

**ESCOLA SUPERIOR ABERTA DO BRASIL  
LATO SENSU EM ENGENHARIA DE SISTEMAS**

**RODRIGO MULINARI**

**QUALIDADE NA TRANSMISSÃO DE VÍDEO NA INTERNET**

**VILA VELHA – ES**

**2010**

**RODRIGO MULINARI**

**QUALIDADE NA TRANSMISSÃO DE VÍDEO NA INTERNET**

MONOGRAFIA APRESENTADA AO CURSO DE  
PÓS-GRADUAÇÃO EM ENGENHARIA DE SISTEMAS DA  
ESCOLA SUPERIOR ABERTA DO  
BRASIL COMO REQUISITO PARA OBTENÇÃO DO  
TÍTULO DE ESPECIALISTA EM ENGENHARIA DE  
SISTEMA, SOB ORIENTAÇÃO DA PROFESSORA  
BEATRIZ CHRISTO GOBBI.

**VILA VELHA - ES**

**2010**

**RODRIGO MULINARI**

**QUALIDADE NA TRANSMISSÃO DE VÍDEO NA INTERNET**

MONOGRAFIA APROVADA EM .... DE ..... DE 2010.

**BANCA EXAMINADORA**

---

---

---

**VILA VELHA – ES**

**2010**

## **DEDICATÓRIA**

Meus pais, irmãos, cunhados, sobrinhos e amigos pelo apoio incondicional e minha esposa, pela compreensão e suporte.

## **AGRADECIMENTOS**

À minha família, e especialmente aos meus pais e irmãos, que me incentivaram nos momentos mais difíceis e sempre me motivaram e ajudaram de todas as formas em minha educação e formação.

À minha esposa Emanuele, pela compreensão, apoio, confiança e inspiração, que foram fatores fundamentais para a conclusão desta etapa.

Aos professores, pela atenção e orientação dedicada, pelo conhecimento transmitido e por todas as reuniões, revisões e demais colaborações essenciais ao desenvolvimento desse trabalho.

Aos membros da banca, pelo interesse e pelas contribuições com correções e sugestões.

Aos amigos que me apoiaram em todas as horas.

Aos colegas do Banco do Brasil pela compreensão, apoio e torcida.

A todas as pessoas que, direta ou indiretamente, ajudaram-me a seguir esse caminho e a construir todo o conhecimento necessário para a realização desse trabalho, seja por meios de ensinamentos ou de inspiração.

## RESUMO

**Palavras-chave:** *streaming*; Adaptação de Rede; Adaptação de Conteúdo; Qualidade de Experiência; Qualidade de Serviço; SVC.

O tráfego de vídeo é cada vez mais popular na Internet. Porém, fatores como perdas de pacotes, variações dinâmicas de banda, congestionamento, dentre outros, tornam desafiador o objetivo de entregar vídeos com qualidade em níveis aceitáveis, havendo necessidade de soluções capazes de prover adaptação a essas variações nas condições da rede. Neste trabalho, estudamos o uso de protocolos de rede, em combinação com técnicas de adaptação de rede e conteúdo, a fim de alcançar a melhor qualidade de vídeo possível para o usuário. Este estudo apresenta, propõe a pesquisa dos protocolos de transmissão, QoS- Qualidade de Serviço em redes, QoE – Qualidade de Experiência do usuário em qualidade de vídeo, técnicas de adaptação de rede e de conteúdo, histórico dos padrões de codificação de vídeo e análise do uso da escalabilidade de vídeo para permitir a maior qualidade nas transmissões de vídeo na internet.

## LISTA DE FIGURAS

FIGURA 1: CLASSE DE COMUNICAÇÃO <i>UNICAST</i> .....	6
FIGURA 2: CABEÇALHO UDP (RFC 768, 1980). ....	7
FIGURA 3: CABEÇALHO RTP (RFC 3550, 2003).....	8
FIGURA 4: CABEÇALHO DO RTCP (RFC 3550, 2003).....	11
FIGURA 5: CRONOLOGIA DO DESENVOLVIMENTO DOS DIVERSOS PADRÕES DE COMPRESSÃO DE VÍDEO (BASEADO EM (ITU-T H264 - ISO/IEC14496,2005)).....	21
FIGURA 6: SEQUÊNCIA DE UNIDADES NAL (BASEADO EM (SILVA,2007).) .....	25
FIGURA 7: ARQUITETURA GENÉRICA DA NORMA H.264/AVC (BASEADA EM (SOLER, 2006)). ....	26
FIGURA 8: ESCALABILIDADE TEMPORAL EM CAMADAS (BASEADO EM (SOLER, 2006)).....	30
FIGURA 9: ESCALABILIDADE DOS QUADROS NAS CAMADAS TEMPORAIS ((MAYER,2005)).....	31
FIGURA 10: ESCALABILIDADE ESPACIAL EM CAMADAS (BASEADO EM (SOLER, 2006)).....	32
FIGURA 11: ESCALABILIDADE SNR EM CAMADAS (BASEADO EM (SOLER, 2006)).....	33
FIGURA 12: ESTRUTURA HIERÁRQUICA DA ESCALABILIDADE TEMPORAL (BASEADO EM (SCHAWARZ,2007)). ....	37
FIGURA 13: HIERARQUIA DOS QUADROS DA ESCALABILIDADE TEMPORAL EM CAMADAS (BOTTREAU, 2007).....	38
FIGURA 14: HIERARQUIA DOS QUADROS DA ESCALABILIDADE ESPACIAL EM CAMADAS (BOTTREAU, 2007).....	39
FIGURA 15: ESCALABILIDADES TEMPORAL, ESPACIAL E SNR NAS CAMADAS BASE E DE REFORÇO (MAYER,2005).....	40
FIGURA 16: IMAGEM DE ALTA RESOLUÇÃO, ALTA TAXA DE QUADROS E ALTA QUALIDADE (BASEADO EM (BOTTREAU, 2007)).....	41
FIGURA 17: IMAGEM DE ALTA RESOLUÇÃO, ALTA TAXA DE QUADROS E BAIXA QUALIDADE (BASEADO EM (BOTTREAU, 2007)).....	42

FIGURA 18: IMAGEM DE RESOLUÇÃO MÉDIA, MÉDIA TAXA E QUADROS E QUALIDADE MÉDIA (BASEADO EM (BOTTREAU, 2007)).....	42
FIGURA 19: IMAGEM DE BAIXA RESOLUÇÃO, MÉDIA TAXA DE QUADROS E ALTA QUALIDADE (BASEADO EM (BOTTREAU, 2007)). .....	43
FIGURA 20: ARQUITETURA CONCEITUAL GENÉRICA PARA ADAPTAÇÃO DE CONTEÚDO MULTIMÍDIA (ADAPTADO DE (CARVALHO, 2004)). .....	55

## LISTA DE TABELAS

TABELA 1: VALORES RECOMENDADOS PARA COMPRESSÃO MPEG-2 EM QUALIDADE <i>STANDARD DEFINITION TELEVISION</i> (SDTV) (ADAPTADO DE (DSL FORUM. 2006)).....	16
TABELA 2: VALORES RECOMENDADOS PARA COMPRESSÃO MPEG-4 EM QUALIDADE SDTV. (ADAPTADO DE (DSL FORUM. 2006)). .....	17
TABELA 3: VALORES RECOMENDADOS PARA COMPRESSÃO MPEG-2 EM QUALIDADE HDTV. (ADAPTADO DE (DSL FORUM. 2006)). .....	17
TABELA 4: ESCALABILIDADE DE VÍDEO NOS PADRÕES DE CODIFICAÇÃO (ADAPTADO DE (BOTTREAU, 2007)). .....	29

## GLOSSÁRIO DE SIGLAS

AF	<i>Assured Forwarding</i>
AIMD	<i>Additive Increase–Multiplicative Decrease</i>
ANSI	<i>American National Standards Institute</i>
ASF	<i>Application Specific Functions</i>
ASP	<i>Advanced Simple Profile</i>
AVC	<i>Advanced Video Coding</i>
BA	<i>Behaviour Agregate</i>
BE	<i>Best Effort</i>
BVQM	<i>Batch Video Quality Metric</i>
CAT	<i>Content Adaptation Techniques</i>
CBR	<i>Constant Bit Rate</i>
CGS	<i>Coarse-Grain Scalability</i>
CIF	<i>Common Intermediate Format</i>
CoS	<i>Class of Service</i>
CSRC	<i>Contributing source Count</i>
DiffServ	<i>Differentiated Services</i>
D-ITG	<i>Distributed Internet Traffic Generator</i>
DS	<i>Differentiated Service</i>
DSCP	<i>Differentiated Service Code Point</i>
DVRC	<i>Dynamic Video Rate Control</i>
ECN	<i>Explicit Congestion Notification</i>
EF	<i>Expedited Forwarding</i>
FGS	<i>Fine Grain Scalability</i>
Fps	<i>Frames per second</i>
FR	<i>Full Reference</i>
FTP	<i>File Transfer Protocol</i>

GIF	<i>Graphics Interchange Format</i>
HDTV	<i>High-Definition Television</i>
IEC	<i>International Electrotechnical Commission</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IETF	<i>Internet Engineering Task Force</i>
IntServ	<i>Integrated Services</i>
IP	<i>Internet Protocol</i>
ISSO	<i>International Organization for Standardization</i>
ITS	<i>Institute of Telecommunication Sciences</i>
ITU	<i>International Telecommunication Union</i>
ITU-T	<i>International Telecommunications Union - Telecommunications</i>
JPEG	<i>Joint Photographic Experts Group</i>
JVT	<i>Joint Video Team</i>
Kbps	<i>Kilo bits per second</i>
LDA	<i>Loss-Delay based Adjustment</i>
LRPS	<i>Loss Rate Packet per Second</i>
LR	<i>Loss Rate</i>
Mbps	<i>Mega bits per second</i>
MF	<i>Multi-Field</i>
MGS	<i>Medium Grained Scalability</i>
MPEG	<i>Moving Picture Experts Group</i>
MPQM	<i>Moving Picture Quality Metrics</i>
MSE	<i>Mean Square Error</i>
MTU	<i>Maximum Transfer Unit</i>
NAL	<i>Network Abstraction Layer</i>
NAT	<i>Network Adaptation Techniques</i>
NR	<i>No Reference</i>
NTIA	<i>National Telecommunication and Information Administration</i>
PCF	<i>Point Coordinator Function</i>

PDA	<i>Personal Digital Assistant</i>
PHB	<i>Per Hop Behaviour</i>
PSNR	<i>Peak Signal-to-Noise Ratio</i>
PT	<i>Payload Type</i>
QCIF	<i>Quarter Common Intermediate Format</i>
QoS	<i>Quality-Oriented Adaptive Schemes</i>
QoE	<i>Quality of Experience</i>
QoS	<i>Quality of Service</i>
RAP	<i>Rate Adaptation Protocol</i>
RBSP	<i>Raw Byte Sequence Payload</i>
RR	<i>Receptor-Report</i>
RR	<i>Reduced Reference</i>
RSVP	<i>Resource ReSerVation Protocol</i>
RTCP	<i>Real Time Control Protocol</i>
RTP	<i>Real-time Transport Protocol</i>
RTT	<i>Round Trip Time</i>
SDES	<i>Source description items</i>
SDTV	<i>Standard Definition Television</i>
SLA	<i>Service Level Agreement</i>
SN	<i>Sequence Number</i>
SNR	<i>Signal-to-Noise Ratio</i>
SR	<i>Sender-Report</i>
SSRC	<i>Synchronization Source</i>
SVC	<i>Scalable Video Coding</i>
SVCf	<i>Scalable Video Coding full</i>
SVCt	<i>Scalable Video Coding temporal</i>
TCA	<i>Traffic Conditioning Agreement</i>
TCP	<i>Transmission Control Protocol</i>
TOS	<i>Type of Service</i>

TS	<i>Timestamp</i>
UDP	<i>User Datagram Protocol</i>
UML	<i>Unified Modeling Language</i>
VCEG	<i>Video Coding Experts Group</i>
VCL	<i>Video Coding Layer</i>
VQEG	<i>Video Quality International Expert Group</i>
VQM	<i>Video Quality Metric</i>
VQMg	<i>Video Quality Metric General</i>
VTC	<i>Visual Texture Coding</i>

# SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO.....</b>	<b>1</b>
1.1 - MOTIVAÇÃO.....	2
1.2 - OBJETIVOS.....	2
1.3 ORGANIZAÇÃO.....	3
<b>CAPÍTULO 2 - TRANSMISSÃO DE VÍDEO E QOS EM REDES .....</b>	<b>4</b>
2.1 CLASSES DE COMUNICAÇÃO.....	5
2.2 PROTOCOLOS DE TRANSMISSÃO DE VÍDEO.....	6
2.2.1 <i>Protocolos de Transporte</i> .....	6
2.2.2 <i>Protocolos de Aplicação e Controle</i> .....	7
2.3 QUALIDADE DE SERVIÇO – QoS.....	14
2.3.1 <i>Parâmetros de QoS</i> .....	15
2.1.1.1 Jitter.....	15
2.1.1.2 Atraso.....	15
2.1.1.3 Perda de Pacotes.....	16
2.1.1.4 Vazão.....	16
2.1.1.5 Valores recomendados.....	16
2.3.2 <i>Mecanismos de QoS</i> .....	18
<b>CAPÍTULO 3 - ESCALABILIDADE DE VÍDEO .....</b>	<b>19</b>
3.1 HISTÓRICO DA CODIFICAÇÃO DE VÍDEO.....	20
3.2 CONCEITOS BÁSICOS EM CODIFICAÇÃO DE VÍDEO.....	21
3.2.1 <i>Codificação MPEG e Tipos de Quadros</i> .....	22
3.2.2 <i>Padrão H.264 AVC</i> .....	22
2.1.1.6 <i>Camada de Codificação de Vídeo ou Video Coding Layer (VCL) e Camada de Abstração de Rede ou Network Abstraction Layer (NAL)</i> .....	24
2.1.1.7 <i>Arquitetura H.264 AVC</i> .....	25
3.2.3 <i>Padrão H.264 SVC</i> .....	26
3.3 HISTÓRICO DA ESCALABILIDADE DE VÍDEO.....	28
3.4 TIPOS DE ESCALABILIDADE.....	29
3.4.1 <i>Escalabilidade Temporal</i> .....	30
3.4.2 <i>Escalabilidade Espacial</i> .....	31
3.4.3 <i>Escalabilidade SNR</i> .....	32
3.4.4 <i>Combinação dos tipos de Escalabilidade</i> .....	33
3.5 NÍVEL DE GRANULARIDADE.....	34
3.5.1 <i>Escalabilidade de granularidade baixa</i> .....	34
3.5.2 <i>Escalabilidade de granularidade elevada</i> .....	35
3.5.3 <i>Escalabilidade híbrida</i> .....	36
3.6 ESCALABILIDADE NOS PADRÕES H.264 AVC/SVC.....	36
3.6.1 <i>Escalabilidade no padrão H.264/SVC</i> .....	36
<b>CAPÍTULO 4 - QUALIDADE DE EXPERIÊNCIA – QOE.....</b>	<b>44</b>
4.1 DEFINIÇÕES DE QUALIDADE DE EXPERIÊNCIA (QoE).....	44

4.2	RELAÇÃO ENTRE QoS, QoE E QoV – QUALIDADE DE VÍDEO.....	46
4.3	QUALIDADE DE VÍDEO .....	47
4.3.1	<i>Métodos Objetivos</i> .....	47
4.3.2	<i>MSE - Mean Square Error ou Erro Médio Quadrático</i> .....	48
4.3.3	<i>PSNR - Peak Signal-to-Noise Ratio</i> .....	49
4.3.4	<i>VQM – Video Quality Metric</i> .....	50
4.4	ABORDAGENS PARA QoE .....	52
<b>CAPÍTULO 5 - ADAPTAÇÃO DE CONTEÚDO E DE REDE.....</b>		<b>54</b>
5.1	ADAPTAÇÃO DE CONTEÚDO .....	54
5.1.1	<i>Classificação das Técnicas de Adaptação de Conteúdo</i> .....	55
5.1.1.1	<i>Primeira Classificação (Zhang, 2000)</i> .....	55
5.1.1.1.1	Classificação da finalidade .....	55
5.1.1.1.2	Abstração de Informação .....	56
5.1.1.1.3	Transformação de modalidade .....	56
5.1.1.1.4	Transcodificação de dados .....	56
5.1.1.1.5	Priorização de dados .....	57
5.1.2	<i>Segunda Classificação (Liu, 2003)</i> .....	57
5.1.2.1.1	Controle de taxa .....	57
5.1.2.1.2	Transcodificação .....	57
5.1.2.1.3	Codificação escalável .....	58
5.1.2.1.4	Adaptação Estática .....	58
5.1.3	<i>Adaptação Dinâmica</i> .....	58
5.1.4	<i>Arquiteturas dos Adaptadores de Conteúdo</i> .....	59
5.2	ADAPTAÇÃO DE REDE.....	59
5.2.1	<i>Sensoriamento</i> .....	60
5.2.2	<i>Processamento</i> .....	60
5.2.3	<i>Atuação</i> .....	60
<b>CAPÍTULO 6 - CONCLUSÃO.....</b>		<b>61</b>
<b>CAPÍTULO 7 - REFERÊNCIAS BIBLIOGRÁFICAS.....</b>		<b>63</b>

## INTRODUÇÃO

A grande maioria dos computadores fixos e boa parte dos recentes dispositivos portáteis, como celulares, PDAs e *notebooks*, é capaz de receber *streaming* de vídeo pela *Internet*, contribuindo para que a transmissão de vídeo seja considerada uma das principais candidatas à próxima aplicação de uso maciço na “grande rede”. Porém, a transmissão de vídeo em redes como a *Internet*, em níveis aceitáveis de qualidade, gera uma série de desafios. Por exemplo, fluxos de vídeo comprimidos podem apresentar grandes variações nas suas taxas, algo que numa rede IP, baseada no melhor esforço, leva a situações de dificuldade de gerenciamento e de garantia de recursos necessários à entrega destes vídeos. Além disso, o caráter comumente imprevisível do comportamento relativo às características de tráfego das diversas redes cabeadas e não-cabeadas, interligadas ou não à *Internet*, principalmente em termos de largura de banda, atraso e *jitter* (atraso médio), torna a transmissão de vídeo uma tarefa ainda mais desafiadora.

Para prover a adaptação das transmissões de vídeo às condições da rede buscando assegurar uma boa qualidade de experiência aos usuários, são consideradas técnicas de adaptação de conteúdo de forma integrada a técnicas de adaptação de rede.

Em especial, o uso de escalabilidade de vídeo, como forma de prover a adaptabilidade de conteúdo, pode ser utilizada em conjunto com a adaptabilidade de rede, permitindo o tratamento de demandas que podem variar em decorrência, por exemplo, da heterogeneidade dos terminais clientes. Neste sentido, a escalabilidade de vídeo permite o envio de camadas complementares de vídeo, que são transmitidas de forma a melhor aproveitar a largura de banda disponível e que são utilizadas pelos usuários conforme sua capacidade de processamento e recepção.

## MOTIVAÇÃO

Uma das questões mais críticas na transmissão de vídeo é assegurar a qualidade do serviço (QoS) e a qualidade de experiência do usuário (QoE) em um nível aceitável em redes heterogêneas e suscetíveis a mudanças de condições de operação. A Qualidade de Experiência (QoE) é mais ampla que a Qualidade de Serviço (QoS), que incide principalmente sobre as medidas de desempenho na perspectiva de rede, enquanto a QoE aborda a satisfação do cliente com o serviço e não apenas a qualidade e o desempenho da rede. Para prover QoE em *streaming* de vídeo em níveis aceitáveis, é importante avaliar técnicas de adaptação de rede e de conteúdo, tais como esquemas adaptativos. Para esse fim, o uso de métricas objetivas permite avaliar a qualidade do vídeo entregue ao usuário e conseqüentemente da qualidade percebida por ele.

A QoE, neste trabalho, foi avaliada apenas nos aspectos de qualidade de vídeo. Para esse fim, foram utilizadas as métricas objetivas PSNR e VQM, que permitem avaliar a qualidade do vídeo entregue ao usuário e conseqüentemente da qualidade percebida por ele (Wolf, 2009) (Martinez, 2007) (Gross, 2004) (Rodrigues, 2008) (Kuipers,2010) (Vishwanath,2008) (Piamrat,2009).

## OBJETIVOS

Este trabalho tem por objetivo a pesquisa dos protocolos de transmissão, QoS em redes, QoE em qualidade de vídeo, técnicas de adaptação de rede e de conteúdo, histórico dos padrões de codificação de vídeo e análise do uso da escalabilidade de vídeo para permitir a maior qualidade nas transmissões de vídeo na internet.

O trabalho contempla como objetivos específicos, o estudo dos protocolos de transporte e protocolos de aplicação e controle para transmissão de vídeo, a investigação de parâmetros e mecanismos de QoS em redes, o estudo da QoE, diferenciado-a de QoS e métricas de qualidade de vídeo para QoE, a pesquisa de técnicas de adaptação de conteúdo, estudar técnicas de adaptação de redes e a pesquisa da codificação de vídeo e de escalabilidade de vídeo.

O emprego de diferentes formas de escalabilidade de vídeo com base no uso da codificação H.264/SVC, permitindo a codificação em camadas escaláveis temporais, espaciais e de qualidade. Em especial, o trabalho objetiva avaliar a melhoria da QoE considerando duas formas de escalabilidade: a primeira, envolvendo somente escalabilidade temporal dos quadros; a segunda, considerando vídeos codificados com escalabilidades temporal, espacial e de qualidade.

## **ORGANIZAÇÃO**

O documento está organizado da seguinte forma: no Capítulo 1, é apresentada uma visão da transmissão de vídeo, detalhando o *unicast* como modelo de endereçamento de entrega da transmissão de vídeo. São apresentadas ainda as formas de transmissão de vídeo e os protocolos UDP, RTP e RTCP, utilizados nos experimentos. Ainda neste Capítulo é tratada a Qualidade de Serviço (QoS) e destacada a sua importância, além de algumas abordagens de QoS em serviços de transmissão de vídeo.

No Capítulo 2 é apresentada uma visão geral da escalabilidade de vídeo, com detalhamento das escalabilidades temporal, espacial e de qualidade ou *Signal-to-Noise Ratio* (SNR), e ainda um histórico da codificação de vídeo e do padrão H.264 SVC.

No Capítulo 3 são apresentados aspectos da QoE e as métricas objetivas de qualidade de vídeo *Peak Signal-to-Noise Ratio* (PSNR) e *Video Quality Metric* (VQM).

O Capítulo 4 descreve e classifica as técnicas de adaptação de conteúdo e de rede.

Por fim, são descritas a conclusão do estudo.

## Capítulo 1 - TRANSMISSÃO DE VÍDEO E QoS EM REDES

As pesquisas em transmissão de vídeo são motivadas por diversos fatores, tais como a crescente demanda de serviços de vídeo na *Internet*, a exigência dos usuários por adequada qualidade nos serviços de transmissão de vídeo e a popularização e crescente oferta de novos dispositivos para acesso a vídeo na *Internet* e TV, por exemplo, PDAs, *notebooks*, iPhones, *smartphones*, dentre outros.

Existem dois modelos principais de distribuição de vídeo: o modelo de transferência completa (*downloading*) e o modelo de transmissão em tempo real (*streaming*) (Soler, 2006). No modelo de transferência completa, o cliente seleciona o vídeo que pretende receber, espera que este seja totalmente transferido para seu terminal e, finalmente, utiliza o decodificador adequado para visualizar o conteúdo transferido. Esta solução pode não ser satisfatória, pois o tempo de transferência pode ser elevado ou inaceitável, principalmente quando os arquivos de vídeos são grandes, isto é, quando o vídeo possui longa duração ou elevada taxa de *bits*.

O modelo de transmissão em tempo real, por outro lado, permite que o usuário visualize o vídeo à medida em que este é transferido para o terminal. Este modelo é muito popular na *Internet*, pois a visualização inicia-se pouco tempo depois de o usuário selecionar o conteúdo a que pretende ter acesso, não existindo, comumente, grande tempo de espera. O modelo de transmissão de vídeo em tempo real, conhecido como *streaming* de vídeo, é o foco desta dissertação.

A transmissão de vídeo em tempo real na *Internet* gera uma série de novos desafios, como garantir qualidade nos serviços para os diferentes tipos de receptores, utilizar ao máximo a capacidade de cada receptor, prover o melhor aproveitamento de banda, permitir o maior número de usuários utilizando a mesma banda e permanecendo em níveis aceitáveis de qualidade, evitar ou reagir rapidamente a congestionamentos na rede e atender a variações repentinas do número de usuários, dentre outros.

Para transmissão de vídeo no ambiente de rede é utilizada de forma conjunta uma série de tecnologias, protocolos, codificadores, etc. Este capítulo se propõe a detalhar algumas tecnologias utilizadas nos experimentos propostos neste documento, descrevendo o *unicast* como forma de endereçamento de entrega da transmissão de vídeo, os protocolos *User Datagram Protocol* (UDP), *Real-time Transport Protocol* (RTP) e o protocolo de controle *Real Time Control Protocol* (RTCP). Além disso, neste capítulo serão discutidos aspectos de QoS, tais como suas métricas, mecanismos e sua utilização em redes.

## 1.1 CLASSES DE COMUNICAÇÃO

Existem diversas tipos de classes de comunicação, são elas (Soares, 1995):

- a) **Broadcast.** é o processo pelo qual se transmite ou difunde determinada informação. A principal característica é que a mesma informação é enviada para muitos receptores ao mesmo tempo, numa associação de um-para-muitos entre endereços do servidor e os possíveis receptores;
- b) **Multicast.** permite a entrega de informação simultaneamente para múltiplos destinatários cadastrados em grupos. Há uma associação de um-para-muitos entre endereços do servidor e um grupo de receptores onde toda a informação é replicada;
- c) **Anycast.** é uma forma de encaminhamento onde os dados são distribuídos “ao destino mais próximo” ou “ao melhor destino” definido pelo roteamento da rede. Há também uma associação de um-para-muitos entre endereços do servidor e um grupo de receptores, mas somente um deles é escolhido para receber a informação dos dados enviados pelo servidor;
- d) **Unicast.** é uma classe de comunicação onde as informações são enviadas a partir de um ponto a outro ponto. Há uma associação um-para-um entre o endereço de origem e o endereço de destino, isto é, a troca de mensagens é feita de um remetente para apenas um receptor.

Em um fluxo *unicast* cada cliente recebe um fluxo distinto. Esta classe de endereçamento de comunicação de rede não otimiza o uso da banda disponível, podendo transferir o mesmo conteúdo para dois clientes em fluxos distintos.

A Figura 1 apresenta o funcionamento do *unicast*, adotado nos experimentos deste trabalho.

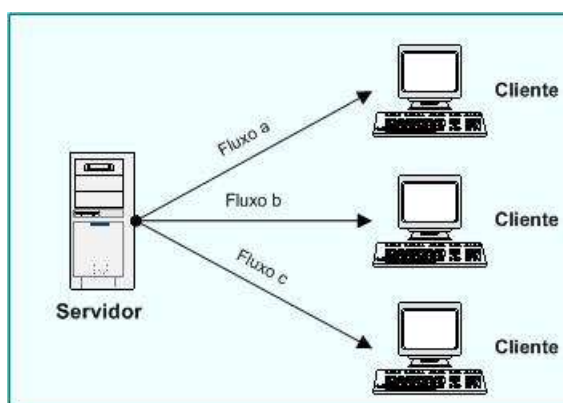


Figura 1: Classe de comunicação *Unicast*.  
Fonte: Elaboração própria(2010).

## 1.2 PROTOCOLOS DE TRANSMISSÃO DE VÍDEO

Neste capítulo são detalhados os protocolos utilizados nos experimentos feitos neste trabalho, sendo eles os protocolos UDP, RTP e o protocolo de controle RTCP.

### 1.2.1 Protocolos de Transporte

Os dois protocolos de transporte mais utilizados são o UDP e o TCP, sendo o UDP o protocolo utilizado neste trabalho. Segue abaixo uma descrição do protocolo UDP.

O protocolo *User Datagram Protocol* (UDP) foi descrito em 1980 na RFC 768 (RFC 768, 1980) e caracteriza-se por não estabelecer uma conexão prévia para o envio/recepção dos dados. Os dados são enviados em blocos chamados datagramas, e cada um destes blocos contém o endereço e a porta de destino dos dados. Além disso, não há garantia de entrega, isto é, não se prevê confirmação do recebimento dos datagramas.

O uso do protocolo de transporte UDP é recomendado para transmissão de vídeos porque não exige confirmação do recebimento do pacote, reduzindo assim o tempo de resposta e aumentando a velocidade de processamento. O protocolo UDP possui um *overhead* de 8 bytes.

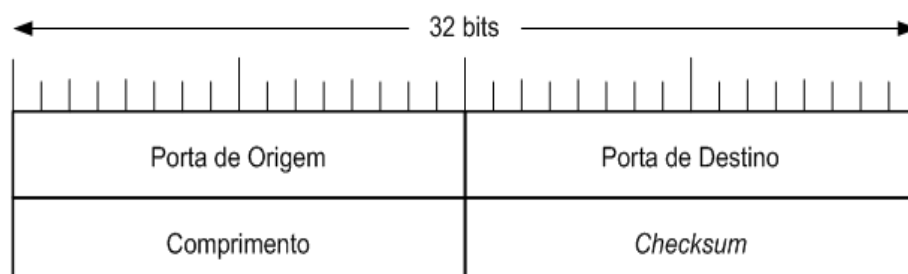


Figura 2: Cabeçalho UDP.

Fonte: RFC 768(1980).

Os campos do cabeçalho UDP apresentado na Figura 2 estão descritos a seguir:

**Porta de Origem:** a porta de origem é um campo opcional. Quando significativo, indica a porta de origem do envio do processo, e pode ser considerada como sendo a porta para que uma resposta seja abordada na ausência de quaisquer outras informações. Se não utilizado, terá o valor zero.

**Porta de Destino:** a porta de destino que irá receber os dados da transmissão.

**Comprimento:** é o comprimento total do datagrama, incluindo o cabeçalho e os dados.

**Checksum:** é calculado usando o *header* UDP e também a área de dados, e destina-se à verificação de erros de transmissão.

### 1.2.2 Protocolos de Aplicação e Controle

Os protocolos RTP e RTCP, utilizados para transmissão de vídeo, estão descritos a seguir:

O protocolo *Real-Time Transport Protocol* (RTP) foi definido pela RFC 3550 (RFC 3550, 2003) do IETF - *Internet Engineering Task Force*, para utilização em aplicativos de transmissão de dados em tempo real, como vídeo e áudio, tanto

em redes *unicast* como nas redes *multicast*, sem, entretanto, garantir a qualidade de serviço QoS ou reservar recursos de endereçamento.

O protocolo RTP (RFC 3550, 2003) é implementado sobre os protocolos UDP/IP, utilizando os serviços de multiplexação e *checksum* do UDP e estabelecendo uma comunicação fim-a-fim. Os pacotes de áudio e de vídeo, produzidos pelo aplicativo remetente, são encapsulados em pacotes RTP. Estes, por sua vez, são encapsulados em um segmento UDP. O protocolo RTP apresenta um *overhead* de 12 *bytes*. Porém, nessa arquitetura, o RTP é transmitido com utilização do protocolo UDP, que tem um *overhead* de 8 *bytes*.

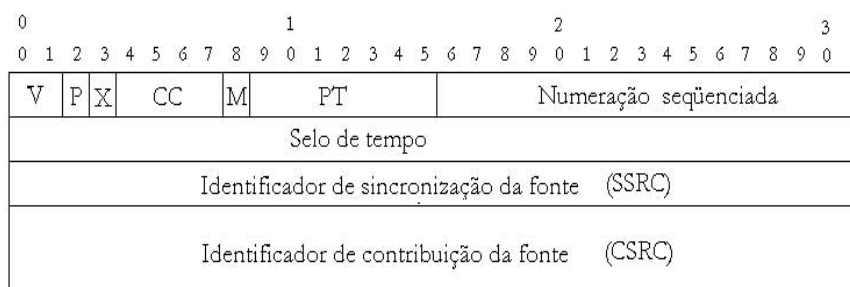


Figura 3: Cabeçalho RTP.

Fonte: RFC 3550(2003).

Os campos do cabeçalho RTP apresentados na Figura 3 estão descritos a seguir:

**Version (V):** 2 *bits*, este campo identifica a versão do RTP.

**Padding (P):** 1 *bit*. Se o *bit* de *padding* está ligado, o pacote contém um ou mais novos octetos adicionais no final que não fazem parte da carga útil. O último octeto do *padding* é uma contagem de quantos octetos adicionais devem ser ignorados.

**Extension (X):** 1 *bit*, Se o *bit* de extensão está ligado, o cabeçalho fixo é seguido por um cabeçalho de extensão.

**CSRC Count (CC):** 4 *bits*, o CSRC - *Contributing SouRce Count* contém o número de identificadores CSRC que seguem o cabeçalho fixo.

**Marker (M):** 1 *bit*, a interpretação do marcador é definido por um perfil adicional. Destina-se a permitir que eventos significativos, como os quadros (*frames*) limites, sejam marcados no pacote de dados.

**Payload Type (PT):** 7 bits, este campo identifica o formato do pacote RTP, bem como a carga e o regime de codificação/compressão. A partir deste tipo de identificador da carga, a aplicação interpreta e reproduz a carga de dados. Tipos de carga são definidos pela RFC 1890, tais como MPEG. Em um determinado momento da transmissão, um remetente RTP só pode enviar um tipo de carga, embora este possa mudar durante a transmissão, por exemplo, para ajustar-se ao congestionamento da rede.

**Numeração Sequenciada - Sequence Number (SN):** 16 bits. Cada pacote enviado é incrementado por um número sequencial e pode ser usado pelo receptor para detectar a perda ou para restaurar a ordem dos pacotes. O valor inicial do número sequencial é fixado aleatoriamente para dificultar ataques e tornar a criptografia mais segura.

**Selo de tempo - Timestamp (TS):** 32 bits. O *timestamp* reflete o instante de amostragem do primeiro octeto no pacote de dados RTP. As amostragens instantâneas devem ser provenientes de um relógio que é incrementado monotonicamente e linearmente no tempo, para permitir os cálculos de sincronização e *jitters*. A frequência do relógio depende do formato dos dados transportados como carga. O valor inicial é fixado aleatoriamente. O RTP por si só não é responsável pela sincronização, isto deve ser feito no nível da aplicação.

**Identificador de Sincronização da fonte - SSRC:** 32 bits. O campo SSRC - *Synchronization Source*, identifica a fonte de sincronização (fluxo). Este identificador é escolhido aleatoriamente, com a intenção de não sincronizar duas fontes dentro de uma mesma sessão RTP a fim de evitar que elas tenham o mesmo identificador SSRC.

**Identificador de Contribuição da fonte - CSRC:** 0 - 15 itens, 32 bits cada. O número de identificadores é dado pelo campo CC.

O protocolo *Real Time Control Protocol* (RTCP) foi definido pela RFC 3550 (RFC 3550, 2003) para prover funcionalidades de controle para o RTP. Em uma sessão

RTP, os participantes enviam periodicamente pacotes RTCP para transmitir *feedbacks* sobre a qualidade da entrega dos dados e informações dos membros.

Cinco diferentes pacotes foram definidos para o RTCP prover controle:

**SR** – *Sender-Report* ou relatório do remetente: é usado para a transmissão e recepção das estatísticas dos remetentes ativos na sessão;

**RR** – *Receiver-Report* ou relatório do receptor: é usado para transmissão de estatísticas da recepção dos pacotes pelo receptor;

**SDES**- *Source description items* ou descrição dos itens da origem: descreve os itens da origem do pacote;

**BYE** - Indica o fim da participação do receptor ou remetente na sessão;

**ASP**- *Application specific functions* ou funções específicas da aplicação.

A função primordial do RTCP é proporcionar qualidade de serviço e acompanhamento do controle de congestionamento. Estas informações são usadas tanto pelo remetente, quanto pelos receptores. O remetente pode ajustar a sua transmissão baseada nos *feedbacks* enviados nos pacotes RR. Os receptores podem determinar se um congestionamento é local, regional ou global, com base nos *feedbacks* providos pelos pacotes SR e RR.

O pacote RR será utilizado nos esquemas adaptativos experimentados neste trabalho, e é apresentado na Figura 4. O formato do pacote *Receiver-Report* (RR) é o mesmo que o dos pacotes SR, exceto pelo campo *Type Packet*, que contém a constante 201 e omite alguns campos com informações do remetente.

V	P	RC	PT=RR=201	Comprimento
Transmissor (SSRC)				
SSRC_1				
Fração de perda		Número cumulativo de pacotes perdidos		
Número sequencial recebido mais significativo				
Jitter do intervalo				
Último SR (LSR)				
Atraso desde o último SR (DLSR)				
SSRC_n				
Dados RR adicionais				

Figura 4: Cabeçalho do RTCP.

Fonte: RFC 3550(2003).

Os campos do cabeçalho RTCP apresentados na Figura 4 estão descritos a seguir:

**Versão - Version (V):** 2 bits. Identifica a versão do RTP, que é a mesma tanto em pacotes RTCP como em pacotes de dados RTP;

**Padding (P):** 1 bit. Se o bit *Padding* estiver ligado, este RTCP contém alguns pacotes adicionais na parte final, que não fazem parte da informação de controle;

**Reception count:** 5 bits. Contém o número de blocos do RR deste pacote;

**Tipo do Pacote – PT - Packet Type:** 8 bits. Contém a constante 201 para identificar um pacote RR;

**Comprimento - Length:** 16 bits. Este campo contém o tamanho do pacote incluindo o cabeçalho e os *padding*s adicionais;

**SSRC:** 32 bits. O identificador da fonte de sincronização indica o remetente do pacote RR;

**SSRC\_n (src ident):** 32 bits. O SSRC é o identificador da origem das informações contidas neste pacote RR;

**Fração de Perda - Fraction Lost:** 8 bits. A fração de pacotes de dados RTP da origem SSRC\_n perdidos desde que o pacote anterior de RR foi enviado. Esta fração é definida para ser o número de pacotes perdidos dividido pelo número de pacotes esperado, tal como definido no parágrafo seguinte;

**Número cumulativo de pacotes perdidos - *Cummulative number of packets lost*:** 24 *bits*. É o número total de pacotes de dados RTP da fonte SSRC\_n que foram perdidos desde o início da recepção. Este número é definido como sendo a diferença entre o número de pacotes esperados e o número de pacotes recebidos efetivamente, em que o número de pacotes recebidos inclui todos os que estão atrasados ou duplicados. Assim, os pacotes que chegarem atrasados não são contados como perdidos, e as perdas podem ser negativas, se existirem pacotes duplicados. O número de pacotes esperados é definido como o sequencial do último pacote recebido menos o número de sequência do pacote inicial recebido;

**Número Sequencial Recebido Mais Significativo - *Extended highest sequence number received*:** 32 *bits*. Os 16 *bits* menos significativos contêm o maior número de sequencia recebido dos pacotes de dados RTP da origem SSRC\_n, e os 16 *bits* mais significativos estendem o número da sequencia com o número sequencial correspondente na contagem de ciclos. Note que diferentes receptores, na mesma sessão, gerarão diferentes extensões para o número de ordem se o seu início diferir;

***Jitter* do Intervalo entre chegadas - *Interarrival jitter*:** 32 *bits*. É uma estimativa do *jitter* dos pacotes de dados RTP, medido em unidades *timestamp* e expressa como um *unsigned integer*;

***Timestamp* do último pacote SR - *Last SR timestamp (LSR)*:** 32 *bits*. É o *timestamp* do NTP (campo do pacote SR) recebido como parte do mais recente pacote *Sender-Report* (SR) da fonte SSRC\_n. Caso não tenha sido recebido nenhum SR, o campo será definido como zero;

**Atraso desde o último pacote SR - *Delay since last SR (DLSR)*:** 32 *bits*. É o atraso, expresso em unidades de 1 / 65536 segundos, entre a recepção do último pacote SR-*Sender-Report* da fonte SSRC\_n e o envio deste bloco do relatório RR. Se nenhum pacote SR foi recebido ainda a partir de SSRC\_n, o campo DLSR é fixado em zero;

O *jitter*  $J$  é definido como a média de desvio (valor absoluto arredondado) da diferença “D” dos espaçamentos dos pacotes recebidos no receptor em comparação com um par de pacotes enviados pelo remetente. Conforme apresentado na Equação 1 abaixo, isto é equivalente à diferença na "relação tempo de trânsito" para os dois pacotes; a relação tempo de trânsito é a diferença entre um *timestamp* do pacote RTP e o relógio do receptor no momento da chegada do pacote, medida nas mesmas unidades.

Se  $S_i$  é o *timestamp* do pacote RTP  $i$ , e  $R_i$  é a hora de chegada em unidades de *timestamp* do pacote RTP  $i$ , então para dois pacotes “ $i$ ” e “ $j$ ”, “D” pode ser expresso por (RFC 3550, 2003):

$$D(i, j) = (R_j - R_i) - (S_j - S_i) = (R_j - S_j) - (R_i - S_i) \quad (1)$$

O *jitter* deve ser calculado continuamente para cada pacote  $i$  recebido da fonte  $SSRC_n$ , utilizando esta diferença “D” para o pacote anterior  $i-1$ , por ordem de chegada, não necessariamente em seqüência, de acordo com a Equação 2 (RFC 3550, 2003):

$$J(i) = J(i-1) + (|D(i-1, i)| - J(i-1)) / 16 \quad (2)$$

Sempre que for emitido um pacote RR, o atual valor de  $J$  é amostrado.

### 1.3 QUALIDADE DE SERVIÇO – QoS

Atualmente, o conceito de Qualidade de Serviço (QoS) é amplamente utilizado, especialmente entre as redes baseadas em *Internet Protocol* (IP).

O ITU-T Rec. E.800 (ITU-T E800,1994) define QoS como "o efeito coletivo do desempenho do serviço para determinar o grau de satisfação de um usuário do serviço". Outra definição da QoS é a da norma G.1010 multimídia fim-a-fim da *International Telecommunication Union* (ITU) (ITU-T G1010,2001): "A QoS representa o conjunto de parâmetros, tecnologias e mecanismos que devem ser implementados na infra-estrutura de rede, a fim de satisfazer os requisitos de desempenho, que incide principalmente sobre parâmetros de rede em vez dos requisitos do usuário".

Serviços realizados por aplicações com requisitos semelhantes são agrupados em diferentes Classes de Serviços (COS). A QoS é então executada por meio do tratamento diferenciado de cada tráfego associado às COSs na rede, a fim de satisfazer o desejado nível de serviço de aplicação (exemplo: voz, vídeo e dados). O serviço de desempenho é, então, gerido por intermédio de um conjunto de parâmetros de rede. Os parâmetros mais comuns são: atraso (*delay*), variação do atraso (*jitter*) e perda de pacotes (ITU-T E800,1994).

Tais métricas não representam a qualidade da experiência pela perspectiva do usuário. Em vez disso, elas representam os parâmetros de rede que podem afetar o desempenho e prejudicar a experiência do usuário. Em especial, note-se que para os serviços de transmissão de vídeo na *Internet*, o usuário não tem conhecimento das deficiências de rede, tais como perdas de pacotes ou *jitter* (Zapater,2007).

O usuário está preocupado com a qualidade de vídeo, com o atraso da troca de canais, com a demora na sincronização de áudio e vídeo, entre outros (Heavyreading,2009). Essas características, ainda que dependentes do adequado funcionamento da infra-estrutura de rede e da qualidade de serviço associada,

estão fortemente ligadas à experiência vivenciada pelo usuário, remetendo ao conceito de QoE, que será melhor discutido no Capítulo 4.

### 1.3.1 Parâmetros de QoS

Podem-se destacar vários parâmetros que afetam a QoS na transmissão de vídeo em tempo real na *Internet*, dentre eles: *jitter*, atraso, perda de pacotes e vazão, apresentados a seguir.

#### 2.1.1.1 Jitter

Em uma rede baseada em protocolo IP, o *jitter* pode ser entendido como a variação no tempo e na sequência de entrega das informações (pacotes) devido à variação nos atrasos da rede. O *jitter* é importante para as aplicações executadas em rede, cuja operação adequada depende, de alguma forma, da garantia de que as informações, os pacotes de vídeo, sejam processadas em períodos de tempo bem definidos (Kurose, 2008).

#### 2.1.1.2 Atraso

É o tempo consumido entre o envio (origem) de um pacote e a recepção no seu destino. Numa transmissão de dados, os pacotes são enviados de um *host* de origem, passam por uma série de equipamentos de rede (roteadores, comutadores, etc...), e chegam no *host* de destino. Cada pacote segue de um nó (*host* ou roteador) para o nó subsequente (*host* ou roteador) ao longo do caminho, e sofre de diversos tipos de atrasos em cada nó até chegar ao *host* destino. A soma destes atrasos forma o atraso total do pacote. Os atrasos mais importantes são: o **atraso de processamento**, que é o tempo necessário para examinar o cabeçalho e determinar para onde direcionar o pacote; o **atraso de enfileiramento**, determinado pelo tempo que o pacote aguarda na fila até ser transmitido para o enlace; o **atraso na transmissão**, que é a quantidade de tempo necessário para o roteador “empurrar” todos os *bits* do pacote para fora, sendo uma função do comprimento do pacote e da taxa de transmissão do enlace, e não sendo dependente da distância entre dois roteadores; o **atraso de propagação**, definido pela quantidade de tempo que leva para o pacote viajar do

emissor para o receptor pelo meio de transmissão (fibras ópticas, satélite, cabo coaxial, dentre outros). Este último pode ser calculado pela distância entre dois roteadores dividida pela velocidade de propagação do meio de transmissão específico (Kurose, 2008).

#### 2.1.1.3 Perda de Pacotes

A perda de pacotes representa um indicador de qualidade na transmissão de pacotes IP entre dois pontos da rede, e que tem influência na qualidade de serviço, podendo ocorrer pelo estouro de *buffers* em roteadores e *switchs*, em decorrência de situações de congestionamento. Quando se tem a utilização dos protocolos UDP e RTP, esses pacotes não podem ser retransmitidos e a própria retransmissão não é tolerável em aplicações de vídeo em tempo real (DSL Forum. 2006).

#### 2.1.1.4 Vazão

É a taxa efetiva de transmissão de pacotes por segundo entre o transmissor e o receptor. A vazão, ou banda, é o parâmetro mais básico de QoS, sendo necessário dispor suficientemente deste recurso para a transmissão adequada do tráfego associado a qualquer aplicação(Kurose, 2008).

#### 2.1.1.5 Valores recomendados

Nas Tabelas 1 a 3 abaixo são apresentados os valores recomendados para latência, *jitter* e perda de pacotes, segundo a recomendação WT-126 (DSL Forum. 2006).

Tabela 1: Valores recomendados para compressão MPEG-2 em qualidade *Standard Definition Television* (SDTV).

Fluxo de Vídeo (Mbit/s)	Latência	<i>Jitter</i>	Perda de pacotes IP	Período de Coleta 30 minutos	Taxa média de perda de pacotes IP
3.0	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 30 minutos	1.9E-06
3.75	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 30	1.6E-06

				minutos	
5.0	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 30 minutos	1.2E-06

Fonte: DSL Forum (2006).

A Tabela 2 abaixo apresenta os valores recomendados para compressão MPEG-4 em qualidade *Standard Definition Television* (SDTV) (DSL Forum, 2006):

Tabela 2: Valores recomendados para compressão MPEG-4 em qualidade SDTV.

Fluxo de Vídeo (Mbit/s)	Latência	Jitter	Perda de pacotes IP	Período de Coleta 30 minutos
1.75	<200ms	<50ms	1 pacote IP	1 evento de erro a cada 30 minutos
2.0	<200ms	<50ms	1 pacote IP	1 evento de erro a cada 30 minutos
2.5	<200ms	<50ms	1 pacote IP	1 evento de erro a cada 30 minutos
3.0	<200ms	<50ms	1 pacote IP	1 evento de erro a cada 30 minutos

Fonte: DSL Forum (2006).

A Tabela 3 apresenta os valores recomendados para compressão MPEG-2 em qualidade *High-Definition Television* (HDTV) (DSL Forum, 2006):

Tabela 3: Valores recomendados para compressão MPEG-2 em qualidade HDTV.

Fluxo de Vídeo (Mbit/s)	Latência	Jitter	Perda de pacotes IP	Período de Coleta 30 minutos	Taxa média de perda de pacotes IP
8	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 4 horas	9.14E-08
10	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 4 horas	7.31E-08
2	< 200ms	< 50ms	1 Pacote IP	1 evento de erro a cada 4 horas	6.09E-08

Fonte: DSL Forum (2006).

### 1.3.2 Mecanismos de QoS

Os mecanismos básicos disponíveis para prover QoS em redes IP são a Reserva de Recursos e a Priorização (Braden, 1994).

Na Reserva de Recursos, componentes e recursos de rede são sujeitos à política de administração da largura de banda e são alocados de acordo com os requisitos de Qualidade de Serviço das aplicações. O protocolo de reserva de recursos ou *Resource ReSerVation Protocol* (RSVP), por exemplo, fornece os mecanismos para implementação de serviços integrados ou *Integrated Services* (IntServ) baseado na reserva de recursos (Braden, 1994).

Na Priorização, o tráfego da rede é classificado e os recursos de rede são divididos entre as classes de acordo com políticas de administração de largura de banda. Para habilitar QoS, os mecanismos de classificação dão tratamento preferencial a aplicações identificadas como tendo requisitos mais exigentes. A arquitetura de Serviços Diferenciados ou *Differentiated Services* (DiffServ) (Black, 1998) provê este tipo de serviço. A arquitetura de DiffServ (Black, 1998) provê este tipo de serviço.

## Capítulo 2 - ESCALABILIDADE DE VÍDEO

Numa transmissão de vídeo clássica, o codificador comprime o sinal de vídeo com uma determinada taxa de *bits* (menor ou igual à capacidade do canal), e o decodificador reconstrói o sinal utilizando todos os *bits* recebidos. Este modelo pressupõe que o codificador conheça a capacidade do canal disponível e o decodificador receba e decodifique todos os *bits* enviados. Este modelo é de difícil implementação, pois geralmente estas premissas não são verdadeiras (p.ex., na distribuição de vídeo pré-codificado a taxa de *bits* disponível no canal não é conhecida durante a codificação (Soler, 2006)).

Uma alternativa na distribuição de vídeos pré-codificados é codificar fluxos independentes, isto é, a mesma informação é codificada de maneira independente, visando diferentes codificações em termos de taxa de *bits*, resolução espacial e temporal. Sendo assim, um servidor de vídeo distribui os *bitstreams* codificados em um ou mais fluxos simultaneamente.

Neste modelo, é necessário codificar o conteúdo múltiplas vezes com os algoritmos de codificação já existentes, levando a uma baixa eficiência, uma vez que não se explora a dependência entre os diversos *bitstreams* independentes. Em face desta dependência, é possível ter uma repetição do conteúdo codificado no servidor e, quando necessário, também durante a transmissão. Outra desvantagem deste modelo consiste no número limitado de representações do vídeo, uma vez que sempre que se deseje obter uma representação correspondente a condições que não tenham sido previamente consideradas, e que o respectivo conteúdo não esteja disponível no servidor, com uma determinada taxa de *bits*, é necessário recodificar esse conteúdo (Schwarz,2007).

Para suprir estas desvantagens, pode ser utilizada a escalabilidade na codificação para obter uma representação escalável do vídeo. A escalabilidade é a capacidade de extrair do *bitstream* total correspondente a um dado conteúdo, codificado em certas condições - por exemplo, com uma determinada qualidade e

resolução - denominado neste caso por *bitstream* escalável, subconjuntos eficientes de *bits* que podem ser utilmente decodificáveis, ou seja, que oferecem imagens ou vídeos decodificados com uma variação (camada) segundo uma dada característica da imagem ou vídeo, como exemplo, a resolução espacial ou qualidade (Soler, 2006).

A escalabilidade é alcançada através de uma camada base e camadas incrementais ou de reforço que exploram a redundância em relação à camada base, e eventualmente, em relação a outras camadas de reforço, que permitem oferecer sucessivas imagens ou vídeo decodificados segundo as dimensões de escalabilidade temporal ou espacial adotadas.

Neste capítulo a escalabilidade de vídeo é apresentada com um breve histórico, os tipos de escalabilidade e sua aplicação na codificação de vídeos com uma sucinta descrição dos codificadores de vídeo H.264 AVC (*Advanced Video Coding*) (ITU-T H264,2007) e H.264 SVC – *Scalable Video Coding* (JVT - a,2005) (JVT - B,2005) (JVT - C,2005).

## 2.1 HISTÓRICO DA CODIFICAÇÃO DE VÍDEO

Os primeiros esforços para a padronização da codificação de vídeo digital datam do final da década de 80 e início da década de 90, com a criação do padrão H.261 pelo ITU-T, evoluído posteriormente para os padrões H.262, H.263, H.263+, H.263++, H.264 (MPEG-4 AVC - *Advanced Video Coding*), H.264 FRExt e, finalmente, para o padrão H.264 SVC.

Os padrões MPEG surgiram em 1988, com o padrão MPEG-1, evoluindo para os padrões MPEG-2, MPEG-4 ASP - *Advanced Simple Profile* e MPEG-4 AVC (H.264 AVC).

A Figura 5 apresenta a cronologia dos padrões de codificação desenvolvidos por tais órgãos.

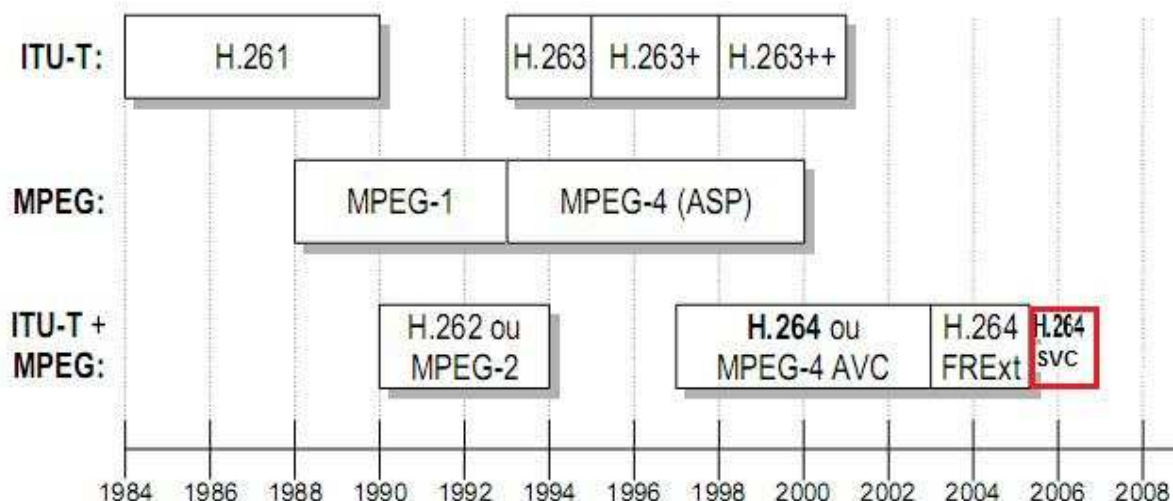


Figura 5: Cronologia do desenvolvimento dos diversos padrões de compressão de vídeo.

Fonte: ITU-T H264 -ISO/IEC14496(2005).

O padrão H.264/AVC tem atraído muita atenção por parte da indústria, sendo cada vez mais utilizado em uma ampla variedade de aplicações. Espera-se que, em curto prazo, o H.264/AVC seja comumente utilizado na maioria das aplicações de vídeo. Dado este elevado grau de aprovação e implantação da nova norma, e levando em conta o melhor desempenho em termos de eficiência de compressão quando comparado com o padrão MPEG-2 e os investimentos que já foram realizados para a preparação e desenvolvimento de produtos baseados no H.264/AVC, é natural o investimento em uma extensão *Scalable Video Coding* (SVC) com a reutilização das principais características do padrão H.264 AVC (Schwarz,2007), visando prover escalabilidade espacial, temporal e de qualidade, como será explanado a seguir.

## 2.2 CONCEITOS BÁSICOS EM CODIFICAÇÃO DE VÍDEO

A padronização de técnicas de codificação de vídeo digital é essencial para permitir que equipamentos de vídeo digital de diferentes fabricantes se comuniquem. Atualmente, os principais órgãos responsáveis por tal padronização são o ITU-T *Video Coding Experts Group* (VCEG) e o ISO/IEC *Moving Picture Experts Group* (MPEG).

O Setor de Padronização de Telecomunicações ou *Telecommunication Standardization Sector* é um setor da *International Telecommunication Union* (ITU), localizado em Genebra, Suíça. O MPEG é um grupo de trabalho conjunto das instituições *International Organization for Standardization* (ISO) e *International Electrotechnical Commission* (IEC).

### 2.2.1 Codificação MPEG e Tipos de Quadros

Existem diferentes tipos de quadros em fluxos de vídeo MPEG (Reichel, 2006):

**Quadros do tipo I ou I-frames - quadro Intra-code:** é um quadro completo que, portanto não precisa dos demais quadros para sua decodificação;

**Quadros do tipo P ou P-frames – quadro Preditivo:** é uma codificação com predição e utiliza informações do quadro de referência (preditivo), geralmente outro quadro I, para decodificação;

**Quadros do tipo B ou B-frames – quadro Bidirecionalmente preditivo:** necessita do quadro (I ou P) posterior e anterior para a correta decodificação, mas não se baseia em outros quadros B.

### 2.2.2 Padrão H.264 AVC

O H.264 AVC é um padrão para a compressão de vídeo, e é equivalente ao MPEG-4 Parte 10, ou MPEG-4 AVC. O padrão ITU-T H.264 e as normas ISO / IEC MPEG-4 Parte 10 (formalmente, a ISO / IEC 14496-10) são mantidos em conjunto para que tenham conteúdo técnico idêntico. O padrão H.264 é um esforço conjunto entre as duas instituições.

A intenção do projeto, alcançada no padrão H.264/AVC, era criar um padrão capaz de fornecer uma boa qualidade de vídeo em taxas de *bits* mais baixas do que as das normas anteriores: com a utilização de metade ou menos que a taxa de *bits* utilizada pelos padrões MPEG-2, H.263, ou MPEG-4 Parte 2, sem aumentar a complexidade da arquitetura para não torná-lo impraticável ou excessivamente caro de implementar. Um objetivo adicional era fornecer

flexibilidade suficiente para permitir que a norma fosse aplicada a uma ampla gama de aplicações em diversas redes, incluindo as redes com baixas e altas taxas de *bits*, baixa e alta resolução de vídeo, transmissão e armazenamento, dentre outras (Schwarz,2007).

Atualmente, o H.264 possui um bom desempenho, com uma taxa de *bits* menor em 50% se comparado ao seu antecessor MPEG-2, e mantendo o mesmo desempenho em termos de erro (PSNR, por exemplo). Além disso, o H.264 possui uma série de avanços em relação a seus antecessores, principalmente em comparação ao MPEG-1, MPEG-2 e H.263, com respeito às etapas de predição, transformação e codificação. Muitos desses avanços apareceram primeiro nos anexos do H.26, resultando no H.263+, H.263++ e no padrão MPEG-4 ASP (Manoel,2007).

As novas características do padrão H.264 trazem, cada uma, pequenas melhorias que, conjuntamente, provêm um grande impacto na taxa de compressão do H.264 em comparação com os seus antecessores. E essas novas características, a serem descritas em seguida, agregam uma complexidade à codificação como também à decodificação, tanto em termos computacionais, sendo necessários *hardwares* muito mais eficientes, principalmente para o codificador, quanto em termos de desenvolvimento, pois existem muito mais processos envolvidos e muito mais interdependências entre tais processos (Manoel,2007).

As principais características definidas para esta norma são (Chiariglione, 2001):

- a) **Desempenho elevado:** Redução em cerca de 50% da taxa de *bits* e mantendo o mesmo desempenho em termos de erro (PSNR, por exemplo) em relação às normas antecessoras;
- b) **Máxima simplificação:** Adoção de uma arquitetura simples, utilizando blocos de complexidade reduzida;
- c) **Adaptação a serviços com diferentes requisitos de atraso:** Deve permitir serviços em tempo real ou com atraso reduzido, tais como a videotelefonia e a transmissão de vídeo na Internet, bem como serviços

sem quaisquer restrições de atraso, tais como o armazenamento e distribuição de vídeo;

- d) **Resiliência a erros:** Deve incluir ferramentas com o intuito de minimizar o impacto resultante da perda de pacotes e erros de *bit* em redes móveis ou fixas;
  
- e) **Adaptação às características da rede de transmissão (*network friendliness*):** Deve haver um conjunto de mecanismos que facilitem o transporte do *bitstream* codificado em redes com características diferentes.

Outra característica desta norma é a compatibilidade com os padrões de codificação já existentes, nomeadamente com a parte de sistema do MPEG-4 (ex: MP4), com a parte de sistema da norma MPEG-2 e com as recomendações H.320 e H.323 (Soler, 2006).

#### 2.1.1.6 Camada de Codificação de Vídeo ou *Video Coding Layer* (VCL) e Camada de Abstração de Rede ou *Network Abstraction Layer* (NAL)

A concepção do padrão de codificação de vídeo H.264/AVC inclui os conceitos de Camada de Codificação de Vídeo (VCL) e de Camada de Abstração de Rede (NAL). A VCL cria uma representação codificada do conteúdo original, enquanto que a NAL apresenta o formato destes dados e informações no cabeçalho, de maneira simples e eficaz e que permite a personalização da utilização de dados VCL para uma ampla variedade de sistemas (Schwarz,2007).

O H.264/AVC faz uma distinção entre a VCL e a NAL. A saída do processo de codificação são os pacotes VCL, isto é, uma seqüência de *bits* que representam os dados do vídeo codificado e que são mapeados para a unidade NAL para uma posterior transmissão ou armazenamento. Cada unidade NAL contém um *Raw Byte Sequence Payload* (RBSP), que é um conjunto de dados correspondentes aos dados do vídeo codificado que permite a transmissão destes dados em pacotes.

Uma seqüência de vídeo codificado é representada por uma seqüência de unidades NAL, como apresenta a Figura 6, e pode ser transmitida através de uma

rede de comutação de pacotes, com transmissão em *bitstream*, ou armazenada em um arquivo. O propósito de especificação separada entre a VCL e a NAL é fazer uma clara distinção entre as características de codificação (VCL) e as características de transporte (NAL) (Silva,2007).



Figura 6: Sequência de unidades NAL.

Fonte: Silva(2007).

#### 2.1.1.7 Arquitetura H.264 AVC

Como já citado anteriormente, a norma H.264/AVC especifica duas camadas principais de representação: uma camada de codificação de vídeo VCL, que permite representar eficientemente o conteúdo de uma seqüência de imagens, e uma camada de adaptação de rede NAL, que encapsula essa representação para a forma adequada ao transporte em qualquer rede de comutação de pacotes ou ao armazenamento em formato digital.

A Figura 7 apresenta a relação entre as camadas VCL e NAL.

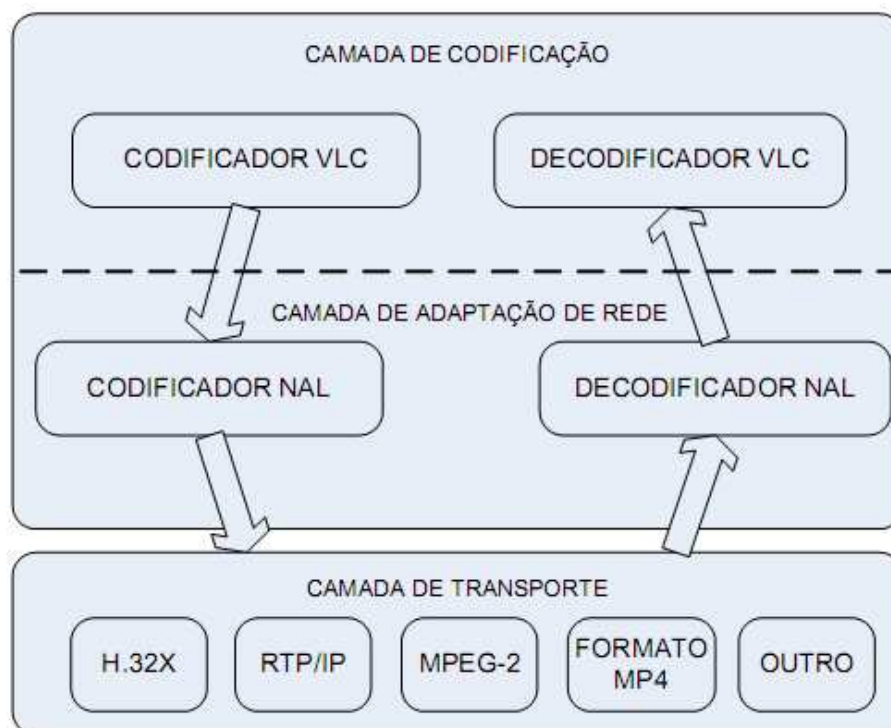


Figura 7: Arquitetura genérica da norma H.264/AVC.

Fonte: Soler (2006).

De acordo com a Figura 3.7, o transporte e o encapsulamento dos dados codificados pela norma H.264/AVC nos diferentes sistemas de transporte, tais como H.32x, MPEG-2 e RTP/IP, estão fora do âmbito da norma e devem ser especificados pelos órgãos internacionais responsáveis (IETF para o RTP/IP). A camada NAL funciona como uma abstração entre o VLC e a rede utilizada no transporte do *bitstream* codificado.

### 2.2.3 Padrão H.264 SVC

O projeto sobre a padronização SVC (JVT - a,2005)(JVT - B,2005)(JVT - C,2005) foi originalmente iniciado pela ISO / IEC MPEG. Baseado em uma avaliação das propostas apresentadas, o MPEG e o ITU-T VCEG concordaram em finalizar em conjunto o projeto de extensão SVC da norma H.264/MPEG4-AVC. O padrão H.264 SVC tem possibilitado a obtenção de melhorias significativas na eficiência utilizando uma codificação com maior grau de escalabilidade em relação as escalabilidades propostas nas normas de codificação anteriores (Schwarz,2007).

A extensão SVC permite a transmissão e decodificação de fluxos de vídeo com taxas de bits inferiores à taxa correspondente ao vídeo completo, para fornecer serviços de vídeo com menor resolução SNR, temporal ou espacial, mantendo uma qualidade elevada na reconstrução do vídeo em relação à taxa de *bits* do *bitstream* considerada. Assim, a SVC oferece diversas funcionalidades, como a suavização da degradação das perdas de transmissão, o poder de adaptação às condições dos receptores e da rede e à baixa taxa de *bits*. Essas funcionalidades fornecem melhorias no transporte e armazenamento digital dos sinais de vídeo.

Em termos de tipos de quadros, a codificação H.264 SVC permite que quadros do tipo B baseiem-se em outros quadros do tipo B. Esta característica é chamada de hierarquia de quadros B (*B-frames*) (ITU-T H264,2007).

Uma importante característica da concepção da SVC é que a maioria dos componentes do H.264/MPEG4-AVC é mantida conforme especificado na norma, contemplando a VCL e a NAL. A camada base de um *bitstream* SVC pode ser decodificada conforme norma H.264/MPEG-4 AVC. Novas ferramentas foram acrescentadas para apoiar principalmente as escalabilidades SNR e espacial.

Na norma H.264/SVC, os conceitos de codificação VCL e de rede NAL são herdados da norma H.264/AVC. O VCL contém as funcionalidades de processamento do codificador, tais como os mecanismos de transformação, quantização, entropia, predição intra-quadro e predição inter-camada. A NAL encapsula cada fatia gerada pela VCL em unidades NAL. Cada camada é formada por unidades SVC NAL, que representam os *bits* de vídeo codificados da camada. Um fluxo RTP, por exemplo, que transporta um *bitstream* completo de vídeo escalável, pode conter unidades NAL de uma camada base e uma ou mais camadas de reforço (Wenger, 2006).

## 2.3 HISTÓRICO DA ESCALABILIDADE DE VÍDEO

Nos padrões internacionais, a escalabilidade temporal foi introduzida pela primeira vez no padrão MPEG-1 (ISO - *International Organization for Standardization* / IEC - *International Electrotechnical Commission* JTC1, 1992) com o conceito de quadros do tipo B (*B-frame*), segundo o qual quadros B podem formar uma camada de reforço ao serem transportados em um canal de transporte independente (Wenger, 2006).

O padrão MPEG-2 (ITU-T - *International Telecommunications Union - Telecommunications* e ISO / IEC JTC1, 1994) manteve esta propriedade de quadros do tipo B com base de escalabilidade temporal, e acrescentou, no seu raramente utilizado perfil escalável, mais duas formas de escalabilidade: a espacial e a *Signal-to-Noise Ratio* (SNR). A escalabilidade espacial utiliza uma camada de reforço em complemento a uma camada base codificada em uma resolução espacial menor. A escalabilidade SNR utiliza um mecanismo similar, exceto que as resoluções da base e as camadas de reforço são idênticas. Normalmente, uma camada de reforço SNR refina a camada base para melhorar a fidelidade na reconstrução da imagem.

Todas as tecnologias mencionadas acima são comumente classificadas como escalabilidade de granularidade grossa ou escalabilidade de granularidade grossa ou *Coarse-Grain Scalability* (CGS)(Coney, 2006).

O MPEG-4 (ISO / IEC JTC1, 1998) mantém todos os mecanismos CGS mencionados, e acrescenta um conceito conhecido como escalabilidade de granularidade fina ou *Fine Grain Scalability* (FGS)(Winken,2005) ou escalabilidade de granularidade fina. A FGS permite a reconstrução de uma fatia de dados onde certo número de *bits*, no final de cada fatia, pode ser perdido. Uma das razões para a remoção dos *bits* FGS em camadas de reforço é a necessidade de aderir aos princípios de controle do congestionamento nas redes IP (Wenger, 2006).

Quando o *Video Coding Experts Group*, formalmente conhecido como ITU-T SG16 Q.6, discutiu os requisitos do que se tornou mais tarde a norma

H.264/AVC, a escalabilidade foi excluída. Talvez a principal razão seja que, naquele momento, considerou-se que a falta de sucesso de escalabilidade como um conceito não justificaria o esforço.

Com a finalização do padrão H.264/AVC versão 1, o foco da investigação ficou centrada na codificação de vídeo escalável. Em Janeiro de 2005, uma primeira proposta de trabalho da extensão SVC foi criada (JVT - a,2005). Este projeto e seus sucessores são comumente conhecidos como SVC (Wenger, 2006).

A Tabela 4 apresenta os padrões de codificação de vídeo que incluem recursos de escalabilidade de vídeo.

Tabela 4: Escalabilidade de vídeo nos padrões de codificação.

<b>Tipo de Escalabilidade</b>	<b>MPEG-2 / H.262</b>	<b>H.263+</b>	<b>MPEG-4 Parte 2 Visual</b>	<b>MPEG-4 Parte 10 AVC /H.264 AVC</b>	<b>H.264 SVC</b>
<b>Espacial</b>	X	X	X		X
<b>Temporal</b>	X	X	X	X	X
<b>Qualidade</b>	X	X	X		X

Fonte: Bottreau, (2007).

Exceto para a escalabilidade temporal, a utilização de escalabilidades de vídeo leva à necessidade de significativa taxa de transmissão adicional e complexidade de implementação, o que dificultou sua aceitação pelo mercado (Bottreau, 2007).

## 2.4 TIPOS DE ESCALABILIDADE

As técnicas de codificação escalável podem ser divididas em três classes: a escalabilidade espacial, temporal e a de qualidade ou SNR. Quando se combinam diversos tipos de características, pode-se alcançar um número superior de representações úteis, o que resulta num novo tipo de escalabilidade: a escalabilidade combinada ou híbrida.

### 2.4.1 Escalabilidade Temporal

A escalabilidade temporal permite a extração de fluxos de vídeo correspondentes a diferentes resoluções temporais a partir do fluxo de vídeo completo. A decodificação da camada base resulta numa versão do vídeo com baixa resolução temporal, enquanto que a decodificação progressiva das camadas de reforço permite um aumento gradual da resolução temporal, como exemplificado na Figura 8.

A codificação escalável deve permitir pelo menos uma ótima qualidade visual na resolução temporal máxima, ao mesmo tempo que permite que uma qualidade visual aceitável seja obtida a partir das resoluções temporais mais baixas. O uso das técnicas de codificação que exploram a correlação entre os quadros (compensação de movimento) e a codificação dos vetores de movimento, permite que se possa extrair diversas resoluções temporais de um único vídeo (Conklin,2000).

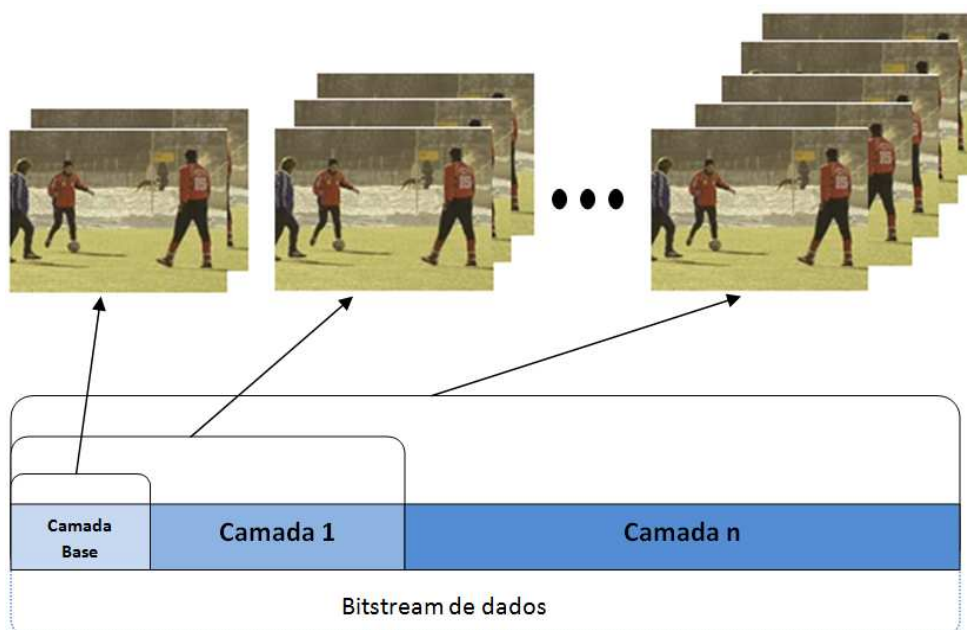


Figura 8: Escalabilidade temporal em camadas.

Fonte: Soler(2006).

A escalabilidade temporal distribui os quadros em várias camadas.

A Figura 9 abaixo apresenta a escalabilidade dos quadros nas camadas temporais:

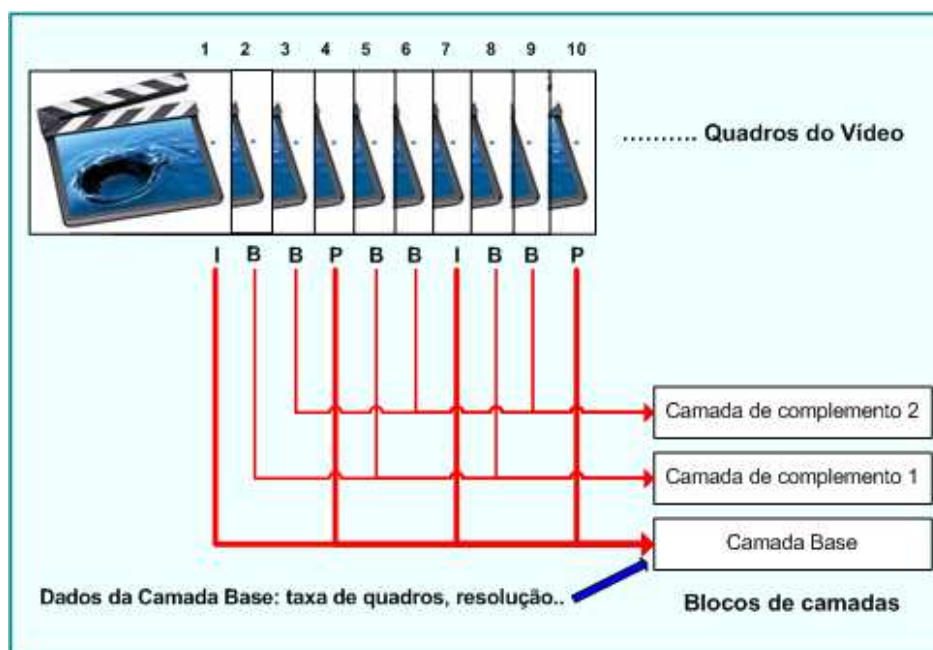


Figura 9: Escalabilidade dos quadros nas camadas temporais.

Fonte: Mayer(2005).

Os fluxos de *bits* com escalabilidade temporal podem ser gerados usando estruturas hierárquicas, conforme ilustrado na Figura 9. Os quadros do tipo I são codificados em intervalos regulares. Os quadros entre os quadros I são hierarquicamente previstos como apresentado na Figura 3.3. Os quadros do tipo P que representam a sequência da imagem se apóiam na resolução temporal a partir dos quadros I. Os quadros do tipo B se apóiam na resolução temporal a partir dos quadros I e/ou P (Schwarz,2006)e, quando na codificação H.264 SVC, podem apoiar-se também em outros quadros do tipo B (ITU-T H264,2007).

## 2.4.2 Escalabilidade Espacial

A escalabilidade espacial permite a extração de *bitstreams* com diferentes resoluções espaciais a partir do *bitstream* completo. Esta técnica permite oferecer conteúdos a terminais com diferentes características (*displays*) em termos de resolução espacial, tais como celular, TV com *High-Definition* (HD), dentre outros, a partir do mesmo *bitstream*.

Na Figura 10, o *bitstream* encontra-se dividido em três camadas de escalabilidade espacial. Com a decodificação da primeira camada, o usuário obtém uma versão da imagem ou vídeo original com a menor resolução espacial possível. A decodificação da segunda camada permite obter mais informações que,

adicionadas à primeira camada, resultam numa imagem reconstruída com uma maior resolução espacial. À medida que se decodificam mais camadas, a imagem reconstruída passa a ter uma resolução espacial cada vez maior, até o nível de resolução espacial máxima. Quando se decodificam mais camadas de reforço, o usuário obtém aumentos sucessivos da resolução espacial da imagem até a resolução original ou resolução máxima estabelecida na fase de codificação (Soler, 2006).

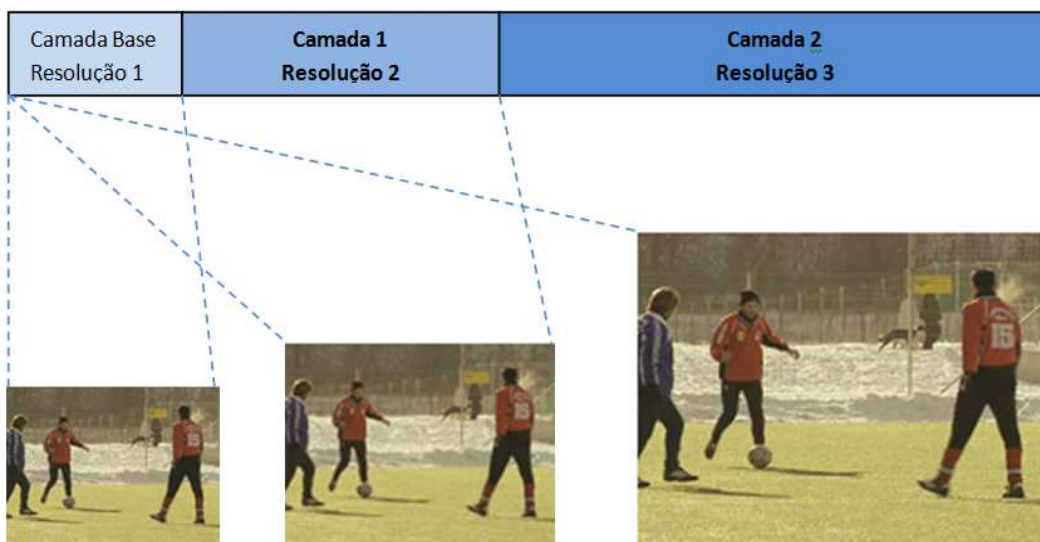


Figura 10: Escalabilidade espacial em camadas.

Fonte: Soler (2006).

### 2.4.3 Escalabilidade SNR

Este tipo de escalabilidade também é chamado escalabilidade de qualidade, uma vez que o erro de decodificação está relacionado com a qualidade da imagem. A SNR é a relação sinal-ruído, onde quanto maior o SNR, maior a qualidade do vídeo. A escalabilidade SNR permite a extração de *bitstreams* correspondentes a diferentes níveis de qualidade a partir do *bitstream* completo. Neste caso, há um aumento da qualidade da imagem ou vídeo sem necessariamente haver variação das resoluções espacial e temporal. Este tipo de escalabilidade codifica sucessivamente o erro de codificação entre a imagem original e a sua reconstrução numa dada camada (Schwarz, 2007) (Wei, 2008).

A decodificação da camada base fornece uma versão de baixa qualidade, em termos de SNR, do vídeo em questão. A sucessiva decodificação das camadas

de reforço resulta num aumento da qualidade, ou seja, da SNR, até o nível de qualidade máxima possível para aquela taxa de *bits*. A Figura 11 apresenta as camadas SNR no fluxo de vídeo.

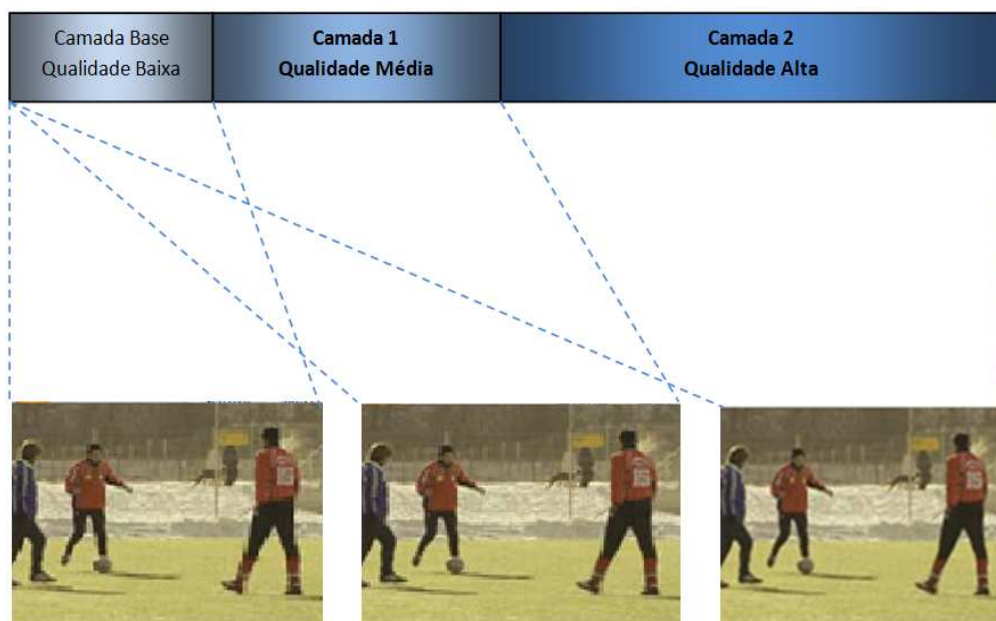


Figura 11: Escalabilidade SNR em camadas.

Fonte: Soler (2006).

#### 2.4.4 Combinação dos tipos de Escalabilidade

Os codificadores de vídeo escaláveis podem combinar os tipos de escalabilidade (espacial, temporal e SNR) para formar qualquer tipo composto. Este tipo de escalabilidade é referido na literatura como **combinada** ou **híbrida** e pode combinar dois ou mais tipos de escalabilidade previamente apresentados. Os tipos híbridos mais utilizados são a espacial/SNR para imagens e a espacial/temporal ou SNR/temporal para seqüências de vídeo.

A principal vantagem da combinação de técnicas básicas de codificação escalável consiste na criação de um *bitstream* com uma granularidade mais elevada, ou seja, com diversos níveis de qualidade. Também permite uma maior flexibilidade por parte do decodificador, pois este pode utilizar mais do que uma característica (resolução, taxa de *bits* ou qualidade) para representar uma imagem ou vídeo. A utilização de mais de uma dimensão de escalabilidade para representar uma imagem ou vídeo é obtida à custa de um acréscimo de

complexidade no codificador e no decodificador. A escolha da combinação de escalabilidade depende da aplicação em questão e dos seus requisitos.

No transporte de vídeo pela *Internet*, a escalabilidade SNR/temporal é a mais utilizada, uma vez que a escalabilidade espacial possui uma complexidade maior (na norma de vídeo MPEG-2) (ITU-T H264,2007).

## 2.5 NÍVEL DE GRANULARIDADE

Uma importante medida na classificação das técnicas de escalabilidade é o nível de granularidade de um *bitstream* escalável, pois indica o número de representações úteis (número de camadas) que podem ser extraídas do mesmo *bitstream*. As técnicas de codificação escalável, referentes à granularidade, são classificadas em três categorias: (1) escalabilidade de baixa granularidade, com um número limitado de níveis de granularidade; (2) escalabilidade de elevada granularidade ou com variação contínua na taxa de *bits*, com um número elevado de níveis de granularidade e (3) a escalabilidade híbrida, que combina a de baixa granularidade (1) com a elevada ou contínua (2) (Conklin,2000).

### 2.5.1 Escalabilidade de granularidade baixa

Na escalabilidade de baixa granularidade o *bitstream* pode ser decodificado segundo um conjunto limitado e pequeno de taxa de *bits* previamente estabelecidos durante a codificação. O esquema de codificação comprime a sequência de vídeo em várias camadas, sendo uma dessas a camada base, que pode ser decodificada independentemente e fornece a qualidade visual mínima. As outras camadas são camadas de reforço ou de melhoria e, apesar de poderem ser decodificadas de forma independente, só podem ser úteis para melhorar a qualidade ou resolução da imagem quando todas as camadas hierarquicamente inferiores forem também decodificadas.

A combinação de todas as camadas fornece a qualidade mais alta do *bitstream*. A decodificação da camada base ou qualquer subconjunto de camadas tem sempre qualidade visual inferior à do *bitstream* total. Esta técnica permite que a camada base seja trabalhada com técnicas mais robustas de codificação de

canal ou, em uma rede que suporte serviços diferenciados, a camada base pode ser transportada com uma prioridade mais elevada.

Fazem parte desta categoria os modos de escalabilidade oferecidos pelas normas MPEG-2 Vídeo e MPEG-4 Visual, tais como a escalabilidade de granularidade grossa ou *Coarse-Grain Scalability* (CGS) e a granularidade de escalabilidade média ou *Medium Grain Scalability* (MGS).

Segundo (Schwarz,2007) (Wei, 2008), a MGS é uma variação da CGS, disponibilizada na extensão SVC, com o objetivo de prover um aumento na flexibilidade de adaptação do fluxo de *bits*. A MGS utiliza o identificador de qualidade Q para controle da qualidade e permite uma adaptação do fluxo de vídeo baseado em unidades NAL de qualidade (Wien, 2007). A MGS será utilizada na codificação dos vídeos transmitidos nos experimentos deste trabalho.

### **2.5.2 Escalabilidade de granularidade elevada**

Esta técnica é muito flexível e permite ao servidor de vídeo adaptar a taxa de *bits* do vídeo em distribuição às disponibilidades da rede usando uma granularidade muito fina, isto é, com grande eficiência, pois todos os *bits* recebidos se tornam úteis. Para que a escalabilidade seja contínua na taxa de *bits* é necessário que todos os dados comprimidos sejam embutidos em um único *bitstream* e de forma a permitir que diferentes taxas de *bits* possam ser obtidas. O decodificador recebe os dados comprimidos desde o princípio do *bitstream* até o ponto onde a taxa de *bits* escolhida seja alcançada. As imagens codificadas com este tipo de escalabilidade podem ser decodificadas progressivamente, já que o decodificador apenas precisa receber um conjunto muito pequeno de dados para começar a visualizar a imagem. Na compressão de uma única imagem, os *bits* que possuem a informação mais importante são incluídos no início do *bitstream*, de forma que a qualidade visual seja maximizada para todas as taxas de *bits*(Conklin,2000).

Fazem parte desta categoria de escalabilidade de granularidade as normas *Joint Photographic Experts Group* (JPEG), JPEG2000 e o modo *Visual Texture Coding* (VTC) da norma MPEG-4 Visual.

### 2.5.3 Escalabilidade híbrida

Esta categoria combina a escalabilidade de baixa granularidade com a escalabilidade contínua. Este esquema de codificação comprime uma seqüência de vídeo em duas camadas: uma camada base e uma camada superior com escalabilidade contínua. No entanto, enquanto na escalabilidade de baixa granularidade a camada superior tem de ser totalmente recebida e decodificada em conjunto com a camada base para que haja melhoria de qualidade, na escalabilidade híbrida a camada superior pode ser truncada em qualquer ponto. A melhoria de qualidade é proporcional ao número de *bits* que o decodificador utiliza da camada superior (Schwarz, 2006).

A escalabilidade de granularidade fina do MPEG-4 Visual, chamada *Fine-Grain Scalability* (FGS) (Winken,2005) (ISO/IEC 14496-2, 2002) faz parte desta categoria.

## 2.6 ESCALABILIDADE NOS PADRÕES H.264 AVC/SVC

### 2.6.1 Escalabilidade no padrão H.264/SVC

A norma H.264/SVC atinge altas taxas de compressão e permite uma ampla gama de combinação de diferentes tipos de escalabilidade (Schwarz, 2007). Assim como as normas anteriores, a SVC distingue uma camada de base de camadas de reforço. Na norma H.264/SVC, a camada de base está prevista para ser igual à camada da norma H.264/AVC. As camadas de reforço seguem a especificação da norma H.264/SVC.

Uma camada de reforço pode aumentar a resolução temporal, a resolução espacial, ou ainda a qualidade do conteúdo de vídeo representada pela camada inferior ou parte dela. Conforme a especificação do padrão H.264/SVC, as camadas escaláveis podem ser agregadas a um único fluxo de transporte ou transportadas de forma independente (Wenger, 2006). Neste trabalho, será adotado um único fluxo de transporte.

A escalabilidade temporal, na norma SVC, é apoiada por uma estrutura de predição temporal hierárquica. Nestas estruturas, os quadros principais são codificados em intervalos regulares, utilizando somente quadros-chave (do tipo I ou P) anteriores como referência. Entre os quadros-chave, ficam os quadros do tipo B, que são bi-direcionalmente preditivos, que podem ser baseados em quadros I, P ou ainda outros B. A camada base contém uma seqüência de quadros-chave (do tipo I e P) enquanto que as camadas de reforço são constituídas hierarquicamente de quadros do tipo B (Wenger, 2006). Conforme apresenta a Figura 12, todos os quadros entre os quadros-chave são dispostos em uma estrutura hierárquica com incremento de quadros B das camadas de reforço bi-direcionalmente entre os quadros-chave. A redução ou incremento da taxa de quadro pode ser obtido incrementando ou decrescendo camadas (Anselmo,2008). Os quadros na cor preta (I ou P) formam a camadas-base, o quadro azul (B1) é a primeira camada complementar, os quadros vermelhos (B2) formam a segunda camada complementar e os quadros verdes (B3) formam a terceira camada complementar.

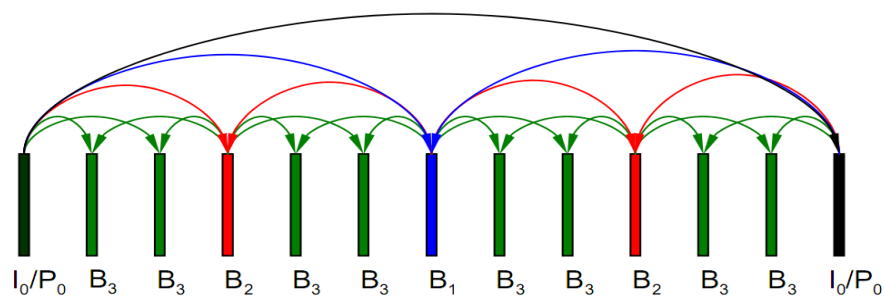


Figura 12: Estrutura hierárquica da escalabilidade temporal.

Fonte: Schwarz(2007).

A Figura 13 abaixo apresenta a hierarquia da escalabilidade dividida em camadas.

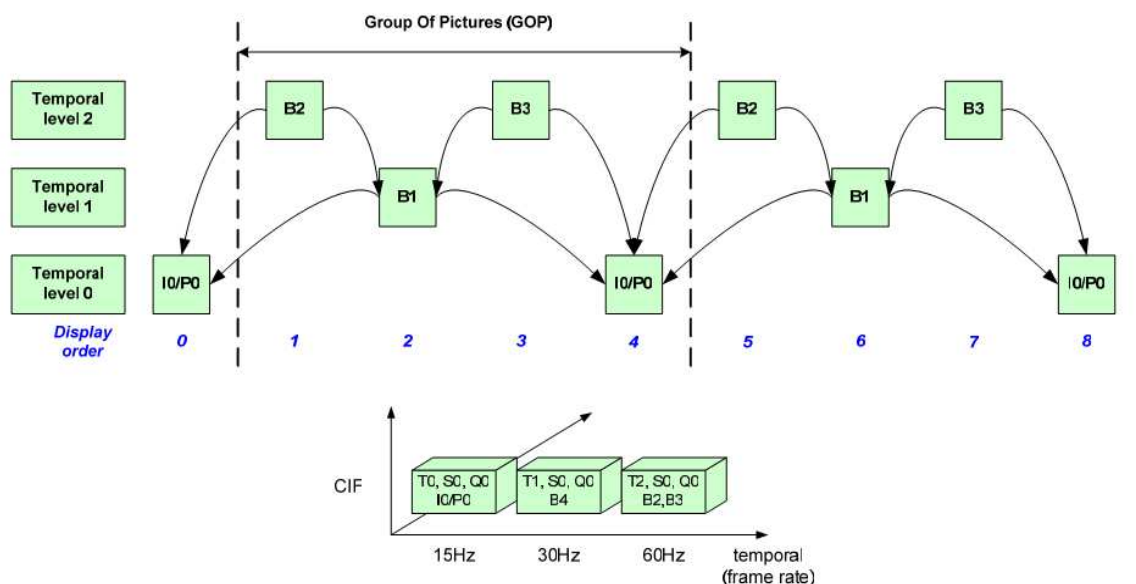


Figura 13: Hierarquia dos quadros da escalabilidade temporal em camadas.

Fonte: Bottreau(2007).

A escalabilidade espacial na SVC permite codificar diferentes resoluções espaciais escaláveis em várias camadas dentro de um único *bitstream*, melhorando a eficiência da codificação do vídeo em comparação com os *streams* separados para cada resolução espacial da norma H.264/AVC. A camada base é codificada em uma baixa resolução espacial e as demais camadas de reforço fornecem imagens de resolução mais elevada em termos de tamanho da tela (Anselmo,2008). A Figura 14 apresenta a hierarquia dos quadros de escalabilidade espacial em três camadas.

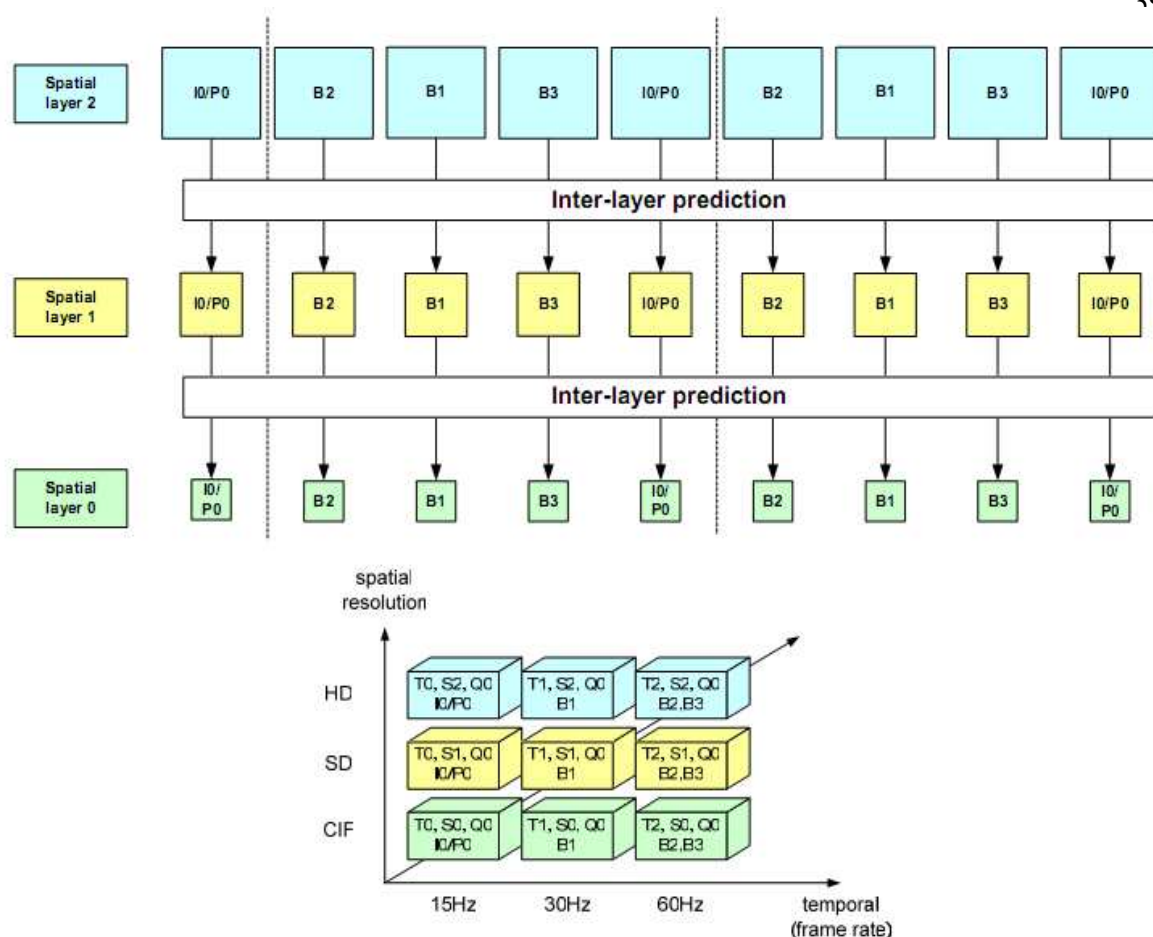


Figura 14: Hierarquia dos quadros da escalabilidade espacial em camadas.

Fonte: Bottreau(2007).

Na Figura 14, para diferentes tipos de resolução espacial (HD – “High Definition”, SD – “Standard Definition” e CIF – “Common Intermediate Format”) e com base em algumas taxas de quadros, são apresentadas as camadas de reforço de resolução espacial (“spatial resolution”) ou S0 para CIF, S1 para SD e S2 para HD. Além disso, são representadas três camadas de escalabilidade temporal, sendo T0 para 15Hz, T1 para 30Hz e T2 para 60Hz. Nesta figura, a escalabilidade de qualidade ou SNR está representada apenas em uma camada Q0.

As cores representam a camada espacial à qual o pacote pertence, sendo que a cor verde representa a camada básica de resolução CIF (S0), a cor amarela representa a camada de reforço de resolução espacial SD (S1) e a cor azul representa a camada de reforço de resolução espacial HD (S2).

Na norma H.264/SVC, a escalabilidade de qualidade ou SNR permite a codificação de mais camadas com a mesma resolução temporal e espacial, mas com o aumento da qualidade do vídeo decodificado, avaliada em termos de pico da relação da potência do sinal pela potência do ruído ou *Peak Signal-to-Noise Ratio* (PSNR), pela correta escala de quantização da camada base até a mais alta camada de reforço (Kodikara,2008). A escalabilidade SNR no padrão H.264/SVC suporta a escalabilidade de granularidade grossa (CGS) e a escalabilidade de granularidade média (MGS). A escalabilidade de granularidade fina não foi incluída na versão inicial do SVC. Atualmente, pesquisas estão em curso para incluir FGS em uma continuação da extensão SVC (Van der Auwera, 2008).

Todos os tipos de escalabilidade podem ser combinados para melhorar a eficiência e a codificação a fim de se adaptarem às especificidades dos requisitos pedidos. A Figura 15 apresenta a hierarquia das camadas base e de reforço com as escalabilidades temporal, espacial e SNR. O quadro verde indica a camada base, com a escalabilidade temporal (T) igual a zero, a escalabilidade espacial (S) igual a zero e a escalabilidade SNR (B) igual a zero. A camada complementar 1 está apresentada em amarelo, a camada complementar 2 está apresentada em azul e a camada complementar 3 está apresentada na cor branco.

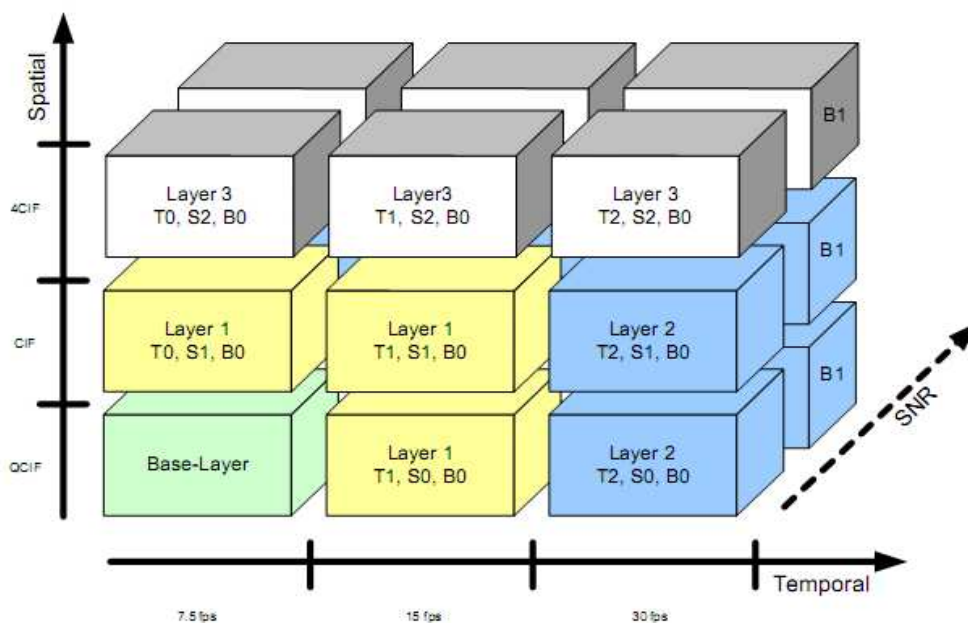


Figura 15: Escalabilidades temporal, espacial e SNR nas camadas base e de reforço.

Fonte: Mayer(2005).

Na Figura 15, para diferentes tipos de resolução espacial (4CIF – “Four Times Common Intermediate Format”, CIF – “Common Intermediate Format” e QCIF – “Quarter Common Intermediate Format”) e com base em algumas taxas de quadros, são apresentadas as camadas de reforço de resolução espacial (“spatial resolution”) ou S0 para 4CIF, S1 para CIF e S2 para QCIF. Além disso, são representadas três camadas de escalabilidade temporal, sendo T0 para 7.5Hz, T1 para 15Hz e T2 para 30Hz. Nesta figura, a escalabilidade de qualidade ou SNR está representada em apenas duas camadas B0 e B1.

As cores representam a que camada o pacote pertence, sendo que a cor verde representa a camada básica (*Base-Layer*), a cor amarela representa a primeira camada de reforço (*Layer 1*), a cor azul representa a segunda camada de reforço (*Layer 2*) e a cor branca representa a terceira camada de reforço (*Layer 3*).

A Figura 16 apresenta uma imagem de alta resolução, alta taxa de quadros e alta qualidade e identifica a hierarquia das camadas base e de reforço com as escalabilidades temporal, espacial e SNR completa que geraram esta imagem.

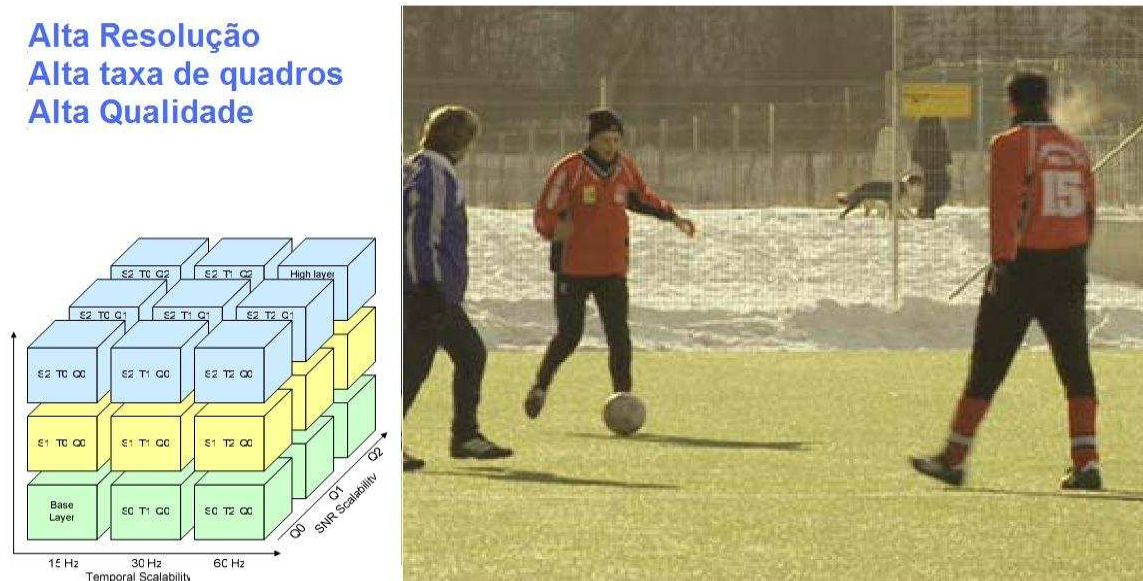


Figura 16: Imagem de Alta Resolução, Alta Taxa de Quadros e Alta Qualidade.

Fonte: Bottreau(2007).

A Figura 17 apresenta uma imagem de alta resolução, alta taxa de quadros e baixa qualidade e identifica a hierarquia das camadas base e de reforço com as escalabilidades temporal, espacial e SNR que geraram esta imagem. Note que apenas uma camada de qualidade foi utilizada.

**Alta Resolução  
Alta Taxa de quadros  
Baixa Qualidade**

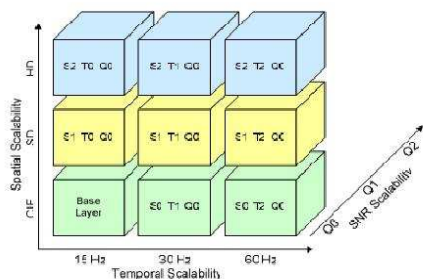


Figura 17: Imagem de Alta Resolução, Alta Taxa de quadros e Baixa Qualidade.

Fonte: Bottreau(2007).

A Figura 18 apresenta uma imagem de resolução média, taxa de quadros média e qualidade média e indica a hierarquia das camadas base e de reforço com as escalabilidades temporal, espacial e SNR que geraram esta imagem. Note que apenas duas camadas de cada tipo de escalabilidade foram utilizadas.

**Média Resolução  
Média Taxa de quadros  
Média Qualidade**

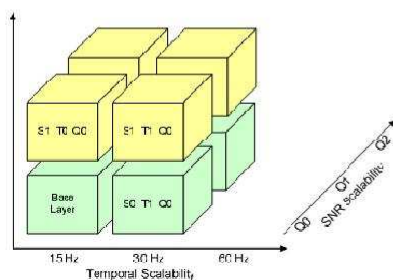


Figura 18: Imagem de Resolução Média, Média Taxa e quadros e Qualidade Média.

Fonte: Bottreau(2007).

A Figura 19 apresenta uma imagem de baixa resolução, alta taxa de quadros e baixa qualidade e indica a hierarquia das camadas base e de reforço com as escalabilidades temporal, espacial e SNR que geraram esta imagem. Note que

apenas uma camada de escalabilidade espacial foi usada, gerando a resolução baixa, duas camadas de escalabilidade temporal foram utilizadas, gerando a taxa de quadros média e todas as camadas de SNR ou qualidade foram utilizadas, gerando uma imagem de alta qualidade.

**Baixa Resolução**  
**Média Taxa de quadros**  
**Alta Qualidade**

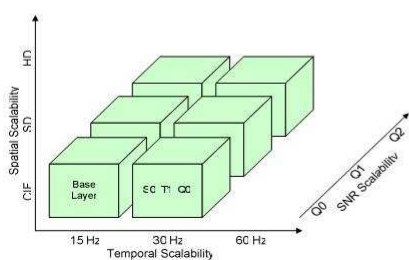


Figura 19: Imagem de Baixa Resolução, Média Taxa de quadros e Alta Qualidade.

Fonte: Bottreau(2007).

## Capítulo 3 - QUALIDADE DE EXPERIÊNCIA – QoE

Embora as medidas em nível de rede permitam garantir os níveis de Qualidade de Serviço (QoS), elas não permitem estimar a qualidade de acordo com a percepção do usuário. Sendo assim, desenvolver um sistema que tenha qualidade de acordo com a percepção do usuário exige que outros aspectos além da QoS (nível de rede) sejam considerados. Ao sistema que agregue, além de métricas objetivas, um conjunto de métricas subjetivas que permitam caracterizar a qualidade percebida pelo usuário, dá-se o nome de Qualidade de Experiência (QoE) (Silva,2007).

Assim, além das métricas objetivas que, tradicionalmente, ajudam a caracterizar o desempenho de uma rede (*jitter*, perdas e latência), torna-se necessário considerar outros aspectos, principalmente quando se trata de aplicações interativas, pois estas têm impacto na Qualidade de Experiência de utilização, tais como o sincronismo entre o áudio e o vídeo e o tempo de inicialização de uma aplicação, dentre outros.

### 3.1 DEFINIÇÕES DE QUALIDADE DE EXPERIÊNCIA (QoE)

O conceito de QoE pode ser considerado relativamente novo para as áreas de conhecimento que envolvem multimídia.

Diferentes definições estão disponíveis na literatura sobre QoE. Por exemplo, em (Patrick,2004), QoE é definida como "as características das sensações, das percepções e as opiniões das pessoas e como elas interagem com seus ambientes. Estas características podem ser gratificantes e divertidas, ou desagradáveis e frustrantes".

Em (O' Neil,2002), QoE é definida como "a totalidade dos mecanismos de Qualidade de Serviço, para assegurar uma boa transmissão de áudio e vídeo sobre redes IP".

Em (Goodchild,2005), é afirmado que "QoE é o desempenho global de um sistema a partir do ponto de vista dos usuários. QoE é uma medida de fim-de-fim em níveis de desempenho e de um indicador do usuário do quão bem o sistema satisfaz as necessidades dos utilizadores ". Em (Goodchild,2005), Goodchild também faz uma distinção entre QoS e QoE no sentido de que QoE abrange a perspectiva do usuário, enquanto QoS representa a perspectiva de rede.

Em (Siller,2003), QoE é definida como "a experiência percebida do usuário do que é apresentado pela camada do aplicativo, onde a camada de aplicação atua como uma interface de usuário que apresenta o resultado global da Qualidade dos Serviços ".

Segundo (kishigami,2007), a definição da QoE é "A global aceitabilidade de um aplicativo ou serviço, tal como o entendimento subjetivo do usuário final". Sendo assim, a QoE é uma noção mais ampla em comparação com a QoS. A QoE irá incluir o total dos efeitos do sistema de ponta-a-ponta. Além disso, a QoE é influenciada pela expectativa e contexto de cada usuário. Assim, a simples classificação não é adequada para este conceito de contexto-dependência (kishigami,2007).

Observa-se que essas definições frisam que a QoE tem uma abrangência maior quando comparada à QoS. Poder-se-á considerar a QoE como uma pseudo-camada (Siller,2003), que está de acordo com a experiência do usuário final, ou como uma extensão para a camada de aplicação definida no modelo OSI (Bauer, 2009). A QoE pode ser então considerada como uma extensão do âmbito tradicional da QoS no sentido de que a QoE fornece informações relativas aos serviços prestados a partir do ponto de vista do usuário final (Lopez, 2006).

Segundo (Jain,2004), experiência é subjetiva. Todos nós temos uma base cultural, situação sociocultural e experiências pessoais diferentes, que alteram a forma de reagirmos a cada situação, como a reação a um filme, por exemplo, ou a uma interface *web*. Às vezes, apenas a troca de uma cor numa interface pode gerar uma experiência diferente ao usuário.

Experiência é, também, dependente de contexto. Alguns livros ou filmes podem gerar experiências diferentes para a mesma pessoa, dependendo do contexto, da situação particular daquela pessoa. Identificar fatores que determinam o contexto que influencia a experiência pode ser difícil. Adicionalmente, ainda pode haver aspectos externos como ambiente físico e questões pessoais, que dificultam a criação de regras claras de QoE (Jain,2004).

### **3.2 RELAÇÃO ENTRE QoS, QoE E QoV – QUALIDADE DE VÍDEO**

Como dito nas definições acima, a QoE refere-se à percepção do usuário sobre a qualidade de um serviço, baseado em sentimentos humanos como 'bom', 'excelente', 'insatisfatório', etc. Por outro lado, a QoS envolve a organização do tráfego da rede definindo prioridades e limites de forma a melhorar a percepção do usuário.

A QoS refere-se aos mecanismos aplicados no nível dos equipamentos de rede, e a QoE é um conceito mais amplo, que engloba QoS e a opinião do usuário quanto a qualidade do serviço fim-a-fim.

A fim de satisfazer expectativas do usuário, a implementação de QoS deve ser centrada na perspectiva do usuário final, de forma a garantir a qualidade de serviço necessária para atender os níveis de QoE do usuário.

No contexto de *streaming* de vídeo, abordado neste trabalho, o principal fator para determinar o QoE, na perspectiva do usuário, é a QoV - Qualidade do Vídeo (Kuipers,2010) (Winkler,2008) (Winkler,2005). Este trabalho engloba apenas questões de QoE referentes a qualidade do vídeo percebido pelo usuário, sem que os demais aspectos sejam medidos.

### 3.3 QUALIDADE DE VÍDEO

Considerando as diferentes dimensões de QoE/QoS mencionadas no presente documento, a qualidade de vídeo é uma característica das mais importantes na perspectiva do usuário. Segundo (ITU-T J149,2004), existem dois conjuntos de métodos para medir e verificar a qualidade de vídeo entregue ao usuário final: são os métodos objetivos e subjetivos.

#### 3.3.1 Métodos Objetivos

Os métodos objetivos de avaliação (ITU-T J143,2000) referem-se à medição da qualidade de vídeo usando métricas objetivas, ou seja, instrumentais, para a obtenção de uma indicação que se aproxime da classificação que seria obtida a partir de um teste de avaliação subjetiva.

Inúmeros testes são realizados e os resultados são usados para criar um modelo de percepção de qualidade humana – o sistema visual humano. Este modelo prevê medidas objetivas de qualidade, calculados algoritmicamente e que podem ser integrados em rotinas automáticas de teste e análise. Embora métricas objetivas possam não refletir exatamente as percepções do usuário, elas têm a importante característica de apoiar o acompanhamento de implementações de serviços em grande escala, pois são reprodutíveis e podem ser realizados rapidamente.

Existem diversas abordagens possíveis para as medições objetivas (ITU-T J143,2000):

- **Referência completa ou *Full Reference* (FR):** o vídeo de referência na entrada do sistema, bem como o vídeo processado na saída do sistema estão disponíveis para comparação e podem determinar a qualidade de vídeo de forma objetiva. Considera-se uma abordagem mais rigorosa e, dada a exigência de disponibilidade de ambos os sinais de entrada e de saída, esta abordagem é também conhecida como "terminação dupla".

- **Referência Reduzida ou *Reduced Reference (RR)*:** esta abordagem difere de FR, pois só são comparados alguns parâmetros selecionados que são extraídos da entrada e da saída. Eles tentam manter a precisão e, ao mesmo tempo, otimizar o transporte de informações adicionais. A NTIA - *National of Telecommunication and Information Administration* definiu um modelo para RR, depois padronizado pela ANSI - *American National Standards Institute* (Ansi,2003), que necessita de parte do canal de vídeo como fonte de comparação.
- **Sem Referência ou *No Reference (NR)*:** utiliza apenas o sinal de vídeo recebido para determinar a qualidade do vídeo de forma objetiva. A vantagem é que ele não precisa de qualquer outra fonte de vídeo para medir qualidade, mas isso pode afetar o nível de precisão.

Em geral, o objetivo é utilizar métricas de avaliação de qualidade, tais como a MSE - *Mean Square Error* e a PSNR - *Peak Signal-to-Noise Ratio*, relativamente fáceis de computar (Robins,2009). Ambas as métricas são indicadores da diferença entre o vídeo recebido e uma referência de sinal, válida para as abordagens FR e RR.

Em função da relativa complexidade e do custo para medir a qualidade subjetiva, procurou-se uma maneira de realizar esta medida através de uma expressão matemática. A métrica mais usada é o PSNR (*Peak Signal-to-Noise Ratio*).

### 3.3.2 MSE - *Mean Square Error* ou Erro Médio Quadrático

Na estatística, a média de erro quadrática ou MSE de um estimador é uma das muitas formas de quantificar o montante pelo qual um estimador difere do valor verdadeiro da quantidade a ser estimada. O MSE realiza a média do quadrático do "erro". O erro é o montante pelo qual o estimador difere da quantidade a ser estimada. A diferença ocorre por causa de aleatoriedade ou porque o estimador não leva em conta informações que poderiam produzir uma estimativa mais exata.

Sejam I e K duas imagens contendo  $m \times n$  pixels, tais que uma é considerada uma aproximação com ruído da outra. A medida de MSE é dada pela Equação 3 abaixo:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i, j) - K(i, j))^2 \quad (3)$$

Em geral, os métodos objetivos são preferidos frente aos subjetivos, considerando a complexidade e custos associados a este último. Embora não exista uma solução definitiva que elimine totalmente a necessidade de uma referência (RN), uma estratégia possível é a combinação de diversos métodos para medir a qualidade de vídeo de forma objetiva (ITU-T J149,2004).

### 3.3.3 PSNR - Peak Signal-to-Noise Ratio

A PSNR é uma relação entre o máximo possível de potência de um sinal, pela potência do ruído, ao se comparar um sinal antes e depois de um processo de degradação, sendo que a unidade utilizada para representá-lo é o dB (decibel). Aplicando este conceito em vídeos e imagens, temos que a PSNR é a relação entre a entrada e a saída de um processo de compressão com perdas, que avalia o quanto de ruídos a compressão introduziu na imagem ou *quadro* original. A medida de PSNR é dada pela Equação 4 abaixo:

$$PSNR = 10 \cdot \log_{10} \left[ \frac{MAX_I^2}{MSE} \right] = 20 \cdot \log_{10} \left[ \frac{MAX_I}{\sqrt{MSE}} \right] \quad (4)$$

Onde MAX é o valor máximo possível de um *pixel*.

Quanto maior o valor da PSNR, maior é a relação entre a potência do sinal pela potência do ruído, o que significa melhor qualidade.

A PSNR pode ser usada como um bom indicador da variação da qualidade de vídeo quando o conteúdo e o padrão de codificação são os mesmos e sujeitos às mesmas condições de teste. Por outro lado, a PSNR é uma métrica de qualidade

de vídeo que pode não ser confiável considerando diferentes condições de ensaio.

Como dito anteriormente, a PSNR é uma métrica objetiva, porém foi mapeada, de forma heurística, a uma das mais populares métricas subjetivas, o MOS – *Mean Opinion Score* (Gross, 2004) (Piamrat,2009). Este mapeamento permite a correlação da métrica objetiva PSNR com a métrica subjetiva MOS, podendo assim, avaliar o QoE a partir da PSNR (Gross, 2004) (Rodrigues, 2008) (Kuipers,2010) (Vishwanath,2008) (Piamrat,2009).

Segundo o mapeamento com o MOS (Gross, 2004) (Piamrat,2009), os valores de PSNR acima de 37dB correspondem a compressões que introduzem perdas imperceptíveis ao olho humano, o que significa uma qualidade excepcional. Podemos considerar que vídeos com PSNR entre 31dB até 37dB tem qualidade boa, entre 25dB e 31dB teremos uma qualidade mediana, entre 20dB e 25dB teremos uma qualidade pobre e abaixo de 20dB a qualidade é ruim.

### 3.3.4 VQM – Video Quality Metric

O *Video Quality Metric* (VQM) é uma métrica objetiva de qualidade de vídeo que prediz as classificações de qualidade subjetiva que seriam obtidos a partir de avaliações de usuários (Martinez, 2007). Devido ao seu desempenho nos experimentos realizados pelo *Video Quality International Expert's Group* (VQEG) Fase II de validação dos ensaios, os métodos VQM foram aprovados pelo ANSI - American National Standards Institute\_ como um padrão (ANSI T1.801.03-2003), bem como as recomendações internacionais da ITU (ITU-T J.144, e ITU-R BT.1683, ambos aprovados em 2004).

O VQM (Wolf, 2009) foi desenvolvido pelo ITS - *Institute of Telecommunication Sciences* e pelo NTIA - National Telecommunications and Information Administration, e fornece uma medida objetiva para a qualidade de vídeo percebida. Ele mede os efeitos perceptivos de vídeo, incluindo ruído, distorção de bloco, distorção de cor, dentre outros, e combina-os em uma única métrica. Os resultados dos testes com VQM apresentam uma elevada correlação com as avaliações subjetivas do vídeo avaliado por usuários. Desta forma, o VQM pode

ser utilizado na avaliação de QoE (Wolf, 2009) (Martinez, 2007) (Rodrigues, 2008) (Kuipers,2010).

Para medir o VQM, segundo (Wang, 2006), o vídeo original e o processado são avaliados nos seguintes passos:

a) **Calibração:** neste ponto são calibrados os vídeos da amostra para extração dos cálculos. A calibração estima e corrige o deslocamento espacial e temporal, o contraste e compensa a luminosidade do vídeo processado em relação ao vídeo original;

b) **Extração de características de qualidade:** nesta etapa é extraído um conjunto de características da qualidade das mudanças perceptivas das propriedades espaciais e temporais da cor e dos blocos de vídeo;

c) **Parâmetros para cálculo de qualidade:** etapa em que se calcula um conjunto de parâmetros que descrevem a qualidade perceptível de mudanças na qualidade de vídeo, por meio da comparação de características extraídas do vídeo processado com aquelas extraídas do vídeo original e;

d) **Cálculo VQM:** é realizado utilizando uma combinação linear dos parâmetros calculados a partir das etapas anteriores.

O VQM está estruturado em cinco modelos objetivos de qualidade de vídeo: (1) geral, (2) televisão, (3) videoconferência, (4) desenvolvedores e (5) PSNR. O modelo geral foi concebido para ser um objetivo geral para sistemas VQM de vídeo que abrangem uma gama muito ampla de qualidade. O modelo de televisão foi especificamente otimizado para particularidades da TV, enquanto que o modelo de videoconferência foi especificamente otimizado para particularidades de videoconferência (por exemplo, H.263, MPEG-4). O modelo para desenvolvedores foi otimizado utilizando a mesma escala de qualidade de vídeo e de taxas de *bits* do modelo geral, adicionando fatores da computação (Martinez, 2007).

O modelo geral (VQM Geral), que será utilizado nos experimentos deste trabalho, foi concebido para medir a qualidade de vídeos e abrange uma vasta gama de

níveis de qualidade. O modelo geral tem parâmetros objetivos para medir os efeitos perceptuais de uma vasta gama de deficiências, como distorções de blocos, o ruído (em ambos os canais de luminância e crominância), blocos de erro, dentre outros.

O VQM Geral, também chamado de VQMg, utiliza uma combinação linear de sete parâmetros. Quatro parâmetros são baseados em elementos extraídos do gradiente espacial da luminância do componente Y, dois parâmetros são baseados em elementos extraídos do vetor formado pelos dois componentes de crominância (Cb, Cr), e um parâmetro é baseado no contraste e nas características temporais, ambas extraídas do componente de luminância Y(Wang, 2006).

Os valores de saída das medições de VQMg variam de 0, que indica a máxima correlação entre os vídeos, até 1, que indica a mínima correlação percebida entre os vídeos.

### 3.4 ABORDAGENS PARA QoE

Segundo (Coney, 2006), as abordagens para prover QoS e QoE nas transmissões de vídeo são classificadas em soluções baseadas em engenharia de tráfego, em protocolos de acesso e em esquemas adaptativos:

a) **As soluções baseadas em engenharia de tráfego:** utilizam basicamente o *Integrated Services (IntServ)*(Braden, 1994) (IETF, 2009) e *Differentiated Services (DiffServ)* (Nichols,1998) (Blake,1998) (Black, 1998) (QoS Forum, 2008) para intervir na camada de rede e prover QoS, e são geralmente usadas em redes controladas, como as redes de IPTV. O DiffServ é uma abordagem de QoS que é implementada localmente em cada roteador.

b) **As soluções baseadas em protocolo de acesso:** incluem mecanismos de QoS do padrão IEEE - *Institute of Electrical and Electronics Engineers* 802.11 (Zhu, 2004) (IEEE 802.11 WG, 2004). Podem ser classificadas em três categorias: diferenciação de serviço, controle de admissão e adaptação de *enlace*.

c) **Esquemas adaptativos:** Os esquemas adaptativos podem abranger uma ou mais técnicas de adaptação de conteúdo e/ou de rede.

Cabe observar que (Coney, 2006) não considera adaptações, por exemplo, na camada física (exemplo: modulação e codificação adaptativas) nem técnicas baseadas em *cross-layer*.

## Capítulo 4 - ADAPTAÇÃO DE CONTEÚDO E DE REDE

Neste Capítulo são apresentados os conceitos e classificações de adaptação de conteúdo e de rede. A adaptação de conteúdo é um conceito que envolve mais do que a execução de uma ou mais operações de transformação sobre o conteúdo. Seu objetivo é utilizar informações descritivas sobre o conteúdo e o contexto em uso para fornecer, de forma tão eficiente quanto possível, a variação para o índice mais adequado para o cenário atual (Carvalho, 2004).

As técnicas de adaptação de rede ou *Network Adaptation Techniques* (NAT) trabalham, principalmente, com adaptação da largura de banda em aplicações multimídia de tempo real, utilizando algoritmos que consideram o estado e/ou carga da rede (Gotz, 2004).

### 4.1 ADAPTAÇÃO DE CONTEÚDO

As técnicas de adaptação de conteúdo ou *Content Adaptation Techniques* (CAT), no contexto de transmissão de *streaming* de vídeo, são métodos comumente utilizados para a adaptação de conteúdos à taxa desejável (Liu, 2003). A Figura 20 apresenta uma arquitetura conceitual genérica para adaptação de conteúdo multimídia que é composta por um mecanismo de adaptação. A entrada para este módulo consistindo de conteúdos multimídia originais e de informação descritiva dos vários fatores que são esperados para condicionar a adaptação. O resultado (saída) é o conteúdo original adaptado para as exigências do contexto do usuário.

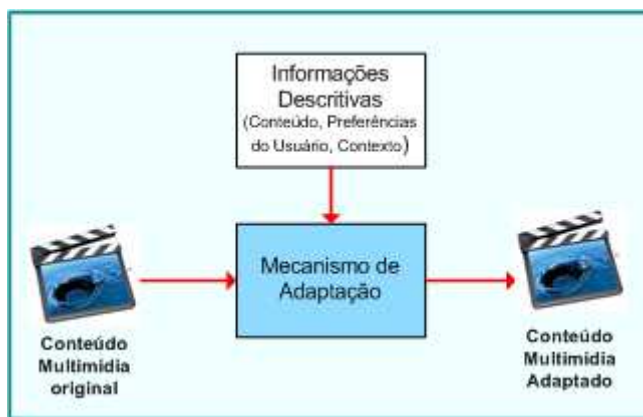


Figura 20: Arquitetura conceitual genérica para adaptação de conteúdo multimídia.

Fonte: Carvalho(2004).

O mecanismo de adaptação é responsável pela análise das informações recebidas do contexto, a decisão sobre a forma mais adequada para adaptar e a execução e consequente adaptação do conteúdo (Carvalho, 2004).

#### 4.1.1 Classificação das Técnicas de Adaptação de Conteúdo

##### 4.1.1.1 Primeira Classificação (Zhang, 2000)

Segundo Zhang(Zhang, 2000), as categorias mais relevantes para adaptação de conteúdo são:

###### 4.1.1.1.1- Classificação da finalidade

É a qualificação e classificação da finalidade de cada objeto de mídia de modo que seja possível melhorar a eficiência da prestação de informações por uma remoção de objetos redundantes, ou priorizando de acordo com sua importância. Uma típica página da *Web* contém uma grande quantidade de informação e objetos de mídia que são redundantes ou podem não ser do interesse de um usuário. Por exemplo, um *site* de *e-commerce* pode ter várias imagens para ligação ao mesmo *site* ou produto na parte superior, inferior e lateral da página. Um *site* de portal geralmente contém muitas imagens de *banners*, logotipos e propagandas. Estes dados geralmente consomem uma boa parte da banda da rede e, conseqüentemente, diminuem a eficiência da prestação da informação. É possível classificar os efeitos de cada objeto de mídia em uma página da *Web* e melhorar a eficiência da prestação de informações pela remoção de objetos redundantes ou priorizá-los de acordo com sua importância.

#### 4.1.1.1.2- Abstração de Informação

O objetivo é o de reduzir a largura de banda necessária para a entrega de conteúdo por meio da compressão dos dados, preservando a informação de maior valor para o usuário. Exemplos de extração de informações incluem sumarização de textos e geração de imagem ou vídeos em resoluções menores. Tais algoritmos também podem ser utilizados para melhorar a experiência do usuário em navegar através de uma visualização do conteúdo. Desta forma, os usuários são capazes de abstrair as informações mais importantes, mesmo que sua velocidade de transmissão seja limitada. A abstração de informações pode ser muito útil quando o dispositivo cliente tem capacidade limitada de exibição, como em *palmtops* e *smartphones*. Por exemplo, resumir cada parágrafo com poucas palavras e reduzir o tamanho/resolução de cada imagem ou vídeo em uma página da *Web* para adequá-la às pequenas telas dos aparelhos.

#### 4.1.1.1.3- Transformação de modalidade

É o processo de transformar o conteúdo de um modo para outro, para que este possa ser útil para um dispositivo cliente em particular. Por exemplo, alguns computadores portáteis não são capazes de lidar com dados de vídeo por causa de limitações de hardware e/ou de software. A fim de tornar a informação contida no vídeo acessível a esses dispositivos, podemos transformar o vídeo em conjuntos de imagens. Desta forma, os usuários serão capazes de receber informação útil, sob a forma que seus dispositivos possam suportar.

#### 4.1.1.1.4- Transcodificação de dados

É o processo de conversão de formato de dados de acordo com a capacidade do dispositivo cliente, mas a modalidade é mantida. Por exemplo, alguns dispositivos clientes podem não ser capazes de exibir as cores de imagens *Graphics Interchange Format* (GIF), devido à falta de software adequado ou à restrição de capacidade do hardware de exibição, como uma tela em preto-e-branco. Nesses casos, há uma necessidade de decodificar as imagens originais em outro formato apropriado, como GIF-para-JPEG – *Joint Photographic Experts Group*, de modo que elas possam ser vistas no dispositivo cliente. Aplicada a *streaming* de vídeo,

a transcodificação decodifica o fluxo de vídeo, processa-o e então o recodifica de acordo com as características da rede e dos clientes.

#### 4.1.1.1.5- Priorização de dados

O objetivo da priorização de dados é distinguir as partes mais importantes dos dados da parte menos importante, para que diferentes níveis de qualidade de serviço possam ser prestados na entrega dos dados através da rede. Por exemplo, podemos permitir que os dados menos importantes possam ser descartados em redes com limitações de largura de banda ou oferecer a entrega progressiva de qualidade ao enviar os dados mais importantes em primeiro lugar, como imagens de baixa resolução e, em seguida, fornecer os dados menos importantes para melhorar a informações posteriores, tais como a reconstrução de imagens em alta resolução. Desta forma, pode-se melhorar a experiência do usuário e utilizar eficientemente a largura de banda disponível. A priorização de dados pode ser alcançada dentro de um único tipo de mídia com a utilização de regimes especiais de codificação, como a codificação em camadas.

### 4.1.2 Segunda Classificação (Liu, 2003)

Outra classificação de adaptação de conteúdo, quando relacionada à adaptação de *streaming* de vídeo, segundo Liu(Liu, 2003), pode envolver codificação com controle de taxa, transcodificação e codificação escalável. Segue abaixo uma breve descrição de cada técnica.

#### 4.1.2.1.1- Controle de taxa

Consiste basicamente em alternar os diversos vídeos gravados em múltiplas taxas, que podem diferir em qualidade, resolução, taxa de quadros, dentre outros, de acordo com o contexto medido da rede.

#### 4.1.2.1.2- Transcodificação

Numa abordagem simples, a transcodificação decodifica o fluxo de vídeo num nó intermediário ou no receptor, e o recodifica de acordo com as características da rede e dos clientes. O processo de transcodificação exige capacidade de

processamento e pode acarretar em atraso na entrega do vídeo. O transcodificador converte um fluxo de vídeo, geralmente completo, em um fluxo novo com um diferente formato ou taxa.

#### 4.1.2.1.3- Codificação escalável

Consiste da extração de informação do vídeo original, com a inclusão ou diminuição de camadas auxiliares a partir de uma camada base. Geralmente, a qualidade de um vídeo depende da taxa de quadros, tamanho do quadro (resolução), e qualidade de quadro. Com o uso de codificação escalável é possível obter escalabilidade temporal, espacial e de qualidade ou a combinação destas escalabilidades. A escalabilidade de vídeo foi detalhada no Capítulo 2, incluindo detalhamento do codificador H.264 SVC, utilizado neste trabalho.

#### 2.1.1.8 - Terceira Classificação (Carvalho, 2004)

Considerando o tempo em que as diferentes alternativas de vídeos são criadas, pode-se classificar a adaptação de conteúdo em duas categorias (Carvalho, 2004):

#### 4.1.2.1.4- Adaptação Estática

Também conhecida como adaptação de conteúdo por seleção. Previamente, os vídeos são pré-codificados e armazenados em múltiplas versões que diferem em qualidade, resolução, requisitos de transmissão, dentre outros fatores. Durante a transmissão, a adaptação do conteúdo é feita alternando a transmissão de cada versão do vídeo.

### 4.1.3 Adaptação Dinâmica

Também conhecida como adaptação por transformação. Apenas uma cópia de cada vídeo é armazenada e adaptada durante a transmissão para as diversas necessidades, alternando a qualidade, taxa de quadros e resolução.

#### 4.1.4 Arquiteturas dos Adaptadores de Conteúdo

Segundo Carvalho(Carvalho, 2004), os mecanismos de adaptação de conteúdo podem ser executados em vários locais, sendo as mais utilizadas as arquiteturas baseadas no servidor, baseadas no cliente e baseadas em um componente intermediário (*Proxy*). Outras possibilidades são as arquiteturas em que o mecanismo de adaptação de conteúdo executa de forma distribuída entre o servidor, o cliente e o Proxy, ou ainda, em nós da rede (*hubs, switches e roteadores*).

Nos experimentos executados neste trabalho, o mecanismo de adaptação de conteúdo foi implementado na arquitetura baseada no servidor. O fluxo de informação desta arquitetura assume que a informação de contexto é obtida na requisição. O servidor analisa as informações da requisição do cliente, tais como, tipo de vídeo, resolução, dentre outros, e informações associadas, como largura de banda disponível, e então, ajusta e envia o conteúdo (vídeo) adaptado ao cliente.

## 4.2 ADAPTAÇÃO DE REDE

As técnicas de adaptação de rede (NAT) trabalham com a adaptação dos parâmetros de rede para as necessidades de uma aplicação multimídia em tempo real utilizando algoritmos que consideram o estado e/ou carga da rede (Gotz, 2004).

Segundo (Liu, 2003), os requisitos básicos das técnicas de adaptação da rede são: (1) fornecer as informações mais precisas possíveis referentes à carga da rede, (2) distinguir os erros e a gravidade do congestionamento da rede e (3) estimar e adequar a taxa de transmissão do fluxo de vídeo à largura de banda disponível. A diferenciação qualitativa de erros e congestionamento é necessária para reduzir a taxa de transmissão apenas quando se justifique (por exemplo, quando realmente há um congestionamento).

A adaptação de rede pode ser dividida em três fases: sensoriamento, processamento e atuação. Segue abaixo a descrição destas fases e suas aplicações pelos esquemas adaptativos.

#### **4.2.1 Sensoriamento**

Na fase de sensoriamento, os esquemas adaptativos capturam informações dos componentes da rede que indicam as condições desta. Estas informações podem ser coletadas no cliente, tais como o *jitter*, a taxa de pacotes perdidos, o percentual de utilização do *buffer*, dentre outros. Informações podem ser coletadas, também, nos componentes da rede, tal como o ECN.

#### **4.2.2 Processamento**

A fase de processamento é responsável por receber e processar as informações coletadas na fase de sensoriamento e diagnosticar a situação atual da rede. A fase de processamento pode ser executada no servidor, em um componente intermediário ou no cliente.

#### **4.2.3 Atuação**

A fase de atuação é responsável por ajustar os recursos de rede às condições medidas a partir do diagnóstico da rede mensurado na fase de processamento. Para transmissão de *streaming* de vídeo, pode-se citar, não exaustivamente, a atuação na priorização da classe de pacotes, no caso de redes que implementam QoS, por exemplo o DiffServ (Nichols,1998) (Blake,1998) (Black, 1998) (QoS Forum, 2008), ou pela atuação na largura de banda disponível para transmissão do vídeo.

## CONCLUSÃO

A transmissão de vídeo é uma aplicação cada vez mais requisitada na *Internet*. Mas a transmissão de vídeo em tempo real com níveis aceitáveis de qualidade gera uma série de desafios. No caso de redes baseadas no melhor esforço, como a *Internet*, as transmissões podem apresentar grandes variações nas suas taxas, o que dificulta muito o gerenciamento da transmissão.

Nestas circunstâncias adversas, as aplicações de transmissão de vídeo precisam se adaptar às mudanças das condições na rede. Assim, precisam executar técnicas escaláveis e adaptativas em termos de conteúdo e taxas de transmissão para poder lidar com as variações e heterogeneidades da rede.

Numa rede que provê QoS, as condições de rede são ajustadas para suportar o tráfego, seja por reserva de recursos, seja por priorização de tráfego.

Em especial, o uso de escalabilidade de vídeo, como forma de prover a adaptabilidade de conteúdo, pode ser utilizado em conjunto com a adaptabilidade de rede, permitindo o tratamento de demandas que podem variar em decorrência, por exemplo, da heterogeneidade dos terminais clientes. Neste sentido, a escalabilidade de vídeo permite o envio de camadas complementares de vídeo, que são transmitidas de forma a melhor aproveitar a largura de banda disponível e que são utilizadas pelos usuários conforme sua capacidade de processamento e recepção.

Uma das questões mais críticas na transmissão de vídeo é assegurar a QoE em um nível aceitável. Neste trabalho, foram estudadas métricas objetivas de QoE que permitem avaliar a qualidade do vídeo entregue ao usuário e, conseqüentemente, da qualidade percebida por ele. As métricas utilizadas foram a PSNR, bastante conhecidas no meio acadêmico, e o VQM no seu módulo geral.

Neste contexto, este trabalho apresentou alguns protocolos de transmissão, definiu e descreveu a QoS em redes, bem como o QoE em qualidade de vídeo. Permitiu o estudo e apresentação de técnicas de adaptação de rede e de conteúdo, além de um histórico dos padrões de codificação de vídeo e análise do uso da escalabilidade de vídeo para permitir a maior qualidade nas transmissões de vídeo na internet.

## REFERÊNCIAS

HEAVYREADING, **ASSURING QUALITY OF EXPERIENCE FOR IPTV**, *HeavyReading, Whitepaper*, 2006. Disponível em: <http://www.heavyreading.com/> Acessado em Junho 2009.

ROBINS, M., **Delivering optimal Quality of Experience (QoE) for IPTV success**, Spirent Communications, Whitepaper, 2006. Disponível em: [www.spirentcom.com/documents/4002.pdf](http://www.spirentcom.com/documents/4002.pdf). Acessado em Junho de 2009.

SCHAWARZ, H., MARPE D., e WIEGAND T., **Overview of the Scalable Video Coding Extension of the H.264/AVC Standard**, IEEE Transactions Circuits and Systems for Video Technology, Vol. 17, No. 9, Pg 1103-1120, 2007.

JAIN, R., **Quality of Experience**, IEEE Multimedia, Volume 11, Issue 1, pp. 96-97, 2004.

KISHIGAMI, J., **The Role of QoE on IPTV Services**, Ninth IEEE International Symposium on Multimedia, Pg.11-13, 2007.

ITU-T *Recommendation G. 1010, End-User Multimedia QoS Categories*, 2001.

SILVA, F., BADRUDINO, N. e SOUSA, T., **IPTV – Arquiteturas, Protocolos e Serviços**, Instituto Superior Técnico - Taguspark – Portugal, dissertação de mestrado, 2007.

PATRICK, A. S., et al., **A QoE Sensitive Architecture for Advanced Collaborative Environments**, in Proceedings of the First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks, Pg. 319-322, 2004.

O' NEIL, T. M., **Quality of Experience and Quality of Service for IP Video Conferencing**, Polycom, Whitepaper, 2002.

GOODCHILD, J., **Integrating Data, Voice and Video – Part II**, IP Video Implementation and Planning Guide, United States Telecom Association, 2005.

PIAMRAT, K., VIHO, C., KSENTINI, A. e BONNIN, J.; **Quality of Experience Measurements for Video Streaming over Wireless Networks**, 2009 Sixth International Conference on Information Technology: New Generations, 2009.

LOPEZ, D., GONZALEZ, F., BELLIDO, L. and ALONSO, A. **Adaptive Multimedia Streaming Over IP Based on Customer Oriented Metrics**, in Proceedings of the Seventh IEEE International Symposium on Computer Networks, pp. 185-191, 2006.

ANSELMO, T., ALFONSO, D., **Buffer-Based Constant Bit-rate Control for Scalable Video Coding**, Whitepaper. Advanced System Technology, STMicroelectronics, Agrate Brianza, *Italy*, 2008.

ITU-T *Recommendation H.264, Advanced Video Coding for Generic Audiovisual Services*, 2007.

MANOEL, E. T. M., **Codificação de Vídeo H.264 – Estudo de Codificação Mista de Macroblocos** Dissertação de Mestrado, Universidade Federal de Santa Catarina, 2007.

DA SILVA, A. M. C., **Um Estudo Sobre o Padrão H.264/AVC de Compressão de Vídeo**, Dissertação de Mestrado, Universidade Católica de Pelotas, Pelotas RS, 2007.

ITU-T Recommendation J.149, **Method for Specifying Accuracy and Cross-Calibration of Video Quality Metrics (VQM)**, 2004.

ITU-T *Recommendation H.264 | ISO/IEC 14496-10:2005 (MPEG-4), Advanced Video Coding for Generic Audiovisual Services*, Int'l Telecomm. Union | Int'l Organization for Standardization, Geneva, Switzerland, 2005.

MAYER, A. e LINDER, H., **A Survey of Adaptive Layered Video Multicast using MPEG-2 Streams** Department of Scientific Computing, University of Salzburg, Austria, 2005.

ITU-T *Recommendation J.143, User Requirements for Objective Perceptual Video Quality Measurements in Digital Cable Television*, 2000.

WINKLER, S; **Digital Video Quality – Vision Models and Metrics**, John Wiley & Sons, 2005.

ANSI T1.801.03: **Digital Transport of One-Way Video Signals – Parameters for Objective Performance Assessment**, 2003.

WINKLER, S. e MOHANDAS, P; **The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics**, IEEE Trans. Broadcasting, Volume 54, N. 3, 2008.

KODIKARA, A. H., HEWAGE, C.T.E.R., DOGAN, S., MRAK, M., BARBOSA, V., M.T. Andrade e A.M. Kondoç, **Context-Aware Adaptation of SVC Scalability Structure for Improved Coding Efficiency**, 50<sup>th</sup> International Symposium ELMAR-2008, Pg. 443-446, Zadar, Croatia, 2008.

VISHWANATH, A; DUTTA, P; CHELTLUR, M; GUPTAB, P; KALYANARAMANB, S. e GHOSH, A; **Perspectives on Quality of Experience for Video Streaming over WiMAX**, Mobile Computing and Communications Review, Volume 13, N. 4, 2008.

ZAPATER, M. N., BRESSAN, G., **A Proposed Approach for Quality of Experience Assurance for IPTV**, First International Conference on the Digital Society, IEEE, 2007.

WENGER, S., WANG, Y. e HANNUKSELA, M. M., **RTP Payload Format for H.264/SVC Scalable Video Coding**, Journal of Zhejiang University, n. 7, pg. 657-667 China, 2006.

ITU-T *Recommendation E.800, QoS Terms and Definitions Related to Quality of Service and Network Performance Including Dependability*, 1994.

SILLER, M. and WOODS, J., **Improving Quality Experience for Multimedia Services by QoS Arbitration on a QoE Framework**, in Proceedings of the 13th Packed Video Workshop 2003, Nantes, France, 2003.

BAUER, B. and PATRICK, A. S., **A Human Factors Extension to the Seven-Layer OSI Reference Model**. [Online]. Disponível em: <http://www.andrewpatrick.ca/OSI/10layer.html>. acessado em Fevereiro de 2009.

SCHWARZ, H., MARP, D. and WIEGAND, T., **Overview of the Scalable H.264/MPEG4-AVC Extension**, Fraunhofer Institute for Telecommunications, Berlin, Germany, 2006.

REICHEL, J., SCHWARZ, H. and WIEN, M. (eds.), **Joint Scalable Video Model JSVM-6**, Joint Video Team, Doc. JVT-S202 Geneva, Switzerland, 2006.

DSL Forum, **Triple-play Services Quality of Experience Requirements and Mechanism**, Working Text WT-126 version 0.5, 2006.

SOLER, L., **Compactação de Vídeo Escalável**, Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, 2006.

CHIARIGLIONE, L., **Terms of Reference for a Joint Project between ITU-T Q.6/SG16 and ISO/IEC JTC1/SC29/WG11 for the Development of new Video Coding, Recommendation and International Standard ISO/IEC JTC1/SC29/WG11 N 4400**, Pattaya, 2001.

CONKLIN, G. J. e HEMAMI, S.S., **A Comparison of Temporal Scalability Techniques**, IEEE Transactions on Circuits and Systems for Video Technology, v.9, n.6, p.909–919, 2000.

NICHOLS, K., BLAKE, S., BAKER, F., BLACK, D., **Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers**, RFC 2474, 1998.

BLAKE, S. BLACK, D., CARLSON, M., DAVIES, E., WANG, Z., WEISS, W., **An Architecture for Differentiated Services**, RFC 2475, 1998.

KUIPERS, F, KOOIJ, R; De VLEESCHAUWER, D. e BRUNNSTRÖM, K., **Techniques for Measuring Quality of Experience**, Proc. of the 8th International Conference on Wired/Wireless Internet Communications (WWIC 2010), 2010.

WINKEN, M., SCHWARZ, H., MARPE, D., e WIEGAND, T., **Adaptive motion refinement for FGS slices**, Joint Video Team, Doc. JVT-Q031, Nice, France, 2005.

Joint Video Team, **2005a. Scalable Video Coding - Working Draft 1**, 2005.

Joint Video Team, **2005b. Scalable Video Coding - Working Draft 4**, 2005.

Joint Video Team, **2005c. Scalable Video Coding - Working Draft 4**, 2005.

RODRIGUES, D., CERQUEIRA, E., MONTEIRO, E., "**Quality of Service and Quality of Experience in Video Streaming**", in Proc. of the International Workshop on Traffic Management and Traffic Engineering for the Future Internet (FITraMEn 08), 2008.

CONEY, E., MUNTEAN, G., **A Priority-Based Adaptive Scheme for Wireless Multimedia Delivery**, IEEE International Conference on Multimedia and Expo, 2006.

ZHU, H., LI M., CHLAMTAC I. e PRABHAKARAN B., **A Survey of Quality of Service in IEEE 802.11 Networks**, IEEE Wireless Communications, 2004.

IEEE 802.11 WG, **Draft Supplement to Telecommunications and Information Exchange Between Systems LAN/MAN Specific Requirements Part 11: Wireless MAC and PHY Specification: MAC Enhancements for QoS**, IEEE 802.11e/draft 11.0, 2004

BLACK, D.; BLAKE, S.; CARLSON, M.; DAVIES, E.; WANG, Z. e WEISS, W., **An Architecture for Differentiated Services**, Request for Comments 2475 <ftp://ftp.nic.it/rfc/rfc2475.txt>. 1998.

BRADEN, R.; CLARK, D. e SHENKER, S., **Integrated Services in the Internet Architecture: an Overview**, Request for Comments 1633. <http://www.ietf.org/rfc/rfc1633.txt>. 1994.

IEFT, **Integrated Services**. <http://www.ietf.org/html.charters/intserv-charter.html>., acessado em Fevereiro de 2009.

QoS Forum, **QoS Protocols & Architectures**, Disponível em [http://www.qosforum.com/white-papers/qosprot\\_v3.pdf](http://www.qosforum.com/white-papers/qosprot_v3.pdf). Acesso em: Novembro, 2008.

GROSS. J; KLAUE, J; KARL, H. e WOLISZ A., **Cross-Layer Optimization of OFDM Transmission Systems for MPEG-4 Video Streaming**, Computer Communications, vol. 27, no. 11, pp. 1044–1055, 2004.

VAN DER AUWERA G., PRASANTH T. D., REISSLEIN M. e KARAM L. J., **Traffic and Quality Characterization of the H.264/AVC Scalable Video Coding Extension**, Hindawi Publishing Corporation, Advances in Multimedia, Article ID 164027, 2008.

WEI Y., ZHENG G. L. e RAHARDJA S., **Scalable Video Coding in a Nutshell**, Synthesis Journal, Institute for Infocomm Research, 2008.

WIEN M., SCHWARZ H. e OELBAUN T., **Performance Analysis of SVC**, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, Pg. 1194-1203, 2007.

KUROSE J. F., ROSS K. W., **Computer Networking – A Top Down Approach**, Pearson Addison Wesley, 4 Edição, 2008.

- ZHANG H., ***Adaptative Content Delivery: A New Application Area for Media Computing Research***, International Workshop on Multimedia Data Storage, Retrieval, Integration and Application, Hong Kong, 2000.
- Carvalho P. M. M., ***Multimedia Content Adaptation For Universal Access***, Dissertação de Mestrado, Faculdade de Engenharia da Universidade do Porto, Portugal, 2004.
- RAMZAN Q., ZHANG Q., CHANDRAMOULI K., BORGES P. V. K., IZQUIERDO E., HANJALIC A., KROON B., CLEMENTS M., SCHMIEDEKE S., KELM P., NEUBAUER N., ZAPPE M., DE SIMONE F., USCUMLIC M., OUARET M., LEE J. e NGUYEN D. T., ***State-of-the-art Report***, PetaMedia, 2008.
- CHANG S., VETRO A., ***Video Adaptation: Concepts, Technologies and Open Issues***, Proc IEEE, vol. 93, no. 1, pp. 148-158, 2005.
- GOTZ D., MAYER-PATEL K., ***A General Framework for Multidimensional Adaptation***, ACM Multimedia, Pg. 612 – 619, 2004.
- LIU J., Li B. ZHANG Y., ***Survey: Adaptive Video Multicast Over the Internet***, IEEE Computer Society, Pg. 22-33, 2003.
- Soares, Luiz Fernando G., ***Rede de Computadores***, 6ª Edição, Editora Campus, 1995.
- BOTTREAU V., ***Scalable Video Coding – Scalable Extension of H.264 / AVC, Technology CR Rennes – Content Delivery & Communication Lab, Compression Group***, Thomson, Rennes, França, 2007.
- RFC 3550, ***RTP: A Transport Protocol for Real-Time Applications***, 2003.
- WOLF S. e PINSON, M., ***Fast Low Bandwidth Video Quality Model (VQM) Description and Reference Code***, ITU-T Grupo 9, 2009.
- WANG, Y., ***Survey of Objective Video Quality Measurements***, Tech report, Worcester Polytechnic Institute, EMC Corporation Hopkinton, Estados Unidos, 2006.
- MARTÍNEZ, J. L., CUENCA, P., DELICADO, F. e QUILES, F., ***Objective Video Quality Metrics: A Performance Analysis***, Universidad de Castilla La Mancha, Albacete, Espanha, 2007.
- RFC 768, ***UDP : User Datagram Protocol***, 1980.