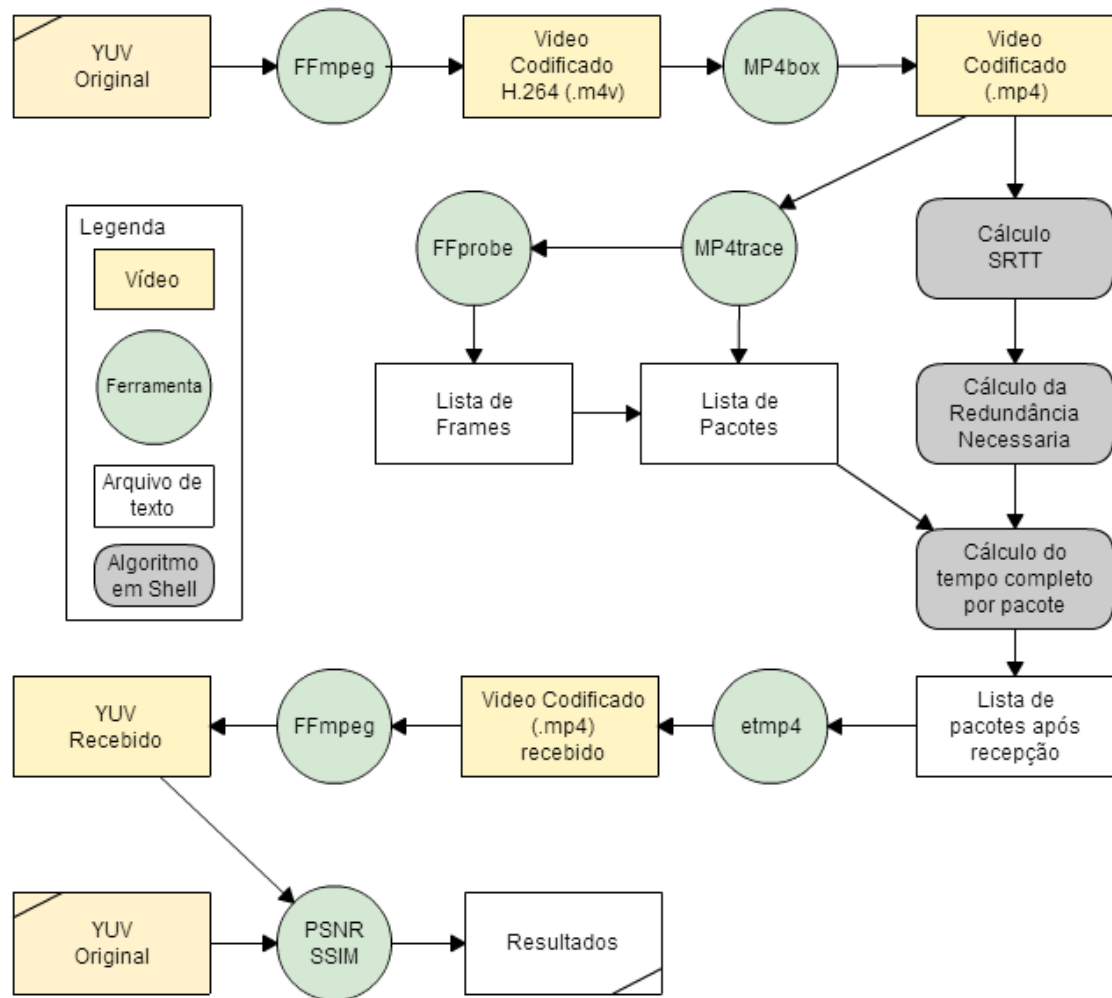


Figura 5 - Fluxograma do método utilizado

## 4 RESULTADOS

### 4.1 Ajuste do Método

#### 4.1.1 Obtenção dos limites $L_I$ e $L_S$

Um dos fatores importantes para determinar quanto de informação redundante o programa deve enviar é o limite da diferença entre o atraso máximo e o atraso suavizado. Este limite define a partir de qual valor e com qual grau será utilizado o envio de informações por caminho secundário.

Para obtenção destes valores foi elaborada uma rotina de testes com vídeos padrões.

#### 4.1.2 Primeiro teste

A princípio foi adotado um valor padrão de 50 ms de diferença entre os valores dos Limites. Após isso, foram feitas simulações onde os valores do Limite Superior e Inferior eram incrementados proporcionalmente, a fim de encontrar em qual período o programa encontrava a melhor eficiência, ou seja, a maior melhoria de qualidade de imagem com a menor quantidade de informações sendo enviadas em redundância. Os valores para os limites testados são apresentados na tabela 3:

Limite Superior (ms)	60	65	70	75	80	85	90	95	100	105	110
Limite Inferior (ms)	10	15	20	25	30	35	40	45	50	55	60

Tabela 3 - Limites para primeiro teste

Essa rotina foi repetida em situações de atraso baixo, médio e alto na rede 3G. Para a situação de alto atraso na rede houve uma influência mínima dos limites na rede como mostram a Tabela 4 e a Figura 6 a seguir:

Tabela 4 - Resultado dos testes para rede 3G com alto atraso

Lim. Sup.	60	65	70	75	80	90	95	100	105	110	
Lim. Inf.	10	15	20	25	30	40	45	50	55	60	
PSNR	3G	20,2	19,9	20,2	19,8	20,0	20,1	19,8	20,0	20,0	20,0
	WiFi	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	ADSL	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	3G+W	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9
	3G+A	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9	36,9
SSIM	3G	0,58	0,57	0,57	0,57	0,57	0,57	0,57	0,57	0,57	0,57
	WiFi	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	ADSL	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+W	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+A	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
Redundancia	99%	99%	99%	99%	99%	99%	99%	99%	99%	99%	

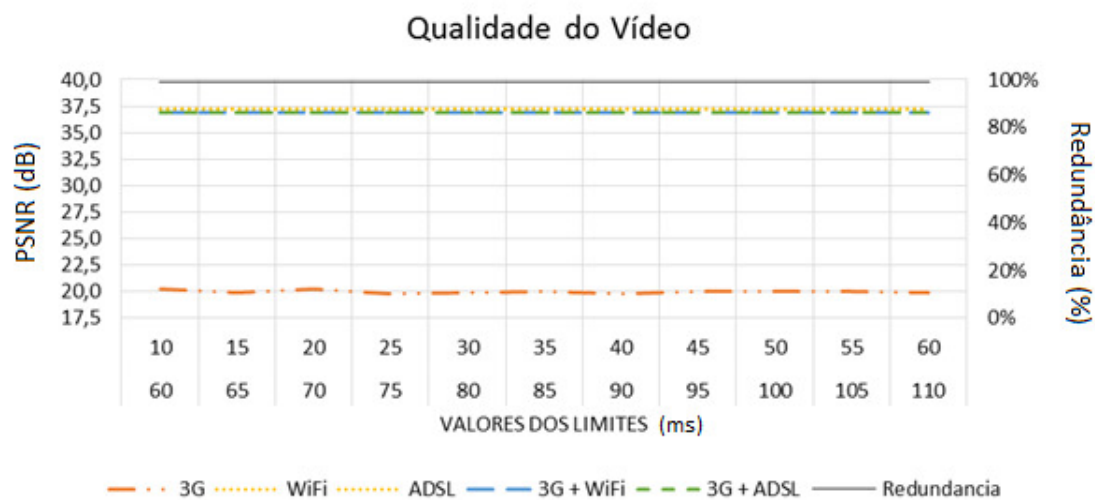


Figura 6 - Resultado dos testes para rede 3G com alto atraso, valores de PSNR a esquerda, valores para redundância a direita

Em sequência, para a situação de baixo tráfego na rede 3G, foi verificado que o método começa a enviar informações desnecessárias pelo caminho secundário a partir dos valores  $L_I = 30ms$  e  $L_S = 80ms$ , como mostram a tabela 5 e Figura 7 na sequência:

Tabela 5 - Resultado dos testes para rede 3G com baixo atraso

Lim. Sup.	60	65	70	75	80	90	95	100	105	110	
Lim. Inf.	10	15	20	25	30	40	45	50	55	60	
PSNR	3G	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	WiFi	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	ADSL	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	3G+W	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	3G+A	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
SISS	3G	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	WiFi	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	ADSL	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+W	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+A	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
Redundancia	0%	0%	0%	0%	0%	1%	4%	9%	12%	41%	

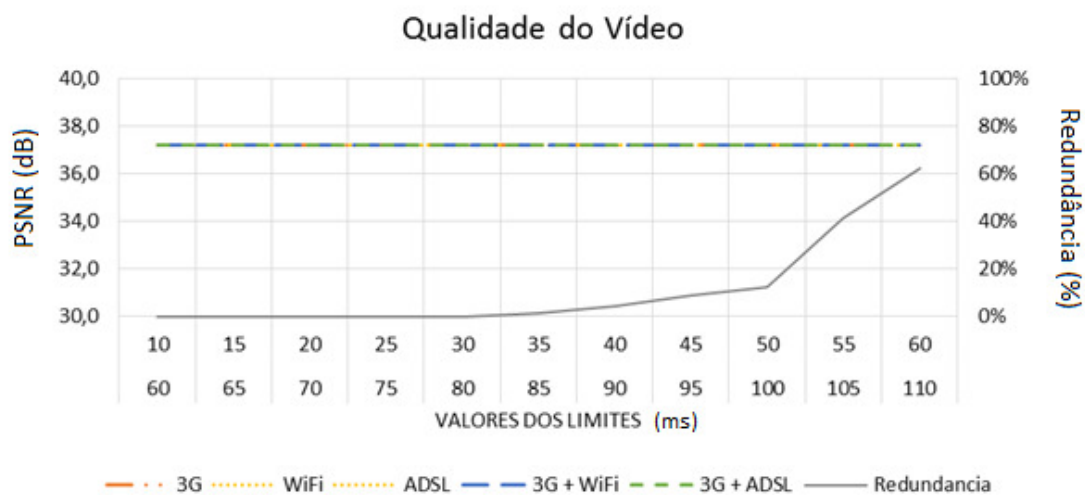


Figura 7 - Resultado dos testes para rede 3G com baixo atraso, valores de PSNR a esquerda, valores para redundância a direita

Enfim, com os resultados do teste com grau médio de atraso na rede 3G foi comprovado que com valores maiores para os limites era obtido uma melhor qualidade de reprodução nos vídeos com envio redundante, como vemos na tabela 6 e Figura 8 a seguir:



Tabela 6 - Resultado dos testes para rede 3G com médio atraso

Lim. Sup.	60	65	70	75	80	90	95	100	105	110	
Lim. Inf.	10	15	20	25	30	40	45	50	55	60	
PSNR	3G	35,6	35,6	35,6	35,5	35,5	35,6	35,5	35,6	35,5	35,5
	WiFi	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	ADSL	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	3G+W	35,4	35,4	36,3	36,3	36,7	36,7	36,6	36,6	36,6	37,1
	3G+A	35,5	35,5	36,3	36,3	36,7	36,7	36,6	36,6	36,6	37,1
SSIM	3G	0,93	0,93	0,93	0,92	0,92	0,93	0,92	0,93	0,92	0,92
	WiFi	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	ADSL	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+W	0,93	0,93	0,94	0,94	0,95	0,95	0,95	0,95	0,95	0,96
	3G+A	0,93	0,93	0,94	0,94	0,95	0,95	0,95	0,95	0,95	0,96
Redundancia	2%	5%	12%	23%	41%	52%	67%	72%	82%	91%	

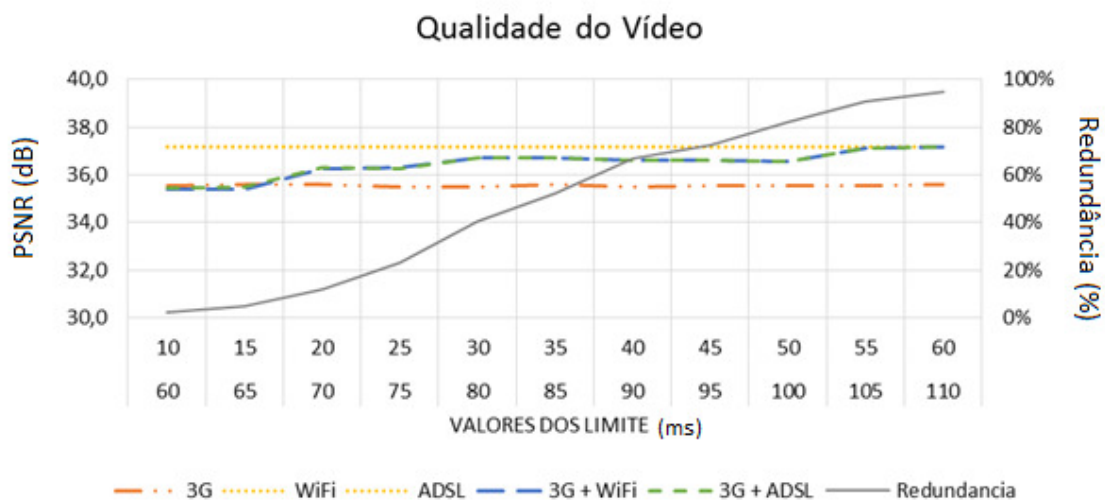


Figura 8 - Resultado dos testes para rede 3G com médio atraso, valores de PSNR a esquerda, valores para redundância a direita

Logo, o valores de  $L_I = 30ms$  e  $L_S = 80ms$  obtiveram a melhor qualidade de reprodução sem o envio de informações desnecessárias.

### 4.1.3 Segundo teste

Considerando que o método apresentou melhor eficiência com os valores de  $L_I = 30ms$  e  $L_S = 80ms$ , foi realizado uma segunda simulação para investigar qual intervalo mais eficiente entre os valores dos limites. Neste caso, foram testados valores para os limites inferior e superior apresentados na tabela 7:

Limite Superior (ms)	55	60	65	70	75	80	85	90	95	100	105
Limite Inferior (ms)	55	50	45	40	35	30	25	20	15	10	5
Delta Limite	0	10	20	30	40	50	60	70	80	90	100

Tabela 7 - Limites para segundo teste

O resultado encontrado indica que o método apresenta maior eficiência para valores de  $L_I = 30ms$  e  $L_S = 80ms$ , como pode-se averiguar pela tabela 8 e figura 9.

Tabela 8 - Resultados de qualidade para segundo Teste

Lim. Sup.	60	65	70	75	80	90	95	100	105	110	
Lim. Inf.	10	15	20	25	30	40	45	50	55	60	
PSNR	3G	35,6	35,6	35,5	35,5	35,5	35,5	35,6	35,5	35,5	35,6
	WiFi	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	ADSL	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2	37,2
	3G+W	36,3	36,3	36,3	36,3	36,3	36,7	36,7	36,7	36,7	36,7
	3G+A	36,3	36,3	36,3	36,3	36,3	36,7	36,7	36,7	36,7	36,7
SSIM	3G	0,93	0,92	0,92	0,92	0,92	0,92	0,93	0,92	0,92	0,93
	WiFi	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	ADSL	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
	3G+W	0,94	0,94	0,94	0,94	0,94	0,95	0,95	0,95	0,95	0,95
	3G+A	0,94	0,94	0,94	0,94	0,94	0,95	0,95	0,95	0,95	0,95
Redundancia	10%	21%	26%	31%	36%	41%	41%	46%	49%	50%	

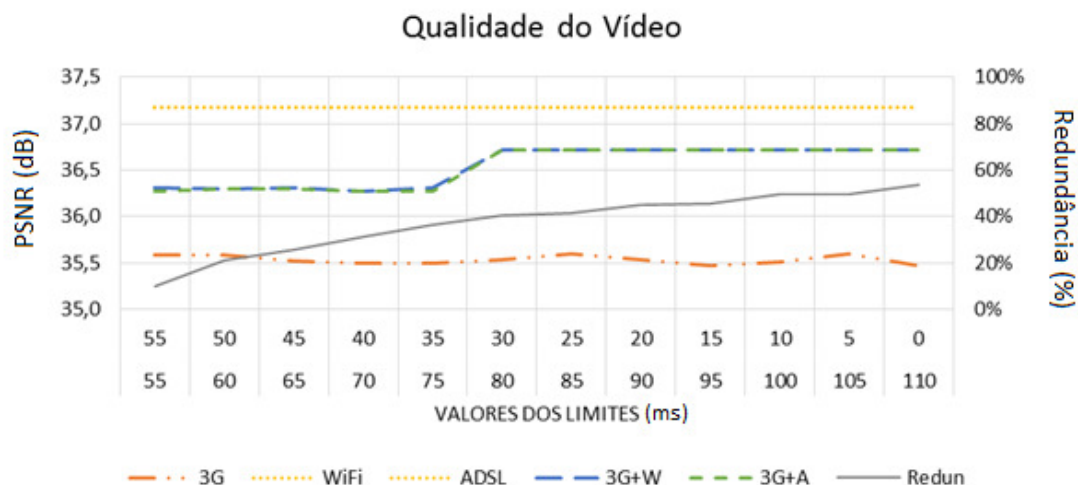


Figura 9 - Resultados de qualidade para segundo Teste, valores de PSNR a esquerda, valores para redundância a direita

#### 4.1.4 Verificação do melhor valor de GoP

Com a aquisição dos limites mais eficientes, foram testados diferentes valores para o GoP para demonstrar a eficácia do método com configurações diferentes. Foram feitos testes com os valores de 12:2, 25:2 e 50:2. No gráfico abaixo pode-se observar que independente do GoP utilizado o programa apresenta uma definitiva melhoria de qualidade. Porém a configuração 12:2 demonstra melhores resultados.

Tabela 9 - Resultados de qualidade para diferentes estruturas de GoP

		GoP	12:2	25:2	50:2
PSNR	3G		30,9	30,9	30,9
	WiFi		37,2	37,1	37,0
	ADSL		36,4	36,3	36,0
	3G+W		36,7	36,7	36,6
	3G+A		36,2	35,9	35,7
SSIM	3G		0,82	0,82	0,82
	WiFi		0,96	0,96	0,96
	ADSL		0,95	0,94	0,93
	3G+W		0,95	0,95	0,95
	3G+A		0,94	0,93	0,92
Redundancia			41%	42%	42%

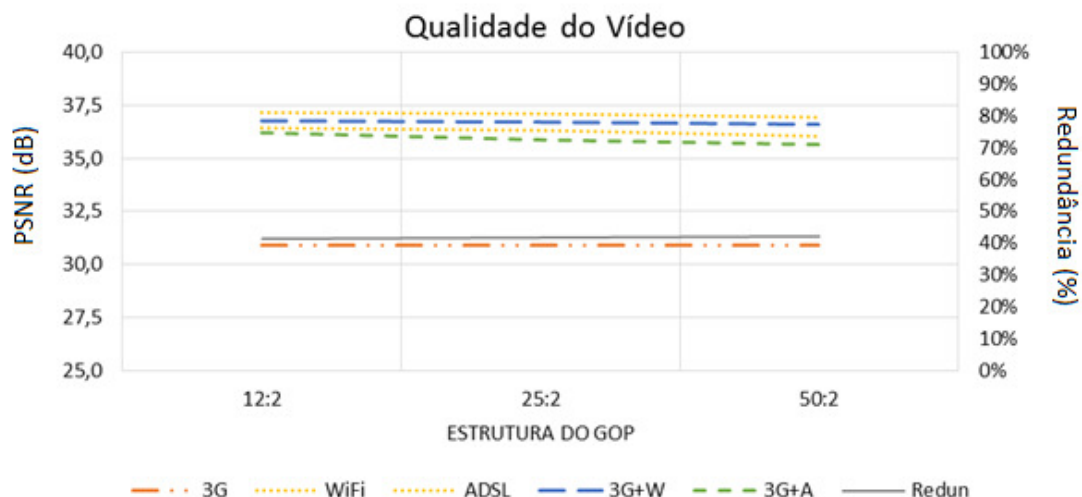


Figura 10 - Resultados de qualidade para diferentes estruturas de GoP, valores de PSNR a esquerda, valores para redundância a direita

#### 4.1.5 Verificação do $\alpha$ para SRTT

O  $\alpha$  é a constante que define o grau de suavização aplicada no cálculo da SRTT. Por último foi feito o ajuste desse valor, para descobrir qual grau de suavização apresentava uma melhor eficiência. Os valores testados foram de  $1/8$ ,  $1/6$  e  $1/4$ . Os resultados estão apresentados da tabela 10 e figura 11 a seguir:

Tabela 10 - Resultados dos testes de qualidade para diferentes valores de  $\alpha$

		GoP	1/8	1/6	1/4
PSNR	3G		35,5	35,5	35,6
	WiFi		37,2	37,2	37,2
	ADSL		37,2	37,2	37,2
	3G+W		36,7	36,3	36,3
	3G+A		36,7	36,3	36,3
SSIM	3G		0,92	0,92	0,92
	WiFi		0,96	0,96	0,96
	ADSL		0,96	0,96	0,96
	3G+W		0,95	0,94	0,94
	3G+A		0,95	0,94	0,94
Redundancia			41%	37%	36%

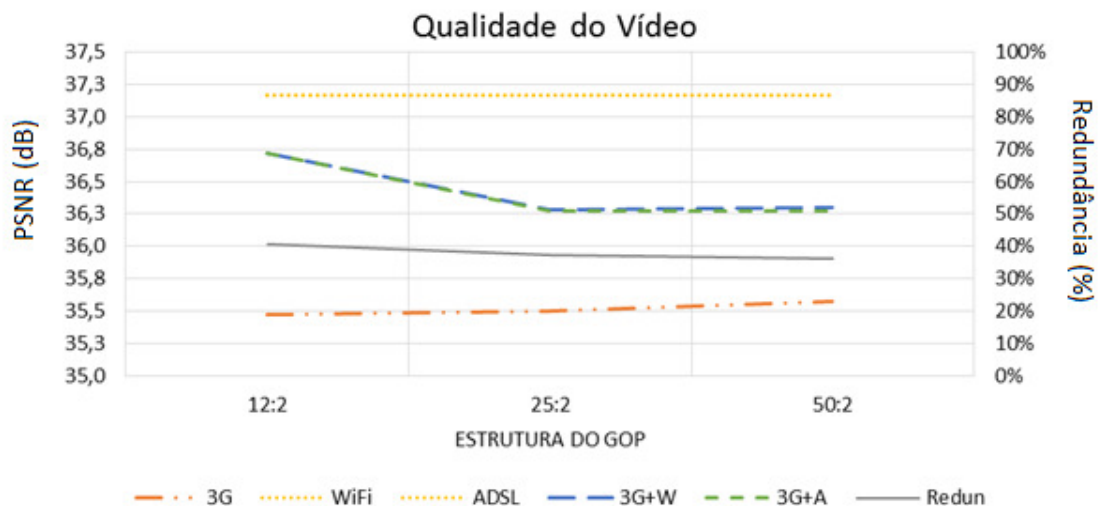


Figura 11 - Resultados dos testes de qualidade para diferentes valores de  $\alpha$

Observa-se que, embora mínimo, existe um ganho adicional de qualidade ao utilizarmos o valor de  $1/8$  para o  $\alpha$  ao criar os atrasos suavizados.

#### 4.3 Resultados Obtidos

Para a obtenção dos resultados foram utilizados 5 vídeos no formato YUV planar 4:2:0. Foram utilizados os vídeos Coastguard, Akyio, Football e Bridge Close, com configuração CIF, de 352x288, e o vídeo Big Buck Bunny na resolução HD1080, de 1920x1080. O vídeo em HD foi utilizado para averiguar como o método se comportaria com um vídeo de alta resolução.

A fim de possibilitar a generalização do resultado através de intervalos de confiança foram efetuadas 15 simulações com cada vídeo, utilizando tempos de atraso em momentos diferenciados da rede. Os resultados para os valores médios do PSNR e SSIM são apresentados na Tabela 11, assim como o valor máximo de PSNR e SSIM que o vídeo atinge após a compressão, indicado por "Base". A rede 3G foi utilizada como canal primário por apresentar uma maior variação nos tempos de atraso, fornecendo uma boa variedade de cenário para os testes. As redes WiFi e ADSL foram selecionadas como canal secundário para envio das informações redundantes.

Tabela 11 - Resultados dos testes de qualidade com os vídeos

		Coastguard	Akyio	Football	Bridge close	Big Buck Bunny
PSNR	Base	38,0	45,0	38,0	38,0	45,0
	3G	30,3 ± 3,9	36,6 ± 4,9	31,1 ± 3,5	33,1 ± 2,9	25,8 ± 4,6
	WiFi	35,1 ± 0,4	42,4 ± 0,5	36,2 ± 0,0	36,1 ± 0,2	41,2 ± 0,7
	ADSL	35,0 ± 0,6	42,3 ± 0,7	35,6 ± 0,8	35,7 ± 0,6	34,2 ± 2,2
	3G+WiFi	34,4 ± 0,6	41,7 ± 0,7	34,5 ± 1,2	35,5 ± 0,4	32,6 ± 4,1
	3G+ADSL	34,2 ± 0,7	41,5 ± 0,8	34,5 ± 1,1	35,5 ± 0,5	30,2 ± 3,2
SSIM	Base	0,96	0,99	0,96	0,95	0,96
	3G	0,80 ± 0,12	0,90 ± 0,07	0,82 ± 0,09	0,89 ± 0,06	0,84 ± 0,04
	WiFi	0,95 ± 0,01	0,98 ± 0,01	0,95 ± 0,00	0,94 ± 0,00	0,95 ± 0,01
	ADSL	0,94 ± 0,02	0,98 ± 0,01	0,94 ± 0,02	0,94 ± 0,01	0,91 ± 0,01
	3G+WiFi	0,93 ± 0,02	0,97 ± 0,01	0,91 ± 0,03	0,93 ± 0,01	0,89 ± 0,02
	3G+ADSL	0,92 ± 0,02	0,97 ± 0,01	0,91 ± 0,03	0,93 ± 0,01	0,88 ± 0,02
Redundancia Enviada	40% ± 25%	40% ± 24%	39% ± 21%	23% ± 18%	34% ± 23%	

A figura 12 apresenta os resultados de PSNR para os testes efetuados com os 5 vídeos citados.

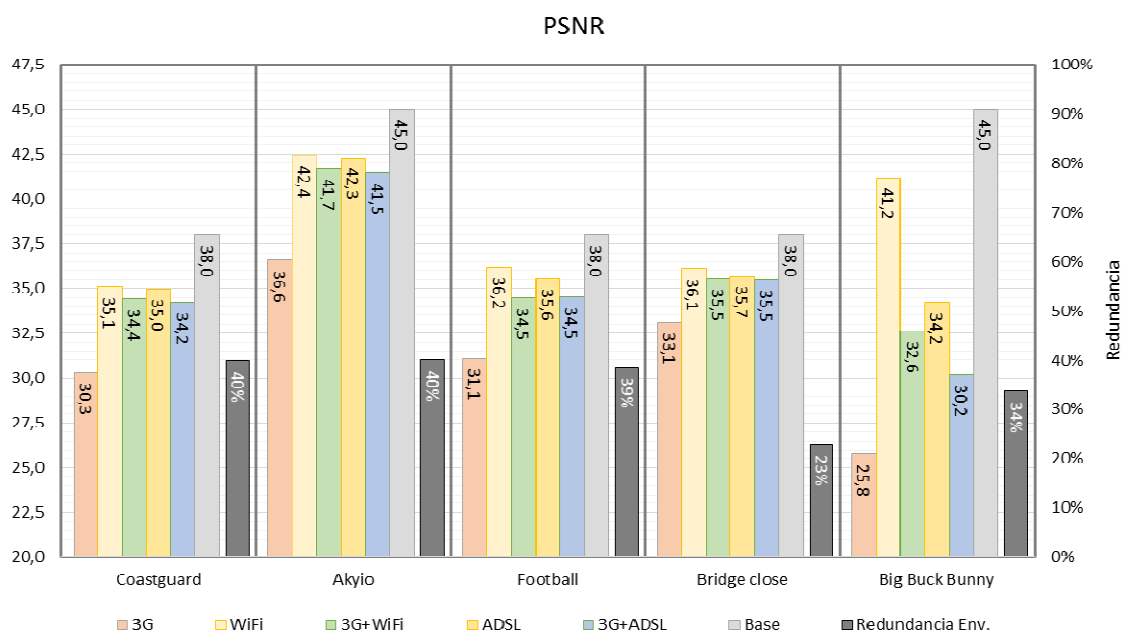


Figura 12 - PSNR após aplicação do método. No eixo esquerdo estão indicados os valores de PSNR e no eixo direito a porcentagem de redundância utilizada.

A figura 13 apresenta os resultados de SSIM para os testes efetuados com os 5 vídeos citados.

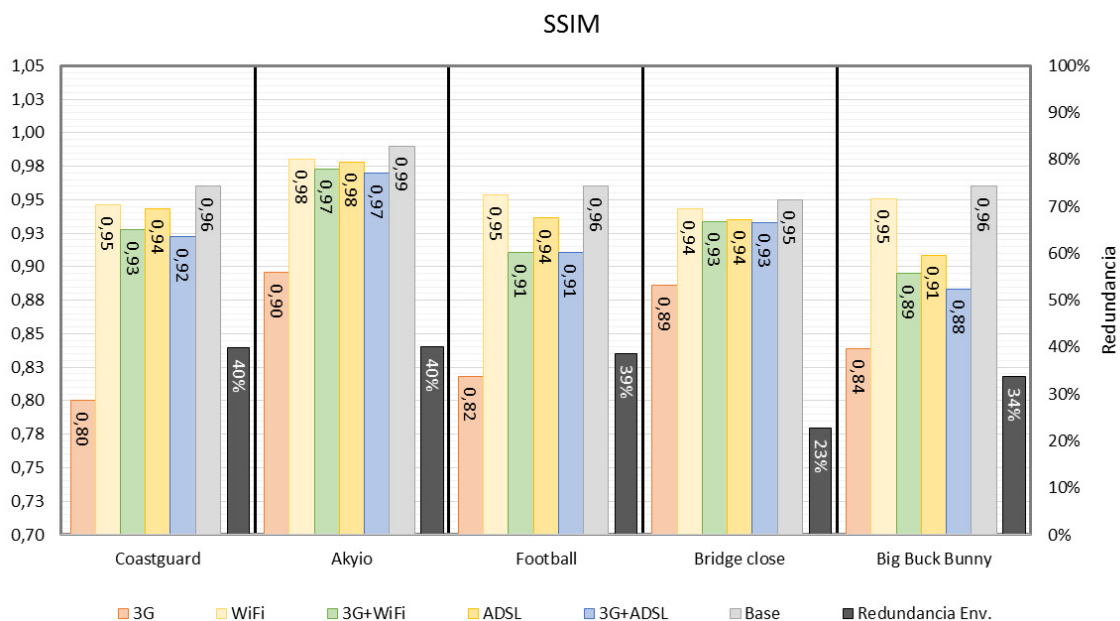


Figura 13 - SSIM após aplicação do método. No eixo esquerdo estão indicados os valores de SSIM e no eixo direito a porcentagem de redundância utilizada.

#### 4.3.1 Coastguard

O Coastguard.yuv foi examinado por ser um vídeo de taxa de movimentação média, favorecendo levemente o algoritmo de compressão. O vídeo apresenta resolução CIF com uma duração de 12 segundos.

Os resultados mostram bom ganho no PSNR e SSIM com uso da transmissão redundante no canal secundário, demonstrando uma boa eficiência do método em situações com atrasos moderados.

#### 4.3.2 Akyio

Akyio.yuv foi escolhido para fazer a análise de um vídeo que apresenta uma taxa de movimentação baixa, favorecendo muito a compressão. O vídeo está na resolução CIF e apresenta uma duração de 12 segundos.

Por apresentar uma baixa taxa de movimentação, obtemos uma PSNR base mais alta. Podemos observar também uma maior PSNR e SSIM, ainda assim com um ganho de qualidade similar, se comparado com o vídeo anterior.

#### 4.3.3 Football

O vídeo Football.yuv foi selecionado como um caso extremo por exibir uma alta variação de frames devido a cenas com muita movimentação, além de conter mais cortes de cena que os vídeos anteriores, o que dificulta o processo de compressão. O vídeo está em formato CIF e apresenta 10 segundos de duração.

Nota-se uma melhoria nos índices de qualidade dos vídeos com informações enviadas em redundância, elevando os índices de quase todas as passagens a PSNR e SSIM, assim como nos casos anteriores.

#### 4.3.4 Bridge close

O vídeo bridgeclose.yuv foi selecionado por ser mais longo que os vídeos anteriores e possuir maior tempo de duração. O vídeo apresenta baixa taxa de movimentação e não tem cortes de cena, apresentando uma duração de 30 segundos.

Como o esperado, mesmo tendo uma duração maior, o programa demonstrou resultados similares aos dos vídeos de baixa variação nos frames que analisamos anteriormente, embora um pouco mais suavizados. Isso ocorre pelo fato de o vídeo utilizar mais segundos para ser transmitido, homogeneizando o desvio padrão da série. Percebesse, novamente, uma melhora na qualidade dos vídeos enviados com redundância. Demonstrando a eficácia do programa com vídeos de durações maiores.

#### 4.3.5 Big Buck Bunny

Para análise de como o programa se portaria com um vídeo de maior resolução foi escolhido o vídeo bigbuckbunny.yuv, este vídeo apresenta uma resolução HD de 1920x1080 pixels e uma duração de 30 segundos.

Nota-se que a rede 3G apresentou diversos problemas para o envio de um vídeo com maior resolução, isso se dá pelo fato dos frames terem um tamanho muito maior devido a resolução HD, e assim ficarem mais suscetíveis a atrasos de rede, criando distorções e artefatos maiores na transmissão.

A aplicação do método resultou em melhoria de qualidade em todos os casos, com ganho médio de 3,5 dB na PSNR.



#### 4.3.6 Rede Wi-Fi como primária

Durante os testes, a rede 3G foi selecionada como rota primária devido a sua alta taxa de variação nos tempos de atraso, fazendo com que, ao rodar simulações simultâneas fosse obtidos resultados para diferentes estados da rede primária.

Porém, uma situação com maior aplicação do método seria o uso de uma rede Wi-Fi como caminho primário, e em momentos de estresse desta rede, procurar auxílio através do envio redundante pela rede 3G.

Com o objetivo de se obter uma resposta sobre a eficiência do método nessa situação foram selecionados intervalos de tempo onde a rede Wi-Fi apresentava valores altos de atrasos. Foram selecionados 40 intervalos com duração de 15 segundos onde a rede apresentava essas condições e aplicados na simulação de envio do vídeo Coastguard.yuv. Os resultados são apresentados abaixo.

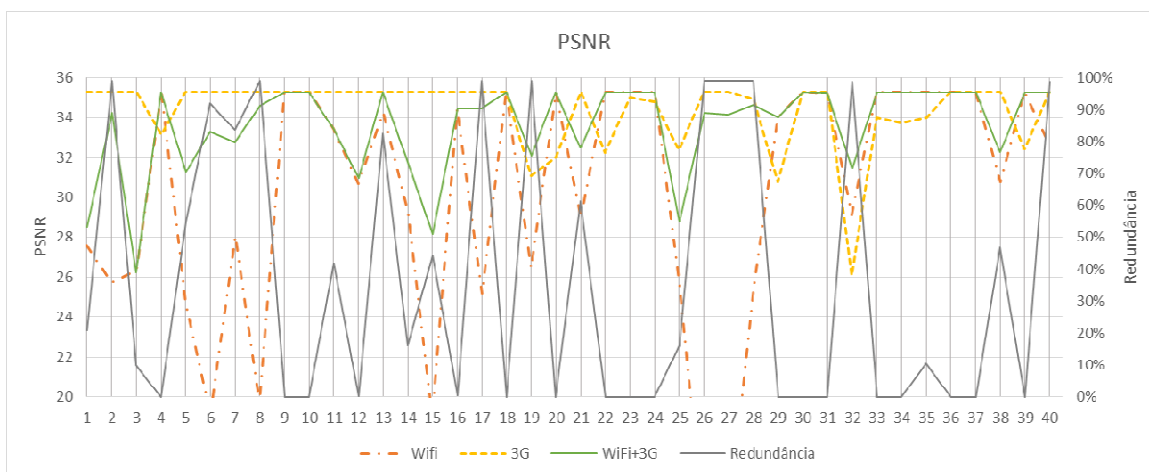


Figura 14 - Resultados do PSNR para os testes utilizando a rede Wi-Fi como caminho primário

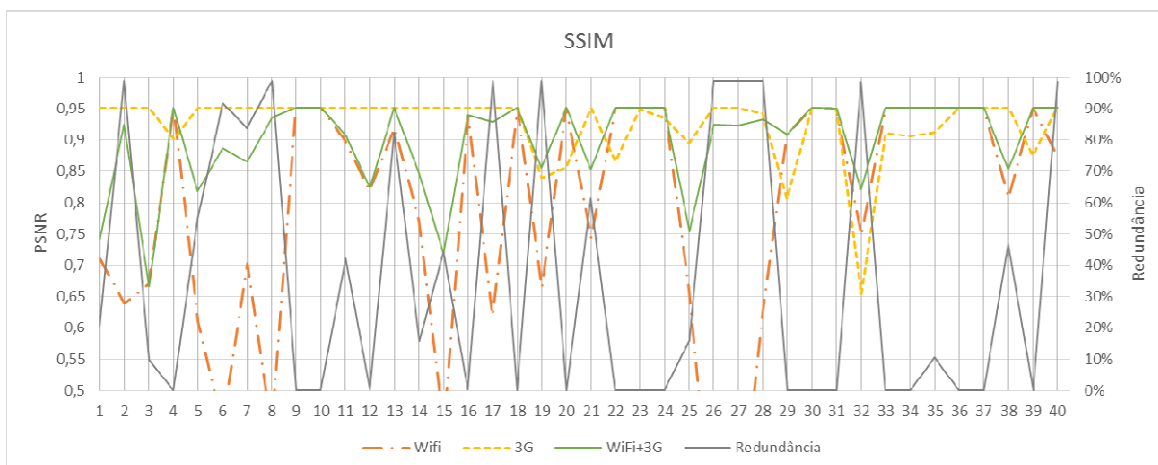


Figura 15 - Resultados do SSIM para os testes utilizando a rede Wi-Fi como caminho primário

Tabela 122 - Valores médios dos resultados obtidos

	WiFi	3G	WiFi+3G
PSNR	30,1 +- 2,4	34,3 +- 0,7	33,6 +- 0,9
SSIM	0,8 +- 0,1	0,9 +- 0,0	0,90 +- 0,03

Redundância	37% +- 16%
-------------	------------

Observa-se que com o auxílio do envio redundante de pacotes importantes pela rede 3G pode-se obter uma melhora média de 4dB na PSNR do vídeo, enviando em média 37% da informação.

Lembrando que esse cenário representa momentos de alto estresse na rede Wi-Fi, nota-se uma boa melhoria na qualidade de vídeo e experiência do usuário, em troca de um moderado envio de informações redundantes.

## 5 CONCLUSÃO

Este trabalho tinha como objetivo avaliar a melhoria de qualidade de reprodução de vídeos através do envio redundante de informações selecionadas utilizando um canal secundário.

Após análise dos resultados foi verificado que o método de envio de informações redundantes pode seguramente ajudar a melhorar a qualidade de reprodução para transmissão de vídeos sobre IP.

Através dos testes foi obtido uma melhoria média de  $4,11 \pm 1,77$  dB na qualidade dos vídeos com a aplicação do método. Com um envio de, em média,  $35 \pm 10\%$  de informações redundantes por um canal secundário.

O método experimentado nesse trabalho apresenta a possibilidade de implementação junto a um método de escolha e seleção permanente de rotas. Assim, enquanto o primeiro visaria melhorar a qualidade do vídeo a ser reproduzido instantaneamente através do envio redundante, o segundo método teria tempo de analisar os caminhos e escolher qual rede deveria ser utilizada como rota primária.

## REFERENCIAS

[1] ZAIA C., "Tráfego na internet dobrará até 2018 no Brasil, diz pesquisa da Cisco". Junho de 2014. Disponível em: <<http://www.valor.com.br/empresas/3581866/trafego-na-internet-dobrara-ate-2018-no-brasil-diz-pesquisa-da-cisco>> - Acesso em 24/10/2014

[2] HARYADI S., PRAMUDITA A. "QoS Measurement of Video Streaming Services Using Agregation Method". Telecommunications Systems, Services and Applications (TSSA), 2012. 7th International Conference, pg 97-101, Outubro de 2012.

[3] GREENGRASS, J., EVANS, J. e BEGEN, A. C. "Not all packets are equal, part 1: The impact of network packet loss on video quality" IEEE Internet Computing, vol. 13, p. 70–75, Março 2009.

[4] WATKINSON, J. "The MPEG Handbook: MPEG-1, MPEG-2, MPEG-4 Broadcasting and communications." Ed. Ilustrada: Taylor & Francis, 2004.

[5] RICHARDSON; I. E. G. "H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia". Editora Wiley, 2003.

[6] SETTON E., GIROD B. "Video Streaming with SP and SI Frames." Disponível em <<http://web.stanford.edu/~bgirod/pdfs/SettonVCIP2005.pdf>> - Acesso em 15 de outubro de 2014.

[7] ASCENSO J. "Codificação Escalável de Vídeo com Elevada Granularidade". Universidade Técnica de Lisboa. Disponível em <[http://www.img.lx.it.pt/publications/Theses/JoaoAscenso\\_MSc.pdf](http://www.img.lx.it.pt/publications/Theses/JoaoAscenso_MSc.pdf)> - Acesso em 15 de outubro de 2014.

[8] SILVA A. M. "Um Estudo Sobre o Padrão H.264/AVC de Compressão de Vídeo". Universidade Católica de Pelotas. Dissertação de Mestrado. Dezembro de 2007. Disponível em <<http://ppginf.ucpel.tche.br/TI-arquivos/2007/PPGINF-UCPel-TI-2007-2-02.pdf>> - Acesso em 17 de outubro de 2014.

[9] MANOEL E. "Codificação de Vídeo H.264 - Estudo de Codificação Mista de Macroblocos". Universidade Federal de Santa Catarina. Dissertação de Mestrado. Março de 2007.

[10] FERREIRA R. "Codificador H.264/AVC com Compensação de Movimento Baseada em Partições Alternativas de Macrobloco". Dissertação (Mestrado em Engenharia Elétrica). Universidade de Brasília, 2009.

[11] GALL D. L. "MPEG: A VideoCompression Standard for MultimediaApplications". Didier Le Gall. Communications of the ACM. Volume 34, Number 4. Pages 46-58, 1991. Disponível em <<http://web.stanford.edu/class/ee398a/handouts/papers/Gall%20-%20MPEG.pdf>> - Acesso em 6 de novembro de 2014.

[12] GREENGRASS, J., EVANS, J. e BEGEN, A. C. "Not all packets are equal, part 2: The impact of network packet loss on video quality" IEEE Internet Computing, vol. 13, p. 74–82, Março 2009.

[13] VENKATARAMAN, M.; CHATTERJEE, M., "Effects of Internet Path Selection on Video-QoE: Analysis and Improvements," Networking, IEEE/ACM Transactions on , vol.22, no.3, pp.689,702, Junho de 2014.

[14] YOON J., ZHANG H., BANERJEE S., RANGARAJAN S. "Video Multicast with Joint Resource Allocation and Adaptive Modulation and Coding in 4G Networks". IEEE/ACM Transactions on Networking, Vol 22, No. 5, Outubro de 2014.

[15] KIM S.T., KOH S. J., KIM Y. J. "Performance of SCTP for IPTV Applications", The 9th International Conference on Advanced Communication Technology (ICACT 2007), pp. 2176-2180, Feb. 2007

[16] KELLY A., MUNTEAN G., PERRY P., MURPHY J. "Delay-centrichandover in SCTP over WLAN", TransactionsonAutomaticControland Computer Science 49 (63) (2004) 1–6.

[17] STEWART R. "Stream control transmission protocol SCTP PROPOSED STANDARD (RFC 4960)". September 2007. Disponível em: <<http://tools.ietf.org/html/rfc4960>>.

[18] WELTE H. "SCTP Headers". Outubro de 2003. Disponível em: <<http://security.maruhn.com/iptables-tutorial/x1736.html>> - Acesso em 20 de setembro de 2014.

[19] WALLACE T., SHAMI A. "A Review of Multihoming Issues Using the Stream Control Transmission Protocol," *Communications Surveys & Tutorials*, IEEE 14 (2) (2012) 565–578.

[20] STEWART R., RAMALHO M., XIE Q., TUEXEN M., CONRAD P. "Stream Control Transmission Protocol (SCTP) - Partial Reliability Extension". RFC 3758 (Proposed Standard), Maio de 2014.

[21] WANG H., JIN Y., WANG We., MA J., ZHANG D. "The Performance Comparison of PR-SCTP, TCP and UDP for MPEG-4 Multimedia Traffic in Mobile Network," in *Communication Technology Proceedings, 2003. ICCT 2003. International Conference on, 2003*, vol. 1, pp. 403-406 vol. 1.

[22] RÜNCOS R., RIBEIRO E. "Avaliação de parâmetros do SCTP para transporte de tráfego VoIP em cenários com perdas". Universidade Federal do Paraná.

[23] SERRAL-GRACI R., CERQUEIRA E., CURADO M., YANNUZZI M., MONTEIRO E., MASIP-BURIN X. "An overview of quality of experience measurement challenges for vídeo applications in IP networks", in: *Wired/Wireless Internet Communications*, Vol. 6074 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2010, pp. 252–263.

[24] ITU-T, Recommendation ITU-R BT.1788, Methodology for the subjective assessment of video quality in multimedia applications (January 2007).

[25] ITU-T, Recommendation ITU-T P.910, Subjective video quality assessment methods for multimedia applications (April 2007).

[26] Peak Signal-to-Noise Ratio as na Image Quality Metric. Disponível em: <<http://www.ni.com/white-paper/13306/en/>> - Acesso em 15 de setembro de 2014.

[27] WANG Y. "Survey of Objective Video Quality Measurements". EMC Corporation. Janeiro de 2006. Disponível em: <<ftp://ftp.cs.wpi.edu/pub/techreports/pdf/06-02.pdf>>

[28] WANG Z. et al. "Image quality assessment: from error visibility to structural similarity". IEEE Transactions on Image Processing, v. 13, n. 4, p. 600–612, 2004.

[29] PANDE G. "Metrics for Video Quality Assessment in Mobile Scenarios". RK Goel Institute of Technology. CoRR abs/1304.3758 (2013). i2, Electronic Edition.

[30] KLAUE J.; RATHKE, B.; WOLISZ, A. "Evalvid - a framework for video transmission and quality evaluation". In: Proceedings of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation. [s.n.], 2003. p. 255–272.

[31] TORRES A. J. F. "Método para Melhoria da Qualidade na Transmissão de Vídeos Sobre o Protocolo SCTP". Universidade Federal do Paraná. Dissertação de Mestrado. Abril de 2014.

[32] ZINNER T.; ABBOUD O.; HOHLFELD O.; HOSSFELD T.; TRAN-GIA P. "Towards QoE Management for Scalable Video Streaming". ITC-SS21, 2010.