



**AVALIAÇÃO DA QUALIDADE DE VÍDEO ESCALÁVEL
EM REDES SEM FIO DE 3ª GERAÇÃO**

VANICE CANUTO CUNHA

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

AVALIAÇÃO DA QUALIDADE DE VÍDEO ESCALÁVEL EM
REDES SEM FIO DE 3ª GERAÇÃO

VANICE CANUTO CUNHA

Orientador: Dr. PAULO ROBERTO DE LIRA GONDIM

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA
ELÉTRICA

PUBLICAÇÃO PPGENE.DM – 470/12
BRASÍLIA/DF, 13 DE FEVEREIRO DE 2012.

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**AVALIAÇÃO DA QUALIDADE DE VÍDEO ESCALÁVEL EM
REDES SEM FIO DE 3ª GERAÇÃO**

VANICE CANUTO CUNHA

ORIENTADOR: Dr. PAULO ROBERTO DE LIRA GONDIM

**DISSERTAÇÃO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA ELÉTRICA DA UNIVERSIDADE DE BRASÍLIA
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA
ELÉTRICA.**

APROVADA POR:

**Paulo Roberto de Lira Gondim, Dr., ENE/UNB
(Orientador)**

Díbio Leandro Borges, Dr., CIC/UNB 1

Joel José Puga Coelho Rodrigues, Dr., UBI/PT

BRASILIA/DF, 13 DE FEVEREIRO DE 2012.

FICHA CATALOGRÁFICA

CUNHA, VANICE CANUTO

**Avaliação da Qualidade de Vídeo Escalável em Redes Sem Fio de 3ª Geração,
2012xv, 119 p. 201x297 mm**

(UnB, Mestre, Engenharia Elétrica, 2012

Dissertação de Mestrado – Universidade de Brasília.

Faculdade de Tecnologia

Departamento de Engenharia Elétrica

1. Qualidade de Vídeo

2. *Streaming* de vídeo

3. Redes de 3ª Geração

4. Qualidade de serviço

REFERÊNCIA BIBLIOGRÁFICA

CUNHA, V. C. (2012). Avaliação da Qualidade de Vídeo Escalável em Rede Sem Fio de 3ª Geração. Dissertação de Mestrado em Engenharia Elétrica. Publicação 470/12, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 119 p.

CESSÃO DE DIREITOS

AUTOR: Vanice Canuto Cunha

TÍTULO: “Avaliação da Qualidade de Vídeo Escalável em Redes Sem Fio de 3ª Geração”

GRAU/ANO: Mestre/2012.

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de Mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor se reserva a outros direitos de publicação e nenhuma parte desta dissertação de Mestrado pode ser reproduzida sem a autorização por escrito do autor.

Vanice Canuto Cunha

UNIVERSIDADE DE BRASÍLIA

Faculdade de Tecnologia.

Departamento de Engenharia Elétrica.

70910-900 – Brasília – DF – Brasil.

AGRADECIMENTOS

A Deus e ao Nosso Senhor Jesus Cristo, que me deu forças e capacidade para que eu chegasse ao final dessa jornada com sucesso.

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.

Ao meu orientador e amigo Paulo Roberto de Lira Gondim, pela oportunidade, confiança, apoio e tempo de orientação dedicada exclusivamente para mim.

As minhas amigas Rafa, Kary, Mary, Juliana, Jaque e Vê e ao meu amigo Jadir pelo apoio que me deram durante esta caminhada, tentando me distrair e me motivando a cada dia.

À minha amiga Lilissanne, que em pouco tempo se tornou uma grande amiga e por várias e várias vezes abriu mão de suas tarefas para estar comigo, me apoiando e me dando forças, a você minha amiga, meu obrigada mais que especial.

Aos meus colegas de laboratório, Helder, Lilissanne, Manoel Campos, Maurício Júnior, Leandro Vaguetti, Rodrigo Mulinari, Valmiro, Vinícius Rios, e em especial ao Cláudio, que se tornou um grande amigo.

Aos meus companheiros de Labredes, Wandemberg, Adriana, Kelly e Diego, além de me apoiarem em todas as horas, me ajudaram muito.

Por último e não menos importante, às outras pessoas, que de certa forma me apoiaram direta ou indiretamente.

Dedico a

Deus, meus pais, irmãos, cunhados, sobrinhos e amigos pelo apoio, compreensão e suporte.

RESUMO

Avaliação da Qualidade de Vídeo Escalável em Redes Sem Fio de 3ª Geração

Autor: Vanice Canuto Cunha

Orientador: Paulo Roberto de Lira Gondim

Brasília, Fevereiro de 2012.

Nos últimos anos a transmissão e a entrega de vídeo tem se caracterizado como uma das mais importantes aplicações das redes de comunicação, e garantir uma boa qualidade do vídeo recebido pelo usuário final tem sido um grande desafio para a maioria das operadoras. Para esse fim, a utilização de escalabilidade de vídeo tem se apresentado como uma importante opção para *streaming* tanto em redes cabeadas quanto as baseadas em tecnologias *wireless*

Esta dissertação apresenta uma análise comparativa entre a qualidade do vídeo codificado nos padrões H.264/SVC (*Scalable Video Coding*) e MPEG-4 (*Moving Picture Experts Group*) parte 2, transmitidos em redes de 3ª geração. Considera-se, para o uso do padrão H.264/SVC, a utilização de esquema adaptativo, que usa lógica *fuzzy* (Nebulosa) e *feedbacks* do receptor, permitindo realizar o ajuste do *bitrate* associado ao padrão H.264/SVC.

São realizadas comparações, com base em métricas objetivas de avaliação de qualidade de vídeo, destinadas a estimar a qualidade das sequências de vídeo recebidas pelo usuário, permitindo inferir sobre a adequação da extensão SVC e do esquema adaptativo usado.

Além do esquema adaptativo utilizado para transmissão do *streaming*, a dissertação apresenta também um conjunto de ferramentas para codificação e decodificação de *streaming*, bem como ferramentas para aferição da qualidade do vídeo.

Palavras-chave: Escalabilidade de Vídeo, Qualidade de Vídeo, Redes Sem Fio de 3ª Geração, *Streaming*, *PSNR*.

ABSTRACT

Avaliação da Qualidade de Vídeo Escalável em Redes Sem Fio de 3ª Geração

Autor: Vanice Canuto Cunha

Orientador: Paulo Roberto de Lira Gondim

Brasília, Fevereiro de 2012

In recent years, video transmission and delivery has been characterized as one of the most important applications of communication networks and ensure a good quality of video received by the end user has been a major challenge for most operators. To this end, the use of video scalability has emerged as an important option for streaming in both based on the wired and wireless technologies

This dissertation presents a comparative analysis of the quality of encoded video standards H.264/SVC (Scalable Video Coding) and MPEG-4 (Moving Picture Experts Group) Part 2, transmitted on 3rd generation networks. It is considered, for the use of standard H.264/SVC, the use of an adaptive scheme, which uses fuzzy logic and feedbacks from the receiver, permitting adjustment of the standard bitrate associated H.264/SVC.

Comparisons are based on objective metrics to evaluate video quality, to estimate the quality of video sequences received by the user, allowing inferences about the adequacy of the SVC extension and adaptive scheme used.

In addition to the adaptive scheme used for transmission of streaming, the dissertation also presents a set of tools for encoding and decoding of streaming, as well as tools to assess the quality of the video.

Keywords: *Scalable Video Coding, Quality of Video, 3rd Generation Networks Streaming, PSNR.*

SUMÁRIO

RESUMO	VII
ABSTRACT	VIII
1 INTRODUÇÃO.....	1
1.1 Motivação e Justificativa	1
1.2 Objetivos.....	2
1.2.1 Objetivo geral	2
1.2.2 Objetivos específicos.....	2
1.3 Organização	3
2 CODIFICAÇÃO E ESCALABILIDADE DE VÍDEO.....	4
2.1 Introdução	4
2.1.1 Conceitos básicos	4
2.1.2 Padrão H.264/AVC.....	7
2.1.3 Padrão H.264/SVC	10
2.2 Tipos de Escalabilidade	14
2.2.1 Escalabilidade Temporal	14
2.2.2 Escalabilidade Espacial	16
2.2.3 Escalabilidade SNR	18
2.2.4 Escalabilidade Combinada	19
2.3 Níveis de Granularidade	20
2.3.1 Escalabilidade de granularidade grossa.....	20
2.3.2 Escalabilidade de granularidade média	20
2.3.3 Escalabilidade de granularidade fina.....	21
3 EVOLUÇÃO DE REDES SEM FIO E REDES DE 3ª	
GERAÇÃO.....	22
3.1 Introdução	22
3.2 Padronização.....	22
3.3 Evolução de redes sem fio	22
3.3.1 HSCSD	23
3.3.2 GPRS	23
3.3.3 EDGE	25
3.3.4 UMTS	25
3.3.5 WCDMA	27
3.3.6 HSDPA	28
3.4 Aspectos do Padrão 3GPP Relativos a <i>Streaming</i> de Vídeo	30
3.4.1 Introdução.....	30
3.4.2 Registros e codificadores.....	31
3.4.3 Perfil de servidor de <i>streaming</i>	31
3.4.4 Perfil de <i>streaming</i> adaptativo.....	31
3.4.5 Informações de <i>buffer</i> de vídeo	32
3.4.6 Protocolos para <i>streaming</i> no 3GPP.....	32
3.4.7 Recursos de <i>streaming</i> adaptativo dentro do PSS	34
3.4.8 Sessão de <i>streaming</i> adaptativo.....	35

4	PROTOCOLOS E ADAPTAÇÃO PARA <i>STREAMING</i> DE VÍDEO	36
4.1	Protocolos Utilizados Neste Trabalho	36
4.1.1	Protocolo de transporte UDP	36
4.1.2	Protocolo RTP	37
4.1.3	Protocolo RTCP.....	38
4.2	Esquema adaptativo para <i>streaming</i> de vídeo SVC.....	39
4.3	Componentes do esquema adaptativo.....	41
4.3.1	Componentes do Servidor	41
4.3.2	Componentes do Cliente.....	41
4.3.3	Algoritmo de decisão.....	42
5	ASPECTOS RELACIONADOS A QOS, QOV E QOE	43
5.1	Métricas de qualidade de vídeo	43
5.1.1	PSNR	43
5.1.2	VQM.....	44
5.2	Métricas de qualidade de serviço.....	45
5.2.1	Atraso	45
5.2.2	<i>Jitter</i>	46
5.2.3	Perda de pacotes	46
5.3	Qualidade de Experiência.....	46
5.4	Ferramentas para aferição da qualidade de vídeo e redes e codificação de vídeo 47	
5.4.1	JSVM.....	47
5.4.2	BVQM	49
5.4.3	FFMPEG.....	49
5.4.4	VLC	49
5.5	Formatos de vídeos.....	49
5.5.1	Sequências de vídeo.....	50
5.5.2	Classificação das Sequências.....	51
6	TRABALHOS RELACIONADOS E DESCRIÇÃO DOS EXPERIMENTOS	53
6.1	Trabalhos Relacionados.....	53
6.2	Descrição dos Experimentos	55
6.2.1	Cenário proposto	55
6.2.2	Arquitetura.....	57
7	ANÁLISE DOS RESULTADOS	59
7.1	Introdução.....	59
7.2	Resultados com base em Y-PSNR	59
7.3	Resultados com base em VQMg	68
8	CONCLUSÕES E TRABALHOS FUTUROS.....	71
8.1	Conclusões.....	71
8.2	Trabalhos Futuros	72
	REFERÊNCIAS	73
	APÊNDICE I	79

I.1 – Artigo publicado no evento COMTEL 2011	79
APÊNDICE II.....	91
II.1– Arquivos de configuração dos vídeos - cfg	91
APÊNDICE III	101
III.1 - Arquivos de configuração do Apache	101

Siglas

3GPP	<i>Third Generation Partnership Project</i>
AuC	<i>Authentication Center</i>
AVC	<i>Advanced Video Coding</i>
BMF	<i>Base Mode Flag</i>
BVQM	<i>Batch Video Quality Metric</i>
CABAC	<i>Context -Adaptive Binary Arithmetic Coding</i>
CAVLC	<i>Context-Adaptive Variable-Length Coding</i>
CGS	<i>Coarse-Grain Quality Scalable Coding</i>
CIF	<i>Common Intermediate Format</i>
CN	<i>Core Network</i>
CS	<i>Circuit-Switched</i>
EDGE	<i>Enhanced Data Rates for Digital Evolution</i>
EIR	<i>Equipment identity Register</i>
FDD	<i>Frequency Division Duplex</i>
FGS	<i>Fine Grain Quality Scalable)</i>
GGSN	<i>Gateway GPRS Support Node</i>
GMSC	<i>Gateway Mobile Switching Center</i>
GOP	<i>Group of Pictures</i>
GPRS	<i>General Packet Radio Service</i>
GSM	<i>Global System for Mobile Communications</i>
HARQ	<i>Hybrid Automatic Repeat Request</i>
HSCSD	<i>High Speed Circuit Switched Data</i>
HSDPA	<i>High Speed Downlink Packet Access</i>
HSS	<i>Home Subscriber Server</i>
IMT	<i>Internacional Mobile Telephony</i>
ISDN	<i>Integrated Services Digital Network</i>
ITU-T	<i>International Telecommunication Union</i>
JVT	<i>Joint Vídeo Team</i>
MBs	<i>Macroblocks</i>
MCP	<i>Motion-Compensated Prediction</i>
MGS	<i>Medium-Grain Quality Scalability</i>
MMS	<i>Multimedia Messaging Service</i>
MPEG	<i>Moving Picture Experts Group</i>
MPF	<i>Motion Prediction Flag</i>
MSC	<i>Mobile Switching Center</i>
NADU APP	<i>Next Application Data Unit Application Packet</i>
NAL	<i>Network Abstraction Layer</i>
PDP	<i>Protocol Data Packet</i>
PR	<i>Progressive Refinement</i>
PS	<i>Packet-Switched</i>
PSNR	<i>Peak Signal-to-Noise Ratio</i>
PSS	<i>Packet Switched Streaming Service</i>
PSTN	<i>Public Switched Telephone Network</i>
QCIF	<i>Quarter CIF</i>
QoE	<i>Qualidade de Experiência</i>
QoS	<i>Quality of Service</i>
QoV	<i>Quality of Vídeo</i>
RA	<i>Routing Área</i>

RNC	<i>Radio Network Controller</i>
RNS	<i>Radio Network System</i>
RPF	<i>Residual Prediction Flag</i>
RTCP	<i>Real Time Control Protocol</i>
RTP	<i>Real Time Protocol</i>
RTT	<i>Round Trip Times</i>
SDP	<i>Session Description Protocol</i>
SGSN	<i>Serving GPRS Support Node</i>
SMS	<i>Short Message Service</i>
SNR	<i>Signal-to-Noise Ratio</i>
SVC	<i>Scalable Video Coding</i>
TDD	<i>Time Division Duplex</i>
TDMA	<i>Time Division Multiple Access</i>
UDP	<i>User Datagram Protocol</i>
UMTS	<i>Universal Mobile Telecommunication System</i>
UTRAN	<i>Universal Mobile Telecommunication System Terrestrial Radio Access Network</i>
VCEG	<i>Video Coding Experts Group)</i>
VCL	<i>Video Coding Layer</i>
VLR	<i>Visitor Location Register</i>
VQEG	<i>Video Quality International Expert Group</i>
VQM	<i>Video Quality Metric</i>
VQMg	<i>Video Quality Metric General</i>
WCDMA	<i>Wideband Code Division Multiple Access</i>

FIGURAS

Figura 2.1: GOP formado por oito quadros.....	5
Figura 2.2 : Quadro com macroblocos destacados.....	6
Figura 2.3: Quadro formado por oito slices.....	7
Figura 2.4: Estrutura do codificador H.264/AVC [Adaptada de [Wiegand <i>et al</i> 2003]...]	8
Figura 2.5: Estrutura em camadas de um fluxo H.264/SVC	11
Figura 2.6: Processo de decisão do decodificador para decodificação do macrobloco das camadas de reforço (Adaptada de [Narvekar <i>et al</i> 2009]).....	12
Figura 2.7: Diferentes camadas temporais.	15
Figura 2.8: Escalabilidade temporal, hierarquia baseada no frame B.[Xiaoqing Zhua <i>et al.</i> 2008].....	15
Figura 2.9: Diferentes camadas de resolução.	17
Figura 2.10: Representação de escalabilidade espacial com duas camadas de reforço [Kang <i>et al.</i> 2009].....	18
Figura 2.11: Escalabilidade de qualidade	19
Figura 3.1: Arquitetura da rede GPRS [Janevski 2003].....	24
Figura 3.2: Arquitetura da rede UMTS [Janevski 2003].....	26
Figura 3.3: Arquitetura básica do sistema WCDMA.	28
Figura 3.5: Arquitetura da HSDPA [Schierl <i>et al</i> 2007].	29
Figura 3.6: Pilha de protocolos utilizada para <i>streaming</i> no 3GPP [Frojdh <i>et al</i> 2006] [Hartung <i>et al</i> 2007].	33
Figura 4.1: Cabeçalho UDP [RFC 768 1980].	36
Figura 4.2: Cabeçalho RTP.	37
Figura 4.3: Cabeçalho do RTCP [RFC 3550 2003].	38
Figura 4.4: Arquitetura do esquema adaptativo.....	40
Figura 5.1: Quadros das sequências de vídeos utilizadas.....	50
Figura 6.1: Cenário proposto para os experimentos.....	55
Figura 6.2: Fluxograma da arquitetura de avaliação	57
Figura 7.1: Média Y-PSNR da sequência de vídeo <i>Foreman</i>	60
Figura 7.2: Média Y-PSNR da sequência de vídeo <i>Container</i>	60
Figura 7.3: Média Y-PSNR da sequência de vídeo <i>Football</i>	60
Figura 7.4: Média Y-PSNR da sequência de vídeo <i>Waterfall</i>	61
Figura 7.5: Média Y-PSNR da sequência de vídeo <i>Paris</i>	61

Figura 7.6: Média Y-PSNR da sequência de vídeo <i>Flower</i>	61
Figura 7.7: Média Y-PSNR da sequência de vídeo <i>Highway</i>	62
Figura 7.8: Média e desvio padrão <i>Foreman</i>	65
Figura 7.9: Média e desvio padrão <i>Container</i>	65
Figura 7.10: Média e desvio padrão <i>Football</i>	66
Figura 7.11: Média e desvio padrão <i>Waterfall</i>	66
Figura 7.12: Média e desvio padrão <i>Paris</i>	67
Figura 7.13: Média e desvio padrão <i>Flower</i>	67
Figura 7.14: Média e desvio padrão <i>Highway</i>	68
Figura 7.15: Resultados média e desvio padrão VQMg com o uso do esquema.	69
Figura 7.16: Resultados média e desvio padrão VQMg para os vídeos MPEG-4.	69
Figura 7.17: Resultados média e desvio padrão VQMg dos vídeos SVC sem o esquema.	70

TABELAS

Tabela 5.1: Parâmetros de configuração do vídeo.....	51
Tabela 6.1: Relação entre PSNR, MOS e grau de deficiência.	56
Tabela 7.1: Mapeamento entre PSNR e MOS.....	64

1 Introdução

1.1 Motivação e Justificativa

Nos últimos anos a transmissão e a entrega de vídeo tem se caracterizado como uma das mais importantes aplicações das redes de comunicação, e garantir uma boa qualidade do vídeo recebido pelo usuário final tem sido um grande desafio para a maioria das operadoras. Para esse fim, a utilização de escalabilidade de vídeo tem se apresentado como uma importante opção para *streaming* tanto em redes cabeadas quanto as baseadas em tecnologias *wireless*, permitindo que haja uma melhor compatibilização entre a disponibilidade de banda e as demandas que envolvem requisitos de tempo real e de elevada banda passante, e levando à necessidade de adoção de esquemas adaptativos.

Dentre as redes sem fio, as de 3ª Geração representam uma evolução dos já bastante conhecidos e explorados sistemas de 2ª geração e de geração 2,5G, com transmissão baseada em pacotes, melhor oferta de banda e de forma compatível com outros sistemas digitais, favorecendo o acesso a uma quantidade crescente de informações disponíveis on-line, frequentemente utilizando a *Web*, e permitindo às pessoas utilizarem seus dispositivos móveis para serviços de dados, vídeo e até TV digital.

Em geral, redes sem fio podem estar sujeitas a altas variações de tráfego da rede, bem como, oscilações, congestionamentos repentinos e heterogeneidade dos dispositivos, com isso as aplicações de *streaming* precisam se adaptar às mudanças das condições da rede e às diferentes demandas nelas geradas.

Dessa forma, a transmissão de vídeo em tempo real em redes sem fio gera uma série de novos desafios, que incluem garantir qualidade de serviços para os diferentes receptores, utilizar ao máximo a capacidade de cada receptor, prover o melhor aproveitamento de banda permanecendo em níveis aceitáveis de qualidade, evitar ou reagir rapidamente a congestionamentos na rede e atender a variações repentinas do número de usuários ativos.

Nestas circunstâncias, as aplicações de transmissão de vídeo precisam se adaptar às mudanças das condições da rede e às diferentes demandas. Assim, elas podem adotar técnicas escaláveis e adaptativas tanto em termos de conteúdo quanto em termos de taxas de transmissão para poderem lidar com as variações de tráfego da rede e heterogeneidades de dispositivos.

Além desses aspectos das redes sem fio causarem um impacto considerável na qualidade do *streaming*, algumas características de compressão espacial e temporal dos diferentes padrões de codificação tais como MPEG-4, H.264/AVC e em especial do padrão H.264/SVC, dos codificadores e decodificadores, podem influenciar consideravelmente na qualidade do vídeo.

Neste contexto esta dissertação discute, avalia e compara os impactos dos padrões de codificação, dos codificadores e decodificadores na qualidade do vídeo, quando utilizados em conjunto com redes de comunicação sem fio de 3ª geração, com especial ênfase nas diferentes formas de escalabilidade de vídeo, e considerando o emprego de esquema adaptativo anteriormente elaborado por integrantes do Laboratório de Televisão Digital da Universidade de Brasília.

1.2 Objetivos

1.2.1 Objetivo geral

Avaliar a qualidade de vídeo escalável recebido por um cliente conectado à internet por meio de uma rede sem fio de 3ª geração, fazendo uso ou não de um esquema adaptativo de vídeo escalável e utilizando métricas objetivas para avaliação da Qualidade de Vídeo (*Quality of Video - QoV*), combinadas a parâmetros de Qualidade de Serviço (*Quality of Service - QoS*).

1.2.2 Objetivos específicos

Como objetivos específicos, propõe-se:

- A) Avaliar e analisar sequências de vídeo que pertencem a diferentes classes de conteúdo, quando transmitidas em redes com limitações de recursos (por exemplo, de banda);
- B) Definir uma arquitetura e um processo de avaliação e aferição dessas sequências de vídeo;
- C) Comparar aspectos ligados ao emprego dos padrões de codificação H.264 SVC (recentemente proposto) e padrão MPEG-4 (o mais utilizado atualmente), quando utilizados para compressão e posterior transmissão em redes de comunicação sem fio de 3ª geração;
- D) Avaliar o desempenho de esquema adaptativo anteriormente proposto para o padrão H.264 SVC, considerando uma rede 3G integrada à Internet.

1.3 Organização

Esta dissertação está organizada da seguinte forma: no Capítulo 2 são apresentados alguns conceitos de codificação e escalabilidade de vídeo, os padrões de codificação H.264/AVC e H.264/SVC. São abordados, também, conceitos de tipos de escalabilidade e níveis de granularidade encontrados em SVC (*Scalable Video Coding*).

No Capítulo 3 é apresentada uma visão geral sobre redes sem fio, com foco predominante em redes de 3ª geração, abordando histórico, evolução, arquitetura, tipos de redes e respectivos padrões.

O Capítulo 4 aborda a transmissão de vídeo em redes de 3ª geração, bem como as arquiteturas e protocolos utilizados para transmissão de vídeo no âmbito da Internet. É mostrado e detalhado, também, o esquema adaptativo utilizado nesta dissertação para transmissão e entrega de vídeo escalável, bem como seu funcionamento e seus componentes.

No Capítulo 5 são abordados aspectos relacionados às métricas de qualidade de vídeo (QoV), às métricas de qualidade de serviço em redes (QoS) e qualidade de experiência (QoE), bem como as ferramentas utilizadas para avaliação, codificação e decodificação do vídeo.

O Capítulo 6 apresenta alguns trabalhos relacionados considerados relevantes para o desenvolvimento desta dissertação e também são descritos os cenários utilizados para os experimentos, as configurações dos equipamentos e dos vídeos utilizados.

No Capítulo 7 são apresentados e analisados os resultados enquanto no Capítulo 8 são apresentadas as conclusões e os trabalhos futuros propostos para a continuidade desta dissertação.

2 Codificação e Escalabilidade de Vídeo

2.1 Introdução

Imagens digitais, processamento de vídeo e técnicas de transmissão, podem ser combinados de forma eficiente e econômica em relação ao armazenamento e ao transporte em enlaces e redes de comunicação. Em um ambiente multimídia, a imagem pode ser descrita como uma representação bidimensional de cenas ou outros dados visuais [Punchihewa *et al* 2003]. O vídeo pode ser tratado como uma sequência de imagens, porém muitas vezes contém informações adicionais tais como cronometragem ou sinais de sincronização.

Uma sequência pode ser formada por um programa completo de vídeo totalmente codificado. A sequência se inicia através de um *sequence header* que contém informações referentes à sequência a ser tratada. [Megrich 2009].

Na codificação MPEG e em outros tipos de codificação como H.264 AVC e SVC, são considerados 3 (três) tipos de "frames" (I, P e B), cujas características serão detalhadas em breve. Um GOP (*Group of pictures*) se inicia, necessariamente, por um quadro I, logo após podem estar alocadas imagens P e/ou B. Uma sequência de vídeo é formada por um ou mais GOP's, que têm seus pontos de início e término comumente delimitados a partir de alterações bruscas na cena tratada [Megrich 2009].

A maioria dos dados da imagem contém uma redundância significativa na forma de valores de *pixels* adjacentes correlacionados (redundância espacial), e entre quadros de vídeo (redundância temporal) [Punchihewa *et al* 2003].

Um dos aspectos importantes no projeto de padronização do SVC (*Scalable Video Coding*) é a utilização de componentes baseados no padrão H.264/AVC e MPEG-4 parte 2, isso inclui a compensação de movimento, predição intra e predição entre camadas, codificação de macroblocos e a NAL (*Network Abstraction Layer*). Tais conceitos serão explanados a seguir.

2.1.1 Conceitos básicos

O modelo de codificação utilizado em padrões mais recentes como o H.264/AVC e o H.264/SVC é baseado na compensação de movimento, na codificação temporal e em

algumas transformações espaciais. Essa codificação é caracterizada pela predição de movimento sincronizada na codificação e na decodificação.

A compensação de movimento faz parte dos componentes utilizados para codificação no padrão SVC. A compensação de movimento utiliza a estimativa de movimento, onde os quadros são preditos através de referências subseqüentes. No processo de reconstrução das imagens, os quadros de referência são empregados para a predição dos quadros correntes, utilizando para isso os vetores de movimento.

A predição intra é formada por meio do quadro atual mais o posterior, corresponde à imagem $n + 1$. A predição *inter* (entre quadros) é gerada pela compensação de movimento a partir de um ou mais quadros de referência. A predição dos macroblocos pode ser desenvolvida com base em um ou mais quadros anteriores ou posteriores. O GOP pode ser formado por quadros dos tipos Intra, Preditivo e Bidirecional, conforme descrição a seguir.

Quadro Intra ou I: É intracodificado, e por essa razão não necessita dos demais quadros para ser decodificado. Nenhuma forma de compensação de movimento é aplicada a ele. O quadro I é empregado como referencial para os quadros P e B.

Quadro Preditivo ou P: Nos quadros do tipo P são aplicadas formas de predição de movimento para frente, já que eles são quadros intercodificados e são baseados no quadro de referência I.

Quadro Bidirecional ou B: Nos quadros do tipo B também são aplicadas formas de predição, pois esses tipos de quadros também são quadros intercodificados. A predição é feita de forma bidirecional (possui como base tanto o quadro I quanto o quadro P). No padrão de codificação H.264/AVC e no padrão de codificação H.264/SVC, descrito na próxima seção, o quadro B também pode possuir como base outro quadro B. A Figura 2.1 representa um *GOP* com quadros dos tipos I, P e B:

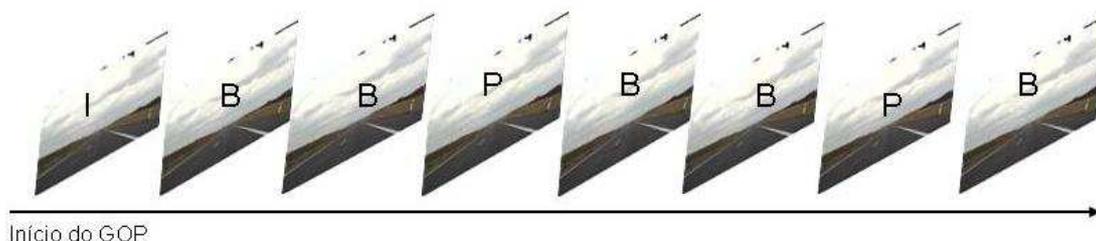


Figura 2.1: GOP formado por oito quadros.

Cada quadro pode representar uma imagem e essa imagem é particionada e formada por macroblocos de tamanho fixo, conforme apresentada na Figura 2.2. Este particionamento já era adotado por outros padrões de codificação de vídeo como o H.261 [Richardson 2003][Wiegand *et al* 2003]. Os macroblocos são agrupados dentro de fatias, também chamadas de *slices*. A Figura 2.3 apresenta um quadro formado por oito fatias (*slices*).



Figura 2.2 : Quadro com macroblocos destacados.



Figura 2.3: Quadro formado por oito slices.

As fatias são sequências de macroblocos agrupados que são processados na ordem para a formação da imagem. Cada fatia possui um cabeçalho para identificar sua localização. Cada macrobloco é formado por seis blocos, sendo quatro blocos de luminância e dois de croma, podendo ser representada por Y:Cb:Cr (YUV). A Figura 2.3 representa um quadro com formação 4:2:0 (YUV). As fatias podem ser codificadas utilizando os tipos I, P, e B[Wiegand *et al* 2003].

I slice: Todos os macroblocos do *slice* ou fatia são codificados usando a predição intra.

P slice: Além de codificado com *I slice* ou fatia, alguns macroblocos do *P slice* podem ser codificados usando predição entre fatias com predição de compensação de movimento.

B slice: Pode ser codificado com *P slice* ou fatia, alguns macroblocos do *B slice* podem ser codificados usando predição entre fatias com duas predições de compensação de movimento.

2.1.2 Padrão H.264/AVC

O Padrão H.264/AVC foi projetado para atender áreas de aplicação tais como:

- *Broadcast* sobre cabo, satélite, modem a cabo, DSL;

- Armazenamento óptico (DVD, por exemplo) e em dispositivos magnéticos;
- Vídeo sob demanda e serviços de *streaming* multimídia sobre ISDN (*Integrated Services Digital Network*), modem, DSL, LAN, redes sem fio;
- Serviços de mensagens multimídia (MMS) sobre ISDN, DSL, ethernet, LAN, rede sem fio e redes móveis;

O H.264/AVC abrange uma VCL (*Video Coding Layer*), que foi projetada para representar eficientemente o conteúdo do vídeo e uma NAL, que formata a representação VCL do vídeo e fornece informações do cabeçalho de forma adequada para transporte por uma variedade de camadas de transporte ou mídia de armazenamento. A Figura 2.4 apresenta a estrutura do codificador de vídeo H.264/AVC [Wiegand *et al* 2003].

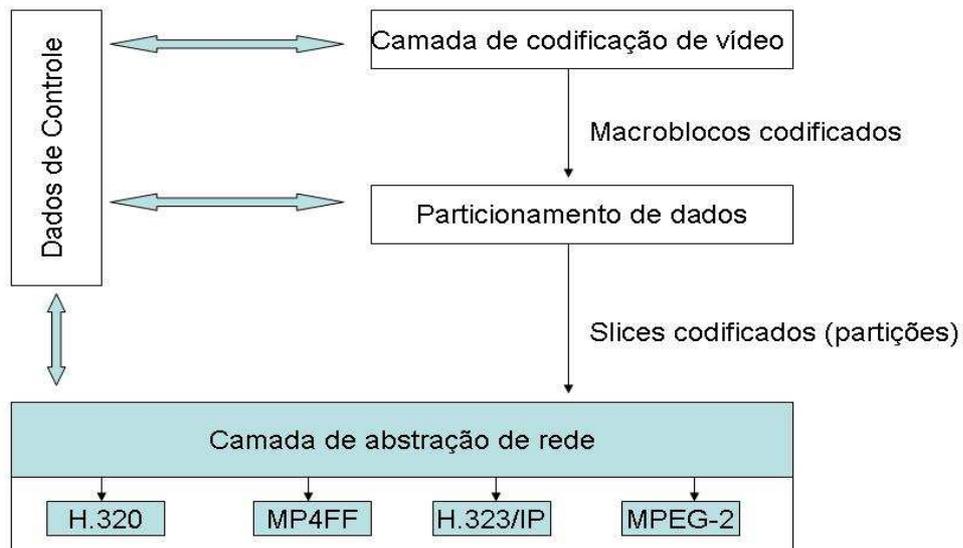


Figura 2.4: Estrutura do codificador H.264/AVC [Adaptada de [Wiegand *et al* 2003].

A NAL foi desenvolvida para fornecer uma interface de rede amigável (*Network Friendliness*) para permitir uma adaptação dos dados vindos da camada VCL para a camada de transporte.

Os conceitos chave referentes à camada de abstração de rede podem ser resumidos da seguinte forma [Wiegand *et al* 2003]:

Unidades NAL: Os dados de vídeo codificados estão organizados dentro de unidades NAL, cada pacote contém um número inteiro de bytes. O primeiro byte de cada unidade NAL é um cabeçalho que contém um indicador do tipo de dados da unidade NAL e os bytes restantes contém o *payload* do cabeçalho indicado. As unidades NAL são classificadas em unidade VCL-NAL, que contém dados que representam a informação do vídeo e as unidades não VCL-NAL, que contém qualquer informação adicional.

Conjunto de parâmetros: Possui informações que raramente mudam. Existem dois tipos de conjunto de parâmetros.

- Conjunto de parâmetros sequencial: que é aplicado a uma série de imagens de vídeo codificadas consecutivas;
- Conjunto de parâmetros de imagens: que é aplicado na decodificação de uma ou mais imagens individuais.

Unidades de acesso: É o conjunto de unidades NAL em uma forma especificada. A decodificação de cada unidade de acesso resulta numa imagem decodificada. Cada unidade de acesso contém um conjunto de unidades VCL-NAL que juntas compõem uma imagem primária codificada.

Na VCL segue a abordagem de codificação de vídeo híbrida baseada em blocos, onde cada imagem codificada é representada por uma unidade em forma de bloco chamada de macrobloco.

Para melhorar os métodos de predição no padrão e melhorar a eficiência da codificação, alguns recursos foram adicionados:

Compensação de movimento de blocos de tamanhos variáveis com blocos de tamanho pequeno: Este padrão proporciona maior flexibilidade na seleção de tamanhos e formas dos blocos para compensação de movimento.

Múltiplas referências para compensação de movimento da imagem: O padrão estende-se sobre a melhor técnica de seleção de imagem de referencia para permitir uma codificação eficiente, permitindo o codificador selecionar, para efeitos de compensação de movimento, um número maior de imagens que foram decodificadas e armazenadas no decodificador.

Codificação de entropia aritmética: Um método avançado de codificação de entropia é incluído ao H.264/AVC, conhecido como codificação aritmética. O uso mais efetivo desta técnica é encontrado no H.264/AVC para criar um poderoso método de codificação de entropia, conhecido como CABAC (*Context-Adaptive Binary Arithmetic Coding*).

Codificação de entropia de contexto adaptativo: Dois métodos de codificação de entropia são aplicados no H.264/AVC, o CAVLC (*Context-Adaptive Variable-Length Coding*) e o CABAC, ambos utilizam contexto baseado em adaptabilidade para melhorar o desempenho do padrão H.264/AVC.

2.1.3 Padrão H.264/SVC

O SVC (*Scalability Video Coding*), considerado como uma extensão do H.264/AVC (*Advanced Video Coding*) é o novo padrão de codificação de vídeo do ITU-T VCEG (*Video Coding Experts Group*) e da ISO/IEC MPEG (*Moving Picture Experts Group*).

O padrão H.264/SVC foca de forma bem específica os aspectos de escalabilidade envolvidos nos processos de codificação e decodificação de vídeo, em oposição à codificação e decodificação não escaláveis, possibilitando a adaptação do *bitstream* de forma compatível com a disponibilidade de banda.

A motivação principal para a sua criação decorre, principalmente, da ampla variedade de dispositivos (palmtops, terminais celulares, netbooks, receptores de TV caseiros, projetores,...), com diferentes resoluções e capacidades, juntamente com a diversidade de redes que podem ser utilizadas para transporte de vídeo (tanto cabeadas quanto redes sem fio, tanto comutadas por pacotes quanto comutadas por circuitos), e comumente baseadas em padrões diferentes.

A escalabilidade é provida com base em uma quantidade de camadas (uma sequência de bits codificados) voltadas para prover adaptabilidade, sendo a camada base (*base layer*) acrescida de uma ou mais camadas de reforço (*enhancement layers*). Essa adaptabilidade permite, quando bem projetada e utilizada, auxiliar no controle de congestionamento realizado por protocolos da camada de transporte, bem como produz reflexos na qualidade de serviço (QoS) provida pela rede e na qualidade de experiência (QoE) vivida pelo usuário ao receber uma ou mais sequências de vídeo.

Conforme [Stretch 2011] “a extensão SVC introduz uma noção não presente no H.264/AVC de camadas dentro do fluxo codificado. A camada base do fluxo de vídeo é codificada com a menor representação temporal, espacial e de qualidade. As camadas de reforço codificadas adicionam informações que, usando a camada base como ponto de partida, podem ser usadas para reconstruir o vídeo com maior qualidade, resolução, ou versões temporais durante o processo de decodificação. A decodificação da camada base e somente as camadas de reforço subseqüentes requerem, um decodificador que possa produzir um fluxo de vídeo com as características desejadas. A Figura 2.5 apresenta a estrutura em camadas de um fluxo H.264/SVC. Durante o processo de codificação, é tomado um cuidado para codificar uma determinada camada usando referencia somente para camadas níveis inferiores. Desta forma, o fluxo codificado pode ser truncado, em qualquer ponto arbitrário e ainda permanece um fluxo, válido decodificado.”

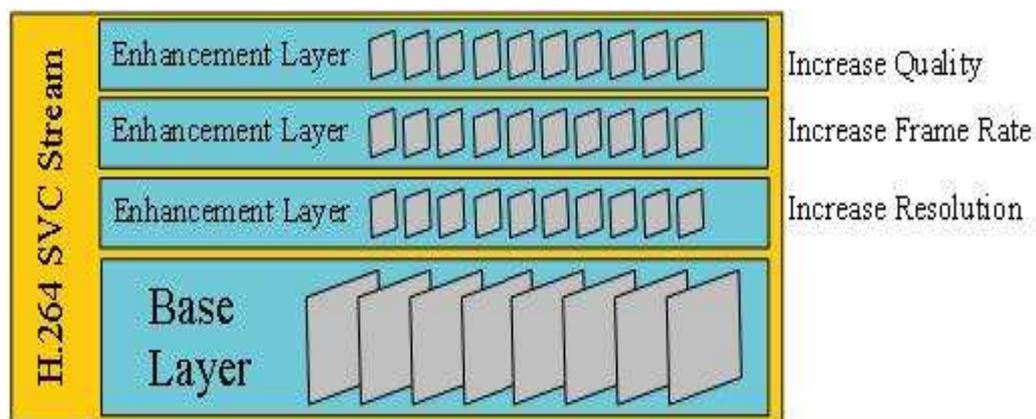


Figura 2.5: Estrutura em camadas de um fluxo H.264/SVC

Ainda, segundo [Stretch 2011] “esta abordagem em camadas permite a geração de um fluxo codificado que pode ser truncado para diminuir a largura de banda consumida ou diminuir os requisitos de decodificação computacionais. O processo de truncagem consiste em extrair as camadas necessárias do fluxo de vídeo codificados sem nenhum processamento adicional no próprio *stream*.”

A escalabilidade considera então a possibilidade de remoção de parte do *bit stream* de vídeo a fim de adaptar o vídeo para as diferentes necessidades dos usuários, condições de redes e terminais com capacidades diferentes. Quando a camada de *bit stream* não permite essa separação é considerada apenas uma camada base [Schwarz *et al.* 2007] [Wien *et al* 2007] [Schwarz and Wien 2008].

Em termos de camada de reforço, é interessante descrever como se dá o processo de decisão do decodificador para decodificação dos macroblocos. Tal processo é mostrado na Figura 2.6:

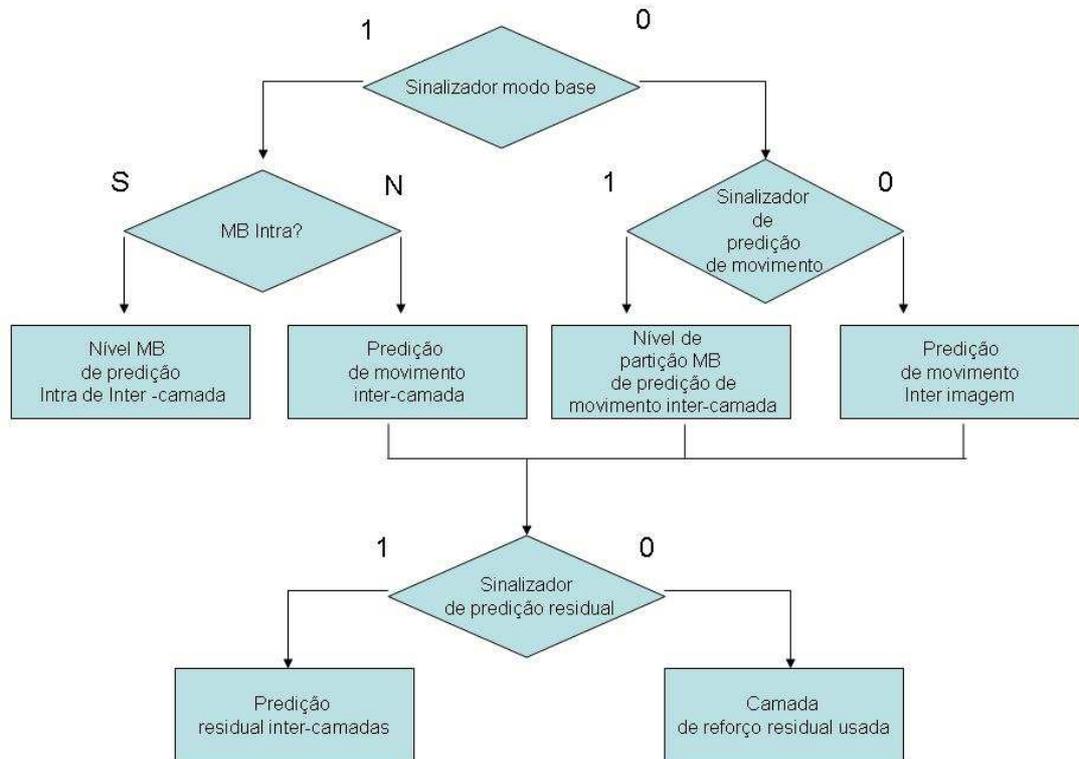


Figura 2.6: Processo de decisão do decodificador para decodificação do macrobloco das camadas de reforço (Adaptada de [Narvekar *et al* 2009]).

O tipo de macrobloco (MB) da camada de reforço é sinalizado usando um elemento da sintaxe chamado sinalizador de modo base (*base mode flag* – BMF). O SVC suporta vários tipos de modos de predição entre camadas [Narvekar *et al* 2009]:

Predição Intra entre camadas: Quando o BMF for igual a 1 (um) e for colocado o bloco da camada de referência na codificação de forma intra, a predição intra entre camadas é usada. Deste modo, a camada de referencia co-localizada no bloco intra codificado é aumentada e adicionada à camada de reforço residual para obter um bloco da camada de reforço [Narvekar *et al* 2009].

Predição de movimento entre camadas: Quando o BMF for igual a 1 (um) e for colocado o bloco da camada de referencia na codificação de forma intra, a predição de movimento inter-camadas é usada. Mas quando o BMF for igual a 0 (zero) e o

sinalizador de predição de movimento (*Motion Prediction Flag* - MPF) for igual a 1 (um), a predição de movimento inter-camadas é usada no nível de partição do MB. Assim, os vetores de movimento da camada de referencia são aumentados e usados para o desempenho da compensação de movimento para obter um bloco da camada de reforço [Narvekar *et al* 2009].

Predição Residual entre camadas: Neste modo, a informação residual da camada de referencia é usada para predição de dados residual de blocos inter codificados na camada de reforço independentemente dos valores de BMF ou MPF. Isso é indicado pelo sinalizador de predição residual (*residual-prediction flag* – RPF) quando o valor for 1 (um) [Narvekar *et al* 2009].

Um dos benefícios do SVC está na transmissão do vídeo onde há a utilização de diversos cenários simultâneos, com clientes heterogêneos, onde podem existir múltiplos fluxos de *bits* vindo da mesma fonte, mas precisam ser diferenciados, de acordo com cada receptor ou banda disponível na rede, na resolução da imagem, na taxa de quadros e de *bits* apresentados ao mesmo tempo.

Outro benefício do uso de vídeo escalável é que um fluxo de *bits* possui diferentes partes em termos de qualidade de vídeo. Em um mesmo fluxo de *bits* podem existir *bits* mais importantes, em termos de qualidade, em relação a outros. Essa característica é importante em cenários de transmissão de vídeo com variações excessivas e elevadas taxas de perda de pacotes.

A escalabilidade de vídeo permite uma fácil adaptação com alta robustez na transmissão em dispositivos e canais, facilitando a transmissão de *streaming*, tanto *unicast* quanto *multicast*, sobre redes heterogêneas sem a necessidade de transcodificação [Narvekar *et al* 2009]. Utilizando uma aplicação com esquema SVC adaptado corretamente, o conteúdo de origem será codificado uma vez, resultando em altas taxas de *bits*, que podem ser representadas com a utilização de baixa resolução.

Com isso, as mensagens de *feedback* referente à capacidade do terminal e às condições de rede podem ser utilizadas para, se possível, remover partes do vídeo que não foram solicitadas a partir de fluxos de *bits* escalável antes de ser transmitido. Sendo assim, um cliente que possui restrições como: baixa capacidade de bateria, baixo poder de processamento ou resolução de vídeo precisará decodificar apenas uma parte dos

fluxos de *bits* enviados. Esses terminais serão atendidos por um fluxo de *bits* único e escalável.

Dessa forma, as perdas devido ao congestionamento ou a erros na transmissão de partes importantes do vídeo, podem ser melhoradas de maneira significativa. Se considerarmos as aplicações de vídeo atuais e as futuras, o sucesso de qualquer um dos vídeos codificados com diferentes tipos de escalabilidade será garantido, considerando a eficiência da codificação do vídeo, com pequeno aumento na complexidade na decodificação se comparada à decodificação de vídeo de camada única.

Há diferentes tipos de escalabilidade que podem ser combinados entre si, gerando assim, diferentes resoluções temporal-espacial, onde diferentes taxas de *bits* podem ser suportadas dentro de um mesmo fluxo de *bit* escalável.

2.2 Tipos de Escalabilidade

Os tipos de escalabilidade de vídeo mais usuais são a escalabilidade temporal, a escalabilidade espacial e a escalabilidade SNR (*Signal-to-Noise Ratio*) ou de qualidade. Além disso, elas podem ser combinadas para formar diversos cenários [Shahid *et al.* 2009] [Schierl *et al* 2007].

2.2.1 Escalabilidade Temporal

A escalabilidade temporal é uma técnica que resulta em uma melhora significativa no comportamento em tempo real, além de ser uma das estruturas em camadas mais flexíveis, com baixo custo de codificação se comparada à sua eficiente codificação. Em redes multimídia, essa técnica possui fluxo de transporte independente, com qualidade de serviço diferenciada, muito eficiente na transmissão de vídeo em tempo-real [Xiaoqing Zhua *et al.*2008]. Para lidar com transmissões em terminais móveis, trabalhar com vídeo escalável é uma excelente opção, pois permite trabalhar com diferentes níveis de qualidade.

Um fluxo de *bits* fornece escalabilidade temporal quando o conjunto de unidades de acesso pode ser dividido em camada base temporal e uma ou mais camadas de reforço temporal. Este processo pode ser identificado da seguinte forma: um identificador temporal identifica a camada base que começa com 0 e é aumentada para 1 a partir da camada de reforço. A Figura 2.7 apresenta um aumento gradual da camada temporal.

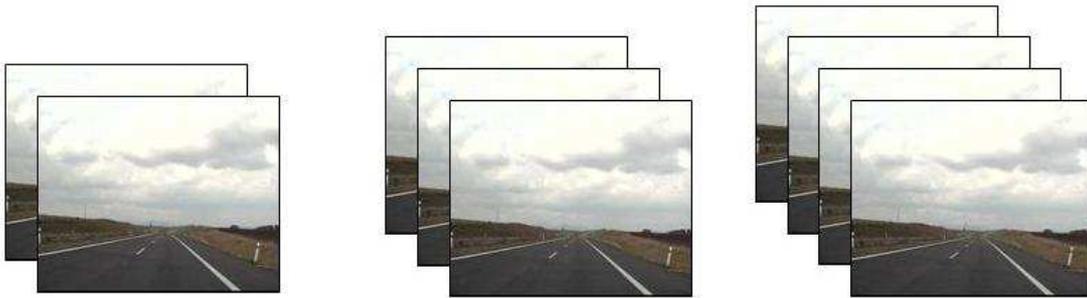


Figura 2.7: Diferentes camadas temporais.

Para codificadores de vídeo híbridos, a escalabilidade temporal pode ser identificada através de uma predição de compensação de movimentos, as imagens da camada temporal de referência possuem uma taxa de *bits* menor ou igual às imagens da camada temporal a ser predita. Em escalabilidade temporal, a taxa de quadros da camada de reforço é maior se comparada com as taxas de quadros das camadas inferiores, inclusive com relação à camada base [Pliakas, T. *et al* 2007]. Qualquer imagem pode ser marcada e utilizada como referência para compensação de movimento, independente do tipo de codificação correspondente à imagem [Schwarz, H. and Wien, M. 2008].

A Figura 2.8 mostra um exemplo de escalabilidade temporal utilizando hierarquia baseada no frame B.

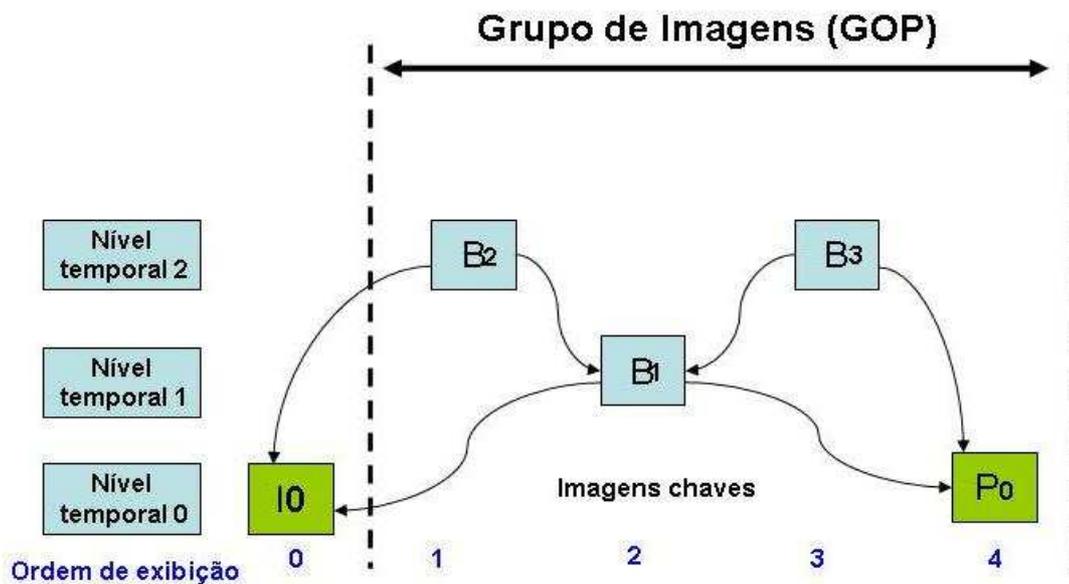


Figura 2.8: Escalabilidade temporal, hierarquia baseada no frame B.[Xiaoqing Zhua *et al.*2008].

Na Figura 2.8 os quadros I e P constituem a camada base, e os quadros B são preditos dos quadros I e P e constituem a segunda camada. Alguns quadros B são preditos de outros quadros B, formando assim a hierarquia baseada no quadro B.

Na escalabilidade temporal, a codificação é baseada na hierarquia do quadro B. Com a introdução do quadro B foi possível conseguir uma melhora significativa na codificação em termos de escala temporal. Essa concepção de utilização do quadro B foi introduzida ainda no padrão no H 264 AVC [Wiegand *et al* 2003].

Os codificadores que utilizam escalabilidade temporal podem ser classificados de acordo com a maneira como a redundância temporal é explorada. Uma das técnicas utilizadas pelos codificadores para explorar a redundância temporal é a compensação de movimento. Caso haja alguma mudança de posição no objeto, pode-se enviar apenas as informações referentes à nova posição, requerendo assim menos banda. O primeiro quadro é decodificado independentemente, como uma imagem parada, os demais são preditos a partir do primeiro quadro, de forma direta ou indiretamente.

2.2.2 Escalabilidade Espacial

As camadas de reforço possuem uma resolução maior ou igual a da camada base ou se comparada com outras camadas inferiores. A escalabilidade espacial corresponde à possibilidade de se transmitir imagens com diferentes resoluções. A camada de reforço pode ser predita com base tanto na camada inferior ou a partir do quadro anterior da mesma camada [Pliakas *et al* 2007]. A cada nova camada dentro de um fluxo de *bits* a resolução é melhorada.

As diferentes camadas espaciais são codificadas independentemente com parâmetros específicos de movimento. Para melhorar a eficiência da codificação das camadas de reforço foram introduzidos mecanismos de predição de inter-codificação. Os mecanismos de predição são feitos de forma comutáveis para deixar livre a escolha pelo codificador de qual informação será explorada da camada base para uma eficiente codificação da camada de reforço [Schwarz *et al.* 2007] [Wien *et al* 2007].

A escalabilidade espacial é alcançada através da codificação utilizando alguns mecanismos de predição inter-camadas, a fim de melhorar a eficiência na codificação. Esse método de predição possibilita ao máximo a reutilização das informações entre uma camada e outra, isso evita a redundância e faz com que a eficiência da codificação seja melhorada.

Para dar suporte à codificação com escalabilidade espacial, o SVC segue com a abordagem de múltiplas camadas de codificação que também é utilizada em outros codificadores: H.264/AVC, MPEG-2, H.263 e MPEG-4 Visual. Cada camada corresponde a uma resolução e possui um identificador [Schwarz *et al.* 2007]. O identificador para a camada base é igual a 0 e vai aumentando para 1 a medida que vão sendo incrementadas as camadas de reforço espacial.

Para diminuir a complexidade do decodificador, o SVC especifica que a decodificação das camadas espaciais deve ser feita na mesma ordem que a codificação.

Vídeos com diferentes resoluções espaciais e em determinados instantes de tempo têm que ser transmitidos, sucessivamente, em ordem crescente correspondente aos seus indicadores na camada espacial. Na predição inter-camadas estão incluídas técnicas de predição de movimento e predição residual. Um codificador deve alinhar todas as estruturas de predição temporal a todas as camadas espaciais.

No SVC existe a possibilidade de uma camada de reforço ser representada apenas por um retângulo de seu quadro de referencia, correspondente a camada que é codificada com uma resolução espacial superior ou semelhante. A Figura 2.9 apresenta um exemplo de escalabilidade espacial.

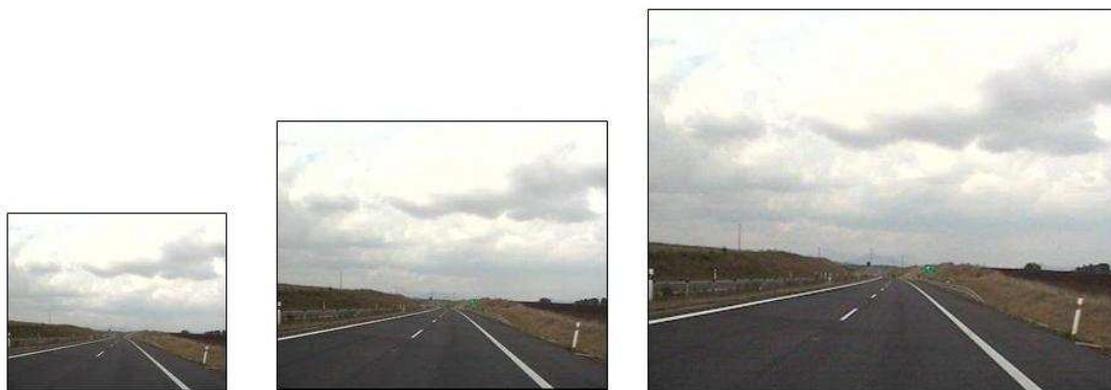


Figura 2.9: Diferentes camadas de resolução.

Para suportar a escalabilidade, o SVC trabalha com múltiplas camadas de codificação. Cada camada possui uma resolução e faz referencia a uma camada espacial. A predição inter-camadas e a predição temporal também são empregadas a cada camada de resolução espacial. A predição inter-camadas é importante para remover as redundâncias entre as camadas espaciais.

A Figura 2.10 representa uma idéia de escalabilidade espacial com a camada S_0 sendo a camada base e S_1 e S_2 suas respectivas camadas de reforço. A Figura mostra também um número de quadros que variam com tempo T_0 , com isso temos uma combinação de escalabilidade temporal e espacial.

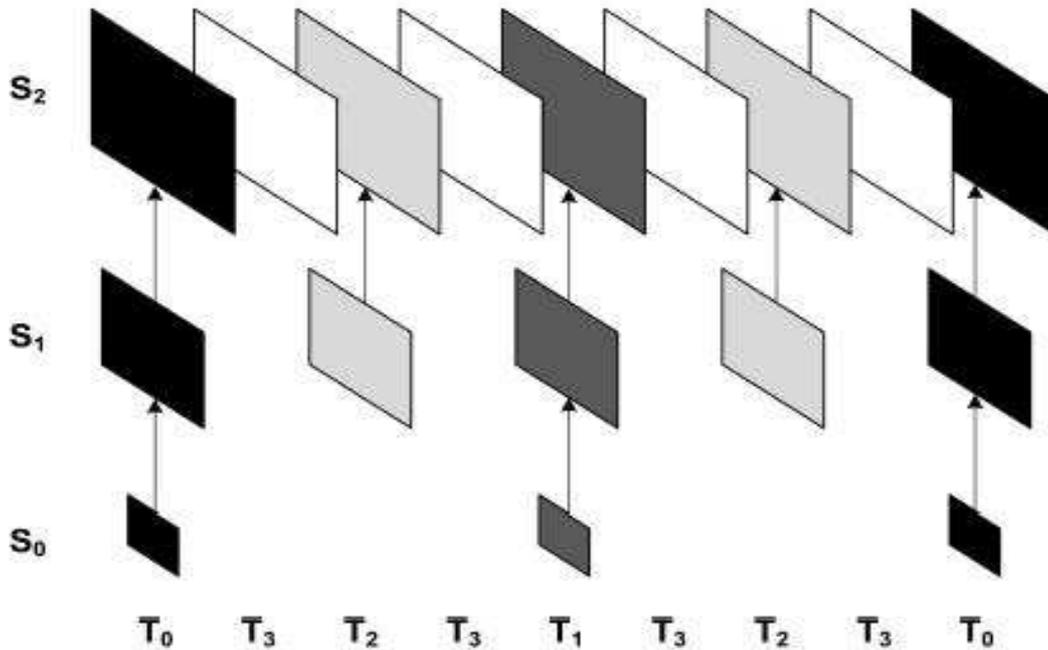


Figura 2.10: Representação de escalabilidade espacial com duas camadas de reforço [Kang *et al.* 2009]

2.2.3 Escalabilidade SNR

Na escalabilidade de qualidade ou SNR as camadas com resolução espacial e temporal permanecem as mesmas e apenas as camadas de qualidade são reforçadas. A escalabilidade SNR também é conhecida como escalabilidade de qualidade, pois permite a extração de *bit stream* com diferentes níveis de qualidade com base no *bit stream* completo. Quanto maior o SNR, maior será a qualidade da imagem produzida. Com a decodificação apenas da camada base, a imagem gerada será de baixa qualidade, à proporção que as camadas vão sendo decodificadas, haverá uma melhora significativa na imagem [Schwarz and Wiegand 2008]. A Figura 2.11 apresenta a codificação de vídeo utilizando escalabilidade de qualidade.



Figura 2.11: Escalabilidade de qualidade

A escalabilidade de qualidade pode ser considerada um caso especial de escalabilidade espacial, pois o tamanho da camada base e da camada de reforço são os mesmos [Schwarz and Wiegand 2008] [Schwarz and Wien 2008].

2.2.4 Escalabilidade Combinada

Os diferentes tipos de escalabilidade podem ser combinados. Cada camada possui resoluções diferentes para a imagem. É acrescentada, na codificação, uma estrutura hierárquica de múltiplos quadros com camadas de escalabilidade temporal, camadas de escalabilidade espacial e camadas de escalabilidade de qualidade. Em cada camada de reforço, é alcançado um alto nível temporal, de resolução e de qualidade para a imagem. São utilizados no processo de codificação os níveis de granularidade de qualidade, com isso é possível formar imagens com diferentes níveis de granularidade, diferentes níveis de qualidade [Schwarz *et al.* 2007].

A estrutura do SVC é organizada em camadas de dependência. Essa dependência normalmente representa uma camada com escalabilidade espacial específica. Em alguns casos, pode haver dependência de duas camadas com resolução espacial. As dependências das camadas são apresentadas pelo número identificador.

A resolução espacial não diminui em relação à primeira camada e à camada seguinte. Quando uma camada de referencia é uma camada de dependência com resolução espacial e esta possui representações de qualidade diferentes, ela precisa ser sinalizada para saber qual será utilizada na predição inter-camadas.

Para permitir uma adaptação simples dos fluxos de *bit*, o SVC fornece alguns meios pelos quais os *substreams* que estão contidos em um fluxo de *bits* escalável completo possam ser identificados facilmente. Um fluxo de *bit stream* não precisa necessariamente fornecer todos os tipos de escalabilidade.

2.3 Níveis de Granularidade

A escalabilidade de qualidade é dividida em três classes, conforme a granularidade, em grossa (*Coarse-Grain Quality Scalable Coding* - CGS), média (*Medium-Grain Quality Scalability* - MGS) e fina (*Fine-Grain Quality Scalable* - FGS). Nesta seção adotaremos a referência [Schwarz *et al.* 2007] como base.

2.3.1 Escalabilidade de granularidade grossa

Os dados comprimidos são embutidos em um único *bitstream* para permitir que diferentes taxas de *bits* possam ser obtidas. Ela pode ser considerada espacial caso a camada base e a camada de reforço possuam a mesma resolução da escalabilidade espacial.

CGS pode ser considerada um caso especial de escalabilidade espacial, quando as camadas consecutivas possuem a mesma resolução [Schwarz *et al.* 2007]. Na escalabilidade de granularidade grossa as operações de aumento da taxa de *bits* no mecanismo de predição entre camadas são omitidos.

Essa técnica permite que o servidor de vídeo faça adaptação da taxa de *bits* do vídeo a ser distribuído de acordo com às disponibilidades da rede usando uma granularidade muito fina, onde todos os *bits* recebidos se tornam úteis.

2.3.2 Escalabilidade de granularidade média

No conceito de MGS algumas unidades NAL da camada de reforço podem ser descartadas a partir de um *bit stream* de escalabilidade de qualidade. Aumenta a eficiência usando um módulo de predição mais flexível, onde ambos os tipos de camada (base e reforço) pode ser referenciado.

Para cada imagem uma sinalização é transmitida, seja na reconstrução da qualidade da camada base ou na reconstrução da camada de reforço, a imagem de referencia é utilizada para predição de compensação de movimento. A predição de compensação de movimento para codificação de escalabilidade de qualidade é sempre aplicada empregando a maior qualidade disponível correspondente a imagem de referencia. Durante o processo de codificação as camadas MGS não sabem que tipo de representação de imagens estarão disponíveis no decodificador.

O codificador decide qual referencia será usada para a estimativa de movimento, (compensação de movimento), essas decisões influenciam na eficiência da codificação.

2.3.3 Escalabilidade de granularidade fina

No FGS a estimativa de movimento no codificador e a compensação de movimento no decodificador utilizam o mesmo nível de qualidade da imagem de referencia. Isso é possível devido ao fato da camada base estar sempre disponível no receptor [Chan *et al* 2008]. No FGS, pode se escolher qual será a camada de qualidade empregada para a predição de movimento.

FGS emprega uma técnica de *bit*-plano avançado onde diferentes camadas são responsáveis pelo transporte de subconjuntos distintos de *bits* correspondente a cada informação dos dados. Nessa técnica é permitido o truncamento de dados em qualquer ponto do fluxo codificado, a fim de apoiar o refinamento progressivo.

3 Evolução de Redes Sem Fio e Redes de 3ª Geração

3.1 Introdução

A ITU-T (*International Telecommunication Union*) desde 1998 vem coletando informações de diversos países para melhorar constantemente o padrão de comunicações móveis de 3ª geração. A primeira e a segunda geração de sistemas celulares de comunicações móveis são para voz e dados de baixa velocidade, enquanto a terceira geração de comunicações móveis (3G) é para serviços de alta velocidade. As redes 3G possuem uma capacidade maior, mais flexibilidade nas transmissões de alta velocidade e oferecem até 2 Mbps para serviços multimídia de alta qualidade [Liu *et al* 2009].

3.2 Padronização

A padronização adotada por organizações normativas como o 3GPP (*Third Generation Partnership Project*) inclui os seguintes aspectos [3GPP TS 2010]:

- Acesso global transparente ao sistema móvel;
- Compatibilidade com sistemas 3G;
- Convergência entre rede móvel e rede fixa.
- Altas taxas de dados para redes de comunicação sem fio;
- Transferência de dados por comutação de pacotes e comutação por circuitos;
- Introdução a aplicações multimídia;

A interface de rádio constitui elemento de crucial importância na padronização global, considerando a escassez de banda e a necessidade de fornecer acesso universal para terminais móveis de diferentes redes 3G.

3.3 Evolução de redes sem fio

Considerando que muitas operadoras voltadas para o provimento de serviços de comunicação móvel mantêm em operação infraestruturas de redes de gerações anteriores a redes 3G (como as gerações 2G e 2,5G), apresenta-se aqui breve discussão a respeito das evoluções das redes sem fio. Nota-se ainda que algumas operadoras buscam otimizar o emprego das diferentes infraestruturas disponíveis, com base em aspectos como balanceamento de carga e alocação de recursos a terminais com base no tipo de tráfego e nos recursos disponíveis em cada rede.

A explosão da Internet teve um impacto na demanda para o avanço dos serviços de comunicação sem fio. Observa-se que a capacidade de transmissão de dados em um sistema móvel 2G, como o GSM (*Global System for Mobile Communications*), é baixa, para muitos serviços de Internet, o que motivou a criação de alternativas como o HSCSD (*High Speed Circuit Switched Data*), GPRS (*General Packet Radio Service*) e EDGE (*Enhanced Data Rates for Digital Evolution*) [Hartung *et al* 2007] [Janevski 2003], a serem descritos a seguir.

3.3.1 HSCSD

O HSCSD (*High Speed Circuit Switched Data*) é um programa de atualização das redes GSM. Nenhum equipamento extra é requisitado. Nas redes GSM, um único *slot* de tempo é deslocado para cada assinante em voz ou dados. A taxa de transferência padrão em GSM é de 9,600 *bps*, podendo chegar a 14,400 *bps*. O HSCSD fornece a um único assinante acesso simultâneo a múltiplos canais (*slots* de tempo), subindo de quatro para oito em um único quadro TDMA (*Time Division Multiple Access*). No entanto é mais caro para os assinantes finais pagarem por ocuparem simultaneamente múltiplos *slots* de tempo [Janevski 2003].

Assumindo uma taxa padrão de transmissão de 14.4 *Kbps* e usando quatro *slots* de tempo com HSCSD a taxa de dados teórica chegaria a 57.6 *Kbps*. Com isso seria permitido acesso à Internet com a mesma velocidade de muitos *modems* discados (56 *Kbps*). Embora o HSCSD seja fácil de ser implementado em redes 2G, existe uma desvantagem quanto à multiplexação, pois os quatro *slots* de tempo permanecem ocupados durante toda a conexão. Outro problema é em relação ao *handover*, a menos que o mesmo *slot* de tempo esteja disponível em canal de tráfego nas diferentes células até o final da chamada.

O HSCSD é baseado na tecnologia de comutação por circuitos, já o GPRS é um complemento para comunicações com outras redes baseadas em comutação por pacotes.

3.3.2 GPRS

O rápido crescimento do uso da Internet aumentou a demanda por serviços de dados em redes sem fio. O GPRS (*GSM Packet Radio Service*) é definido como uma complementação do sistema GSM. É o primeiro passo para integração da Internet com redes móveis celulares. A diferença entre o HSCSD e o GPRS, é que este é aplicado

principalmente em transferência de pacotes de dados [Janevski 2003]. A arquitetura da rede GPRS pode ser representada pela Figura 3.1:

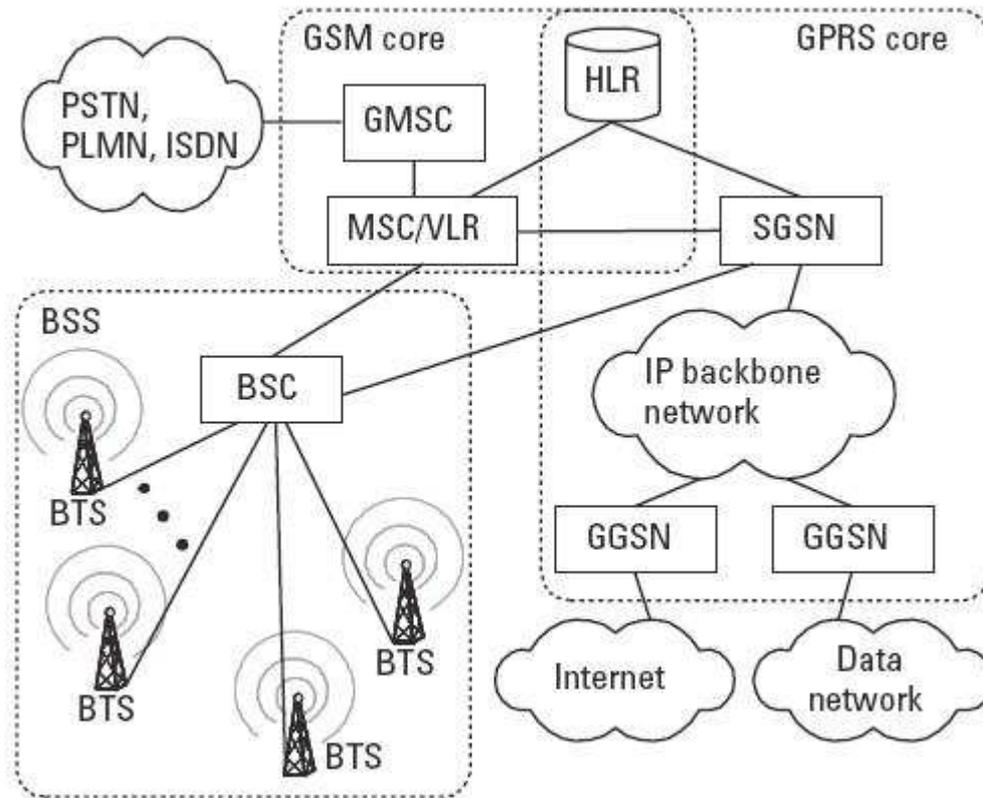


Figura 3.1: Arquitetura da rede GPRS [Janevski 2003].

A rede GPRS considera a existência de uma arquitetura GSM, na qual dois novos nós de rede foram acrescentados: O SGSN (*Serving GPRS Support Node*) e o GGSN (*Gateway GPRS Support Node*). O SGSN é responsável pela entrega dos pacotes de/para a estação móvel dentro da área de serviço, onde algumas de suas principais tarefas são o gerenciamento do móvel, o roteamento de pacotes, o gerenciamento do *link* móvel, a autenticação e as funções de tarifação. O GGSN atua como uma interface entre a rede de pacotes GPRS e a rede externa baseada em pacotes. Para cada sessão em GPRS, um contexto PDP (*Packet Data Protocol*) é criado, e descreve a sessão. Podem existir vários SGSN ou GGSN. No GPRS são permitidas definições de perfis de QoS (*Quality of Service* - Qualidade de Serviço) usando parâmetros de precedência de serviços, confiabilidade, atraso e vazão.

São consideradas três tipos de prioridades: alta, normal e baixa. Para confiabilidade são descritas algumas características tais como de perda probabilidade, duplicação e decomposição de pacotes. O atraso é definido em atraso médio e o atraso máximo em 95% de todas as transferências. A vazão refere-se à taxa de *bit* máxima e

taxa de *bit* principal. Para gerenciamento de localização o GPRS possui três estados: o modo inativo, o modo pronto (*ready*) e o modo de espera. No modo inativo, a rede não sabe a localização da estação móvel. Quando a estação recebe e envia pacotes, ela está no estado de pronto (*ready*), neste estado a rede sabe onde está localizada a estação móvel.

Para gerenciamento de localização no estado de inativo, a área de localização do GSM está dividida em várias áreas de roteamento (RAs – *Routing Area*). Para localizar um móvel no estado de espera a rede executa um processo de *paging* (procura) na área de roteamento corrente. No estado de pronto (*ready*) a estação não precisa executar um processo de *paging*, já no estado inativo a rede faz um *paging* em todas as BTs (*Base Transceiver Station* – Estação Rádio Base) na área de localização da estação móvel.

Também fazem parte da rede o MSC (*Mobile Switching Center*), o VLR (*Visitor Location Register*), o GMSC (*Gateway MSC*), o AuC (*Authentication Center*) e o EIR (*Equipment identity Register*). O GMSC está conectado à PSTN (*Public Switched Telephone Network*).

3.3.3 EDGE

A EDGE (*Enhanced Data Rates for Digital Evolution*) foi criada para fornecer altas taxas de dados em serviços baseados em pacotes com alta largura de banda usando a rede móvel 2G existente. Na EDGE há uma melhoria na vazão por *slot* de tempo tanto na HSCSD, quanto na GPRS. A EGPRS ou a GPRS melhorada é a atualização das redes EDGE, enquanto a ECSD é a HSCSD melhorada. Na ECSD a taxa de dados por *slot* de tempo é de 64 *Kbps* devido às limitações da interface aérea, mas a taxa de dados por *slot* de tempo é triplicada quando todos os *slots* são utilizados para uma única conexão em EGPRS e o pico da vazão ultrapassa a 384 *Kbps*.

A tecnologia EDGE é também usada sobre os sistemas D-AMPS, onde é fornecida taxa de dados de 473 *Kbps* por 30 *kHz* transportados. A tecnologia EDGE é uma opção para serviços 3G. A rede EDGE pode coexistir com a rede UMTS para fornecer serviços de alta velocidade para uma larga área de cobertura [Janevski 2003].

3.3.4 UMTS

O sistema UMTS (*Universal Mobile Telecommunication System*) foi criado para atender novos serviços como: serviços multimídia, fornecimento de conteúdo e *roaming* global.

A rede UMTS possui uma variedade de serviços com diferentes características de QoS, essas características foram divididas em serviços de tráfego em tempo real, que englobam as classes convencionais, classes de *streaming* e serviços de aplicações que não necessariamente são em tempo real, como as classes interativas e as classes de *background*.

O 3GPP especificou a banda de frequência para a rede UMTS para os modos de acesso a rádio FDD (*Frequency Division Duplex*) e TDD (*Time Division Duplex*). A Figura 3.2 apresenta uma arquitetura da rede UMTS.

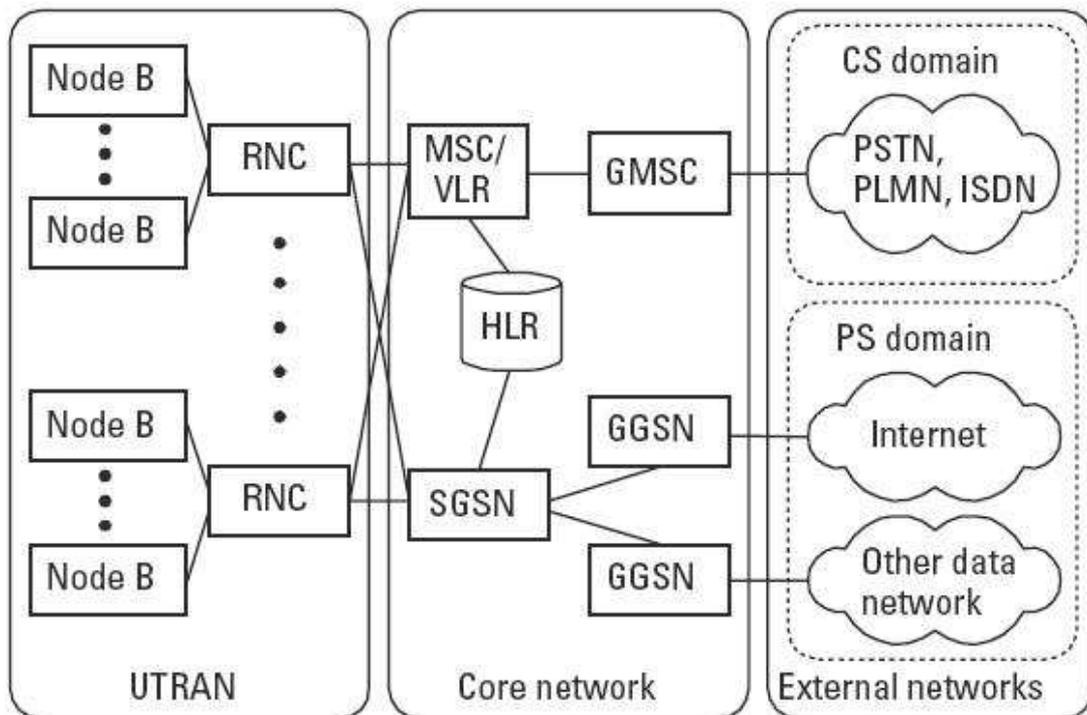


Figura 3.2: Arquitetura da rede UMTS [Janevski 2003].

A arquitetura básica de uma rede UMTS é dividida em equipamentos do usuário e infraestrutura. Os equipamentos do usuário são utilizados para acessar serviços da rede UMTS, isso inclui o módulo de identificação e os equipamentos móveis, dispositivos de hardware e grupos de software.

A infraestrutura da rede é dividida em rede de acesso (*Access Network – NA*) e Núcleo da rede (*Core Network - CN*). Fazem parte do núcleo da rede o módulo de comutação por circuito (*circuit-switched – CS*) e o módulo de comutação por pacotes (*packet-switched – PS*). Estes módulos são sobrepostos no núcleo da rede, sendo que a comutação por pacote é um módulo de operação GSM e a comutação por circuito é um módulo de operação da GPRS. O MSC (*Mobile Switching Center*) e o GMSC (*Gateway MSC*) são entidades específicas do módulo de comutação por circuito da rede GSM.

Outras entidades também fazem parte do módulo de comutação por circuito, porém também são compartilhadas pelo módulo de comutação por pacotes.

As entidades específicas do módulo de comutação por pacotes são somente a GGSN e a SGSN. As demais entidades do núcleo da rede como o HSS (*Home Subscriber Server*), o AuC (*Authentication Center*), o VLR (*Visitor Location Register*), o EIR (*Equipment identity Register*) e o nó de suporte SMS (*Short Message Service*) são comuns aos dois módulos.

O HSS é a base de dados que contém as informações de identificação do usuário, as informações de segurança do usuário (autenticação e autorização), as informações de localização e as informações do perfil do usuário [Janevski 2003]. O HSS fornece também algumas funcionalidades multimídia no núcleo da rede.

A UTRAN (*UMTS Terrestrial Radio Access Network*) é formada pelo RNC (*Radio Network Controller*) e pelas ERB's. Os RNS (*Radio Network System*) são equivalentes as BSS no GSM. Os RNS consiste nos RNC, que controla as ERB's - Estações Rádio Base (também conhecida como *Node B*) . A ERB é um componente da rede que serve uma célula.

3.3.5 WCDMA

O sistema WCDMA (*Wideband CDMA – Code Division Multiple Access*) possui algumas características básicas do sistema CDMA, porém o WCDMA consome, aproximadamente, três vezes mais banda que o sistema CDMA. A taxa especificada para o WCDMA é de 3.84 Mcps (Milhões de chips por segundo) ocupando uma banda de 5 MHz.

O WCDMA opera em dois modos diferentes: o TDD (*Time Division Duplex*) e o FDD (*Frequency Division Duplex*). O modo FDD destina-se a grandes coberturas, a aplicações de banda larga e a aplicações de rádio simétricas, os enlaces de subida e descida são empregados em duas faixas de frequências separadas. O modo TDD fornece altas taxas de dados, destina-se a aplicações assimétricas e de baixa mobilidade, isso inclui os serviços baseados na Internet. As transmissões nos enlaces são feitas na mesma banda, usando intervalos de tempo sincronizado.

Conexões individuais entre os móveis e as estações base são separadas por códigos no sistema WCDMA, enquanto as transmissões são realizadas na mesma banda de frequência. A Figura 3.3 apresenta a arquitetura básica de um sistema WCDMA.

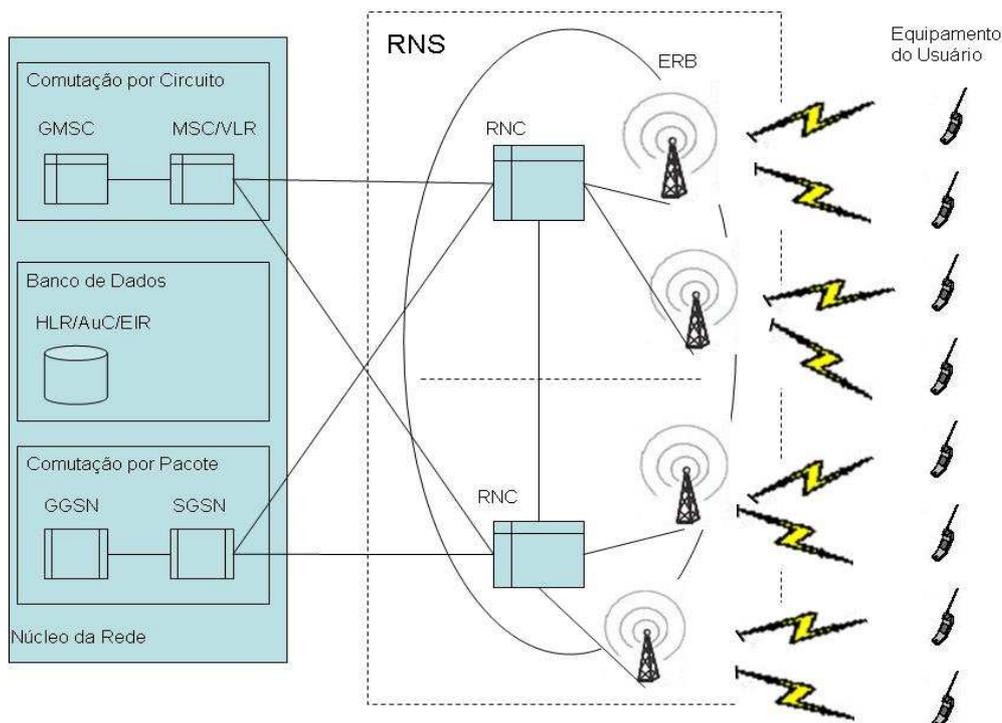


Figura 3.3: Arquitetura básica do sistema WCDMA.

Em um sistema WCDMA a ERB e o RNC formam a UTRAN. O RNC realiza funções de controle e gerenciamento de todos os recursos de rádio do seu domínio. Várias ERB's estão ligadas ao mesmo RNC. O núcleo da rede é formado pela comutação por circuito, comutação por pacotes e pelo banco de dados, onde se encontram o HLR, AuC e EIR.

3.3.6 HSDPA

O HSDPA (*High Speed Downlink Packet Access*) representa um sistema baseado no canal de *downlink* compartilhado, que permite taxa de dados de até 10 Mb/s, projetado em conjunto com o HSDPA (*High Speed Uplink Packet Access*) para suportar serviços que necessitam de altas taxas instantâneas no *downlink* e taxas menores no *uplink*. Nesses sistemas, funções como *Scheduling* e HARQ (*Hybrid Automatic Repeat Request*) foram aplicadas à transmissão na interface aérea.

Este recurso permite um aumento no nível de retransmissão, permitindo a redução no tempo de entrega. Podemos citar a internet e o vídeo sob demanda, como exemplos de serviços usando HSDPA [3GPP 2010].

HSDPA foi criado como uma nova técnica de transmissão em UMTS. Essa tecnologia fornece um melhoramento significativo nos serviços fim-a-fim, baseados em

serviços IP [Schierl *et al* 2007]. O HSDPA permite um aumento significativo na taxa de dados e na capacidade da célula [Hartung *et al* 2007].

O HSDPA suporta alocação de recursos com adaptação de codificação e modulação, para explorar a variação do canal e a variação de interferência, o ARQ híbrido rápido reduz a retransmissão RTT (*Round Trip Times*), a redução do intervalo de tempo de transmissão, para redução de latência e suporte rápido a decisões de agendamento, e um rápido canal de retorno. A Figura 3.4 apresenta a arquitetura HSDPA.

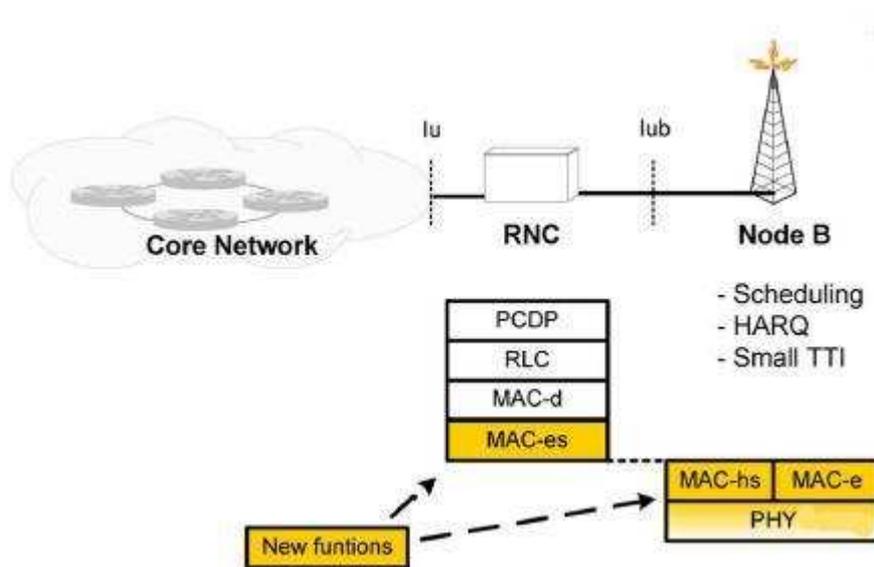


Figura 3.5: Arquitetura da HSDPA [Schierl *et al* 2007].

Funções como *Scheduling*, HARQ e SMALL TTI foram aplicadas a ERB (*Node B*). Uma das grandes vantagens da HSDPA foi transferir algumas funcionalidades do RNC para a ERB, formando uma nova entidade no sistema, o MAC-hs.

O MAC-hs é responsável pelo tratamento do *Scheduling*, HARQ. Algumas atualizações na ERB são necessárias para permitir as funcionalidades do MAC-hs.

Observa-se, assim, uma evolução nos sistemas de comunicações móveis e em redes de comutação de pacotes, em que as redes 3G representam uma evolução dos sistemas de comunicação móvel, sendo totalmente digitais e incluindo transmissão baseadas em pacotes, permitindo que muitas pessoas utilizem seus dispositivos móveis para serviços de dados, vídeo e até TV digital, com base em diferentes requisitos de QoS.

3.4 Aspectos do Padrão 3GPP Relativos a *Streaming* de Vídeo

3.4.1 Introdução

O padrão 3GPP (*3rd Generation Partnership Project*) tem o serviço de *streaming* de pacotes fim-a-fim definido na especificação 26.244 Release 9 v9.2.0 (2010-06). O 3GP (formato de arquivo 3GPP) é usado para mídia contínua ao longo de toda a entrega da sequência prevista pelo MMS (*Multimedia Messaging Service*), independentemente se a entrega final for feita por *streaming* ou arquivo de *download*.

O 3GPP PSS atualmente é o padrão mais maduro nesse campo, e todos os maiores fornecedores de equipamento de telecomunicações móveis bem como fornecedores de equipamento de *streaming* procuram estar em conformidade com ele [Frojdth *et al* 2006].

O 3GPP PSS (*Packet Switched Streaming Service*) fornece toda uma arquitetura para entrega de *streaming* e *download* fim a fim em redes móveis, abrangendo servidor de arquivo, onde ocorre o armazenamento de sessões de *streaming*, servidor de entrega de *streaming*, até o cliente para recepção de *streaming* [Schierl *et al* 2005] [Frojdth *et al* 2006] [Hartung *et al* 2007].

O principal escopo do PSS é definir uma aplicação provendo sincronismo no tempo de mídia. Em particular os seguintes estágios são considerados:

- *Upload* de um terminal de origem para um Proxy MMS;
- Troca de arquivo entre servidores MMS;
- Transferência de conteúdo de mídia recebido de um terminal quer seja de arquivo de *download* ou de *streaming*. No primeiro caso o arquivo é transferido independentemente, considerando que no segundo caso o conteúdo é extraído do arquivo e enviado de acordo com o formato de abertura do cabeçalho.

O 3GPP adota formatos de arquivos de dados de mídia definidos na ISO/IEC 14496-12:2008 | 15444-12:2008. No entanto, a conformidade para os arquivos 3GP, as restrições de endereçamento e as extensões de formato dos arquivos de dados de mídia da ISO, registros de codificadores, arquivo de identificação (extensão de arquivo, identificação de qualidade e tipos MIME) e perfis são definidas em [3GPP TS 2010].

3.4.2 Registros e codificadores

O padrão H.264/AVC está considerado como um dos padrões de codificação utilizados pelo 3GPP de forma a proporcionar economia de banda e uma qualidade de vídeo adequada para diferentes receptores.

No perfil básico o padrão H.264/AVC sofre algumas restrições, entre as quais se encontra restrição no número de faixas de vídeo, isso implica que, em particular, não pode haver faixas alternativas (incluindo troca de faixas) e não pode haver separação de faixas para um conjunto de parâmetros.

O 3GPP fornece uma estrutura para integração entre os diferentes formatos de codificadores, incluindo o H.263, MPEG-4 Visual, H.264/AVC, entretanto o 3GPP não fornece uma estrutura para integração entre o H.264/SVC e os demais codificadores.

3.4.3 Perfil de servidor de *streaming*

No PSS é usado um perfil de servidor de *streaming*, conforme este perfil a interoperabilidade será garantida entre ferramentas de conteúdos criados e servidor de *streaming*. De acordo com o perfil, as seguintes restrições devem ser aplicadas nos arquivos de servidor de *streaming*:

- As faixas RTP indicadas devem ser incluídas para todas as faixas de mídia;
- As faixas RTP indicadas devem cumprir as especificações do PSS;
- As informações SDP devem ser incluídas, onde fragmentos SDP devem ser armazenados em faixas indicadas com nível de controle URLs indicados na faixa;
- As extensões do servidor de *streaming* devem ser usadas para a faixa indicada.

As instruções dadas na faixa RTP indicada devem ser consistente com o PSS. Em particular, tempos de envio de pacotes RTP devem respeitar as restrições de *buffer* e parâmetros usados no SDP.

3.4.4 Perfil de *streaming* adaptativo

O perfil de *streaming* adaptativo no 3GP está registrado como ‘3gh9’. Ele é usado como rótulo de arquivos que são adequados primeiramente para *streaming* adaptativo baseados em arquivos. De acordo com o 3GP, as seguintes restrições devem ser aplicadas para o perfil de *streaming* adaptativo:

- Todos os dados de vídeo devem conter um fragmento de vídeo, não somente amostras;
- Deve conter uma marcação ‘mvex’ na caixa do arquivo para indicar a presença de fragmentos de vídeo;
- Na caixa do arquivo a marcação ‘moov’ deve ser seguida por um ou mais marcações ‘moof’ e opcionalmente o par de caixas do arquivo deve conter a marcação ‘mdat’;
- Cada caixa de arquivo marcada por ‘moof’ deve conter no mínimo uma faixa de fragmento.

3.4.5 Informações de *buffer* de vídeo

Em um arquivo 3GP podem estar incluídos alguns parâmetros de *buffer* de vídeo associados com o vídeo *stream*. Para o caso onde somente um conjunto de parâmetros é associado para todo vídeo *stream*, este pode ser incluído no nível de mídia correspondente ao fragmento SDP. No entanto, dependendo da ordem fornecida de parâmetros de *buffer* para diferentes pontos de operação e para diferentes pontos de sincronização, uma faixa pode conter uma amostra de agrupamento de *buffer*. O tipo de amostra de agrupamento depende do modelo do *buffer* de vídeo que é usado para um determinado codificador de vídeo. Para o H.264/AVC, estes são os dois tipos de *buffers*:

- H.264/AVC Modelo Decodificador de Referência Hipotética (HRD – *Hypothetical Reference Decoder*).
- *Buffer* de desentrelaçamento do modo de pacotização do H.264/AVC de RTP entrelaçado.

3.4.6 Protocolos para *streaming* no 3GPP

Conforme apresentado na Figura 3.6, fazem parte da pilha do 3GPP o protocolo de streaming de tempo real (RTSP- *Real Time Streaming Protocol*), o protocolo de descrição de sessão (SDP – *Session Description Protocol*) para ferramentas de sessão e controle e o protocolo de transferência de tempo real (RTP – *Real Time Protocol*) para transporte de mídia em tempo real, tais como vídeo, voz (fala) e áudio. No esquema adaptativo apresentado neste artigo foram utilizados os protocolos RTP, RTCP e o UDP. Além da definição de mecanismos de transporte, o escopo do PSS também envolve o conjunto de codificadores de mídia.

O PSS também padronizou um mecanismo opcional de proteção de confidencialidade e integridade para o conteúdo por meio de encriptação do *stream* e de arquivos. Outros destaques do padrão incluem seleção de mídia, onde um cliente pode escolher diferentes alternativas de taxas de *bits* e idiomas e reporte de qualidade de experiência (QoE), o qual dá ao fornecedor do serviço medidas para avaliar a experiência do usuário final. A Figura 3.6 apresenta a pilha de protocolos para *streaming* 3GPP.

Video Audio Texto sincronizado com outras mídias	Descrição do cenário Imagens Gráfico bitmap Vetor gráfico Texto	Apresentação da descrição
Formato do Payload	HTTP	RTSP
RTP	TCP	UDP
UDP	IP	UDP

Figura 3.6: Pilha de protocolos utilizada para *streaming* no 3GPP [Frojdth *et al* 2006] [Hartung *et al* 2007].

Além da definição de mecanismos de transporte, o PSS também envolve o conjunto de codificadores de mídia. O PSS inclui várias características, as quais apontam para o melhoramento na qualidade do *streaming* em redes móveis. Uma dessas características é chamada de *streaming* adaptativo. O *streaming* adaptativo permite que os fluxos a serem transmitidos sejam adaptados de acordo com as variações da rede. Embora as redes de melhor esforço possam, eventualmente, fornecer a taxa de *bits* requerida, não há garantia da disponibilidade dessa taxa de *bits* requerida durante o tempo da sessão, especialmente em enlaces móveis, que, geralmente, são caracterizados por uma variação na vazão devido à natureza do canal sem fio [Hartung *et al* 2007].

Uma das características dos *links* móveis é variação da banda devido à natureza do canal sem fio. Diferentes tecnologias de acesso como GPRS, WCDMA, HSDPA colocam limite máximo diferente na média de taxa de *bits* disponíveis, onde *handovers* entre sistemas resultaram em *links* com características diferentes [Hartung *et al* 2007].

O objetivo do *streaming* adaptativo é fornecer sempre entrega de conteúdo em um nível de qualidade que requeira taxa excedente da taxa de transmissão atualmente disponível. Considerando que técnicas de *streaming* adaptativo já foram propostas para *streaming* em internet, as características de link móveis, tais como mudança rápida na variação da banda, atualização de células e *handovers* entre sistemas querem uma solução especial para *streaming* adaptativo em redes móveis.

O *streaming* adaptativo no 3GPP foi projetado para um cenário típico de *streaming* de vídeo sob demanda, em que os dados de mídia são transmitidos e o *feedback* do cliente é recebido, se o *feedback* do cliente não pode ser enviado de volta para a origem da transmissão, outra técnica tem que ser usada. Uma das alternativas é a combinação com codificação escalável.

3.4.7 Recursos de *streaming* adaptativo dentro do PSS

O PSS possui uma coleção de recursos para ajudar no desempenho do *streaming* adaptativo. Os recursos de adaptação são destinados a fornecer a melhor qualidade de conteúdo possível quando livre de interrupções na reprodução sobre o *link* sem fio disponível. Para atingir este objetivo, deve-se evitar o seguinte:

- Transmissão de fluxo de *bit* maior que o *link* disponível é capaz de transportar, causando assim congestão, perda de pacotes que resultarão em redução da qualidade da mídia;
- Por longos períodos de tempo, transmitir conteúdos codificados com taxas de *bits* maiores que as taxas de transmissões de *bits* disponíveis, causando assim *buffer undeflow* no cliente, quando o conteúdo é consumido mais rápido que substituído a partir do servidor.
- Transmitir mais conteúdo para o *buffer* do cliente do que há espaço, causando assim *overflow*, que pode ocorrer quando a taxa de *bit* do conteúdo é menor que a taxa de transmissão de *bit* e o conteúdo é recebido mais rapidamente do que é consumido.

Considerando um *link* sem fio e as suas características, a grande vantagem de se ter um *buffer* completo é que o receptor pode ficar menos sensível a eventos como *handover* ou sombra de rádio, que interrompe temporariamente ou reduz significativamente a taxa de *bit* disponível a partir do servidor para o cliente [Frojdh *et al* 2006].

3.4.8 Sessão de *streaming* adaptativo

Uma sessão de *streaming* adaptativo possui duas fases: A fase de configuração e a fase de *streaming* de mídia. Na fase de configuração, o cliente e o servidor trocam informações usando RTSP e SDP, isso adapta o *streaming* de mídia e permite que o cliente e o servidor negociem os recursos que serão usados [Frojdh *et al* 2006].

A sessão é iniciada pelo cliente com o RTSP, o servidor diz que a conexão do cliente suporta *streaming* adaptativo e que ele pode ser usado. A resposta no servidor inclui um SDP. Com a informação do SDP, o cliente pode configurar o *streaming* de mídia nas sessões de *streaming*.

O pacote de aplicações do PSS, (Pacote de Aplicação de Unidade de Dados - *Next Application Data Unit Application Packet* - NADU APP), possui definido um procedimento de *feedback* único de informações para a solução de *streaming* adaptativo. Usando o *feedback* fornecido no RTSP e mensagens do RTCP, um servidor PSS pode adaptar o fluxo de transmissão e fluxo de conteúdo de acordo com as condições do *link* [Frojdh. *et al* 2006].

4 Protocolos e Adaptação para *Streaming* de Vídeo

4.1 Protocolos Utilizados Neste Trabalho

Nesta seção são apresentados e detalhados os protocolos que serão utilizados nos experimentos realizados neste trabalho, sendo eles o protocolo de transporte UDP, o protocolo de aplicativos em tempo real RTP e o protocolo de controle RTCP, tomando por base o seu emprego em esquema adaptativo anteriormente proposto no âmbito dos trabalhos realizados no Laboratório de Televisão Digital Interativa (LabTVDI/UnB).

4.1.1 Protocolo de transporte UDP

O protocolo UDP (*User Datagram Protocol*) é definido pela RFC 768; ele recebe as mensagens do processo de aplicação, anexa os campos de número de porta da fonte e destino para serviço de multiplexação/demultiplexação, adiciona dois pequenos campos e passa o segmento resultante a camada de rede, que encapsula o segmento dentro de um datagrama IP.

Algumas aplicações se adaptam melhores ao protocolo UDP, devido ao melhor controle no nível de aplicação sobre quais dados são enviados e quando. O protocolo empacotará os dados dentro de um segmento UDP e os passará imediatamente a camada de rede.

O UDP envia mensagens sem o estabelecimento de conexão e assim não introduz nenhum atraso para estabelecer conexões. Além disso, o UDP não mantém o estado de conexão e não monitora buffers de envio e recebimento, parâmetros de controle de congestionamento e parâmetros numéricos de sequência e de reconhecimento. A Figura 4.1 apresenta os campos do cabeçalho UDP, que é formado por 8 *bytes* [RFC 768 1980].

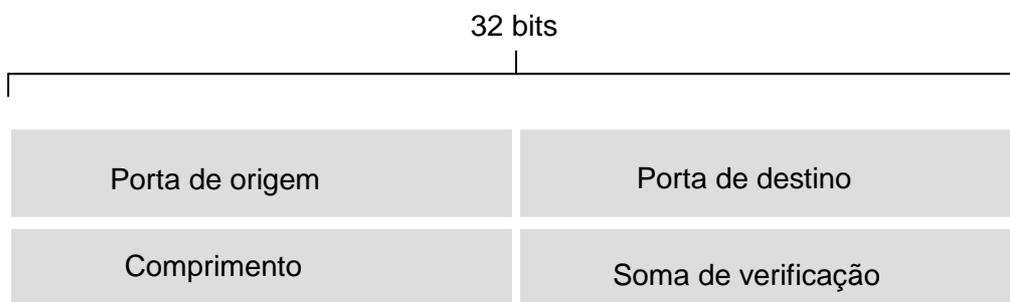


Figura 4.1: Cabeçalho UDP [RFC 768 1980].

Porta de origem: Indica a porta de origem do processo e pode ser considerada a porta de resposta caso seja abordada a ausência de qualquer outra informação. É um campo opcional e quando não utilizado terá valor 0 (zero);

Porta de destino: É a porta onde serão entregues os dados transmitidos, ou seja, a porta de destino;

Comprimento: É o comprimento total do datagrama, incluindo dados e cabeçalho;

Soma de verificação: Destina-se a verificação de erros de transmissão;

4.1.2 Protocolo RTP.

O RTP permite uma gestão dos fluxos multimídia (voz, vídeo) em redes IP. O cabeçalho do RTP comporta informações de sincronização e de numeração.

O RTP e o RTCP são protocolos que se situam em nível de camada de aplicação e utilizam os protocolos subjacentes da camada de transporte UDP ou TCP. Eles podem utilizar tanto o modo de transmissão *unicast* (ponto-a-ponto), quanto *multicast* (multiponto). Ambos permitem, respectivamente, transportar e controlar dados que têm propriedade de “tempo-real”. O RTP utiliza uma porta par e o RTCP uma porta ímpar imediatamente superior [RFC 3550 2003] [Kurose *et al* 2006]. O cabeçalho RTP está representado pela Figura 4.2 e comporta as seguintes informações:

Tipo de carga útil	Número de sequência	Marca de tempo	Identificador de sincronização da fonte	Campos variados
--------------------	---------------------	----------------	---	-----------------

Figura 4.2: Cabeçalho RTP.

Campo tipo de carga útil: Este campo possui 7 (sete) *bits* de comprimento. Em uma transmissão de vídeo, o tipo de carga útil é usado para indicar o tipo de codificação de vídeo (por exemplo, MPEG4, H.264). O remetente pode modificar a codificação durante a sessão;

Campo do número de sequência: Este campo possui comprimento de 16 (dezesseis) bits. A cada pacote enviado, o número de sequência é incrementado e pode ser usado pelo receptor para detectar perda de pacotes e restaurar a sequência de pacotes;

Campo marca de tempo: Este campo possui 32 *bits* de comprimento. Ele aponta o instante do primeiro *byte* no pacote RTP recebido. O receptor pode utilizar o tempo para eliminar a variação de atraso dos pacotes (*jitter*) que acontecem na rede e fornecer uma recepção sincronizada no receptor;

Identificador de sincronização da fonte (SSRS – *Synchronization Source Identifier*):

Este campo possui comprimento de 32 *bits*. O SSRS é um número aleatório atribuído pela fonte quando um novo fluxo é iniciado. Ele serve para identificar a fonte do fluxo RTP.

4.1.3 Protocolo RTCP

O protocolo de controle RTP definido pela RFC 3550 [RFC 3550 2003] é um protocolo baseado no controle periódico de transmissão de pacotes. Sua principal função é fornecer um retorno da qualidade dos dados transmitidos e acompanhamento do controle de congestionamento. Pacotes RTCP são transmitidos por cada participante de uma sessão RTP para todos os outros participantes da sessão.

O pacote RTCP é formado por:

SR: *Sender-Report* – relatório do remetente: Usado para estatísticas de transmissão e recepção dos participantes que estão enviando ativamente.

RR: *Receiver-Report* – relatório do receptor: Usado para estatísticas de recepção.

SDES: *Source description items*- descrição dos itens de origem: Descreve os itens de origem do pacote.

BYE: Indica o fim da participação da sessão.

APP: *Application-specific functions* - funções específicas da aplicação.

A Figura 4.3 apresenta os campos do cabeçalho do RTCP.

V =2	P	RC	PT = RR = 201	Comprimento
Emissor (SSC)				
SSRC_1				
Fração de perda		Número cumulativo de pacotes perdidos		
Número sequencial mais significativo recebido				
Intervalo de jitter				
Último SR (LSR)				
Atraso desde o último SR (DLSR)				
SSRC_2				
Dados RR adicionais				

Figura 4.3: Cabeçalho do RTCP [RFC 3550 2003].

V – Versão: Composta por 2 *bits*, identifica a versão do pacote RTCP;

P – Padding: Composto por 1 *bit*, caso esse *bit* esteja ligado existe algum pacote adicional, que não faz parte da informação de controle;

RC – Reception Count: Número de blocos do RR desde pacote, formado por 5 *bits*;

PT – Packet Type: Formado por 8 *bits*, possui uma constante 201, que identifica o pacote RR;

Comprimento: Formado por 16 *bits*. Contém o tamanho do pacote incluindo cabeçalho e os *padding*s adicionais.

SSRC: Formado por 32 *bits*. Indicador da fonte de sincronização.

SSRC_n: Formado por 32 *bits*.

Fração de perda – Fraction Lost: Formado por 8 *bits*. É definida pelo número de pacotes perdidos dividido pelo número de pacotes esperados.

Número cumulativos de pacotes perdidos: Formado por 24 *bits*. Total de pacotes perdidos desde o início da recepção. É calculado como sendo a diferença entre o número de pacotes esperados e o número de pacotes recebidos.

Número seqüencial mais significativo recebido: Formado por 32 *bits*, sendo que 16 *bits* são menos significativos e contém o maior número de pacotes RTP da origem, os 16 mais significativos estendem o número da seqüência de pacotes RTP e correspondem na contagem de ciclos.

Intervalo de jitter entre chegadas: Formado por 32 *bits*. É uma estimativa do *jitter* dos pacotes de dados RTP.

Timestamp do último pacote SR - Last SR timestamp (LSR): Formado por 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): Formado por 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.

4.2 Esquema adaptativo para *streaming* de vídeo SVC

O esquema adaptativo proposto em [Mulinari e Gondim 2009], foi implementado e comparado com outros dois esquemas ADIVIS [Antoniou *et al* 2007] e “Bouras e Gkamas” [Bouras and Gkamas 2005] com base, inicialmente, em redes WLAN. Os

vídeos utilizados foram codificados no padrão H.264/SVC e o modelo de transmissão foi *unicast*. O mesmo esquema foi utilizado para os experimentos realizados em redes de 3ª Geração neste artigo e será detalhado nesta seção

O esquema baseia-se na incorporação da lógica *fuzzy* e em *feedbacks* do cliente com informações das condições da rede para tomar decisões de ajuste na taxa de bits do vídeo transmitido. O conteúdo de vídeo e os pacotes de controles com os *feedbacks* do cliente para o servidor são transmitidos via protocolos RTP / RTCP. O papel do mecanismo de *feedback* e dos componentes de adaptação é o de associar a qualidade de vídeo desejada às condições de rede. A Figura 4.4 apresenta a arquitetura do esquema adaptativo.

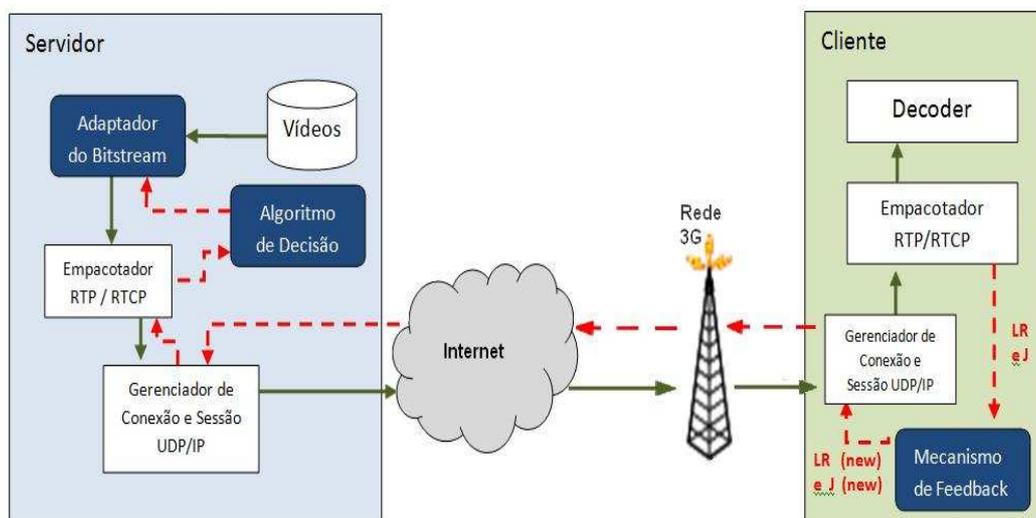


Figura 4.4: Arquitetura do esquema adaptativo.

O mecanismo para processamento da adaptação de rede utiliza um algoritmo *fuzzy* para decisão. O mecanismo de *feedback*, para sensoriamento, combina informações de perda de pacotes e *jitter* enviadas pelo receptor sobre a qualidade percebida, a fim de avaliar a situação atual da rede.

Os valores das variáveis que indicam a qualidade percebida na rede são introduzidos no algoritmo de decisão, que decide com base na melhora ou piora das condições de rede, conforme segue abaixo:

- Taxa de perda (LR- *Loss Rate*): O primeiro valor de *feedback*, a taxa de perda de pacotes, é calculado pela variação da medição atual de perda de pacote e a medição anterior LR;
- Atraso médio: O segundo parâmetro de *feedback* é calculado utilizando o valor

gerado pelo *interval jitter* do pacote RR (*Receiver-Report*);

- Os dois parâmetros coletados pelo mecanismo de *feedback* são enviados para o servidor em pacotes do tipo RR do protocolo RTCP.

4.3 Componentes do esquema adaptativo

4.3.1 Componentes do Servidor

O servidor é composto de:

- Adaptador de *bitstream*: responsável pela adaptação de conteúdo, incluindo o gerenciamento da quantidade de camadas de vídeo que devem ser enviadas a partir da banda disponível;
- Algoritmo de decisão: recebe os *feedbacks* dos clientes (receptores), enviados em pacotes RTCP e decide pela manutenção, diminuição ou incremento da banda disponível;
- Empacotador RTP/UDP (*User Datagram Protocol*) / Desempacotador RTCP/UDP: encapsula os pacotes de vídeo, ou unidade NAL em pacotes no protocolo RTP/UDP; desencapsula os pacotes RTCP/UDP recebidos pelo gerenciador de conexão e sessão e encaminha para o algoritmo de decisão;
- Gerenciador de conexão e sessão: gerencia as conexões e sessões e encaminha os pacotes RTP/UDP e RTCP/UDP.

4.3.2 Componentes do Cliente

O cliente é composto de:

- Gerenciador de conexão e sessão: solicita a abertura de sessão com o servidor, envia pacotes de controle e recebe pacotes de vídeo;
- Mecanismo de *feedback*: recebe os pacotes e faz o sensoriamento das condições da rede. Contabiliza os pacotes perdidos e o atraso dos pacotes durante o intervalo medido;
- Empacotador RTCP/Desempacotador RTP: encapsula os pacotes de controle e encaminhá-lo ao gerenciador de conexão e sessão; desencapsular os pacotes recebidos e armazenar os pacotes.

O intervalo de medição para computar e gerar um relatório com os valores coletados das condições de rede (*feedback*) é de 0,5 segundo e o envio de pacotes RRs é de 1 pacote a cada 0,5 segundo. O mecanismo de decisão informa ao adaptador de

bitstream a nova largura de banda disponível e identifica as camadas que podem ser transmitidas na largura de banda disponível. O esquema faz uso de dois arquivos extraídos com a ajuda do aplicativo JSVM . O arquivo vídeo.bitrate é utilizado para informar quais camadas podem ser transportadas com a banda disponível. Ele identifica a taxa de bits necessária para transportar cada camada. O arquivo de *trace* é utilizado para identificar os pacotes NAL (*Network Adaptation Layer*) que pertencem à camada escolhida ou a camadas inferiores.

4.3.3 Algoritmo de decisão

O algoritmo de decisão recebe os *feedbacks* dos clientes, enviados pelo(s) receptor(es) em pacotes RTCP (RRs), processa-os pela lógica do esquema adaptativo e decide pela manutenção, diminuição ou incremento da banda disponível. Após esta decisão, repassa ao adaptador de *bitstream* a nova banda disponível e armazena, para avaliação no arquivo de log, os valores de *feedback* recebidos dos clientes e a nova banda disponível [Mulinari e Gondim 2009].

5 Aspectos relacionados a QoS, QoV e QoE

5.1 Métricas de qualidade de vídeo

A avaliação da qualidade do vídeo é um aspecto muito importante em codificação de vídeo. Essa avaliação pode ser feita objetivamente e subjetivamente. Métodos de teste subjetivos requerem uma perspectiva humana para avaliar a qualidade ou diferenças de qualidade entre duas imagens ou vídeos.

A avaliação subjetiva pode ser um processo caro e demorado, além do custo esse método não pode ser usado para monitorar qualidade de imagem de vídeo em tempo real. Esse processo requer equipamentos especiais e muitas pessoas. Teste subjetivo de qualidade de vídeo avalia como um vídeo é percebido pelo espectador e designa a opinião do espectador em uma sequência de vídeo particular [Wang 2006], [Rojagopalan 2010].

O método de teste objetivo não depende da avaliação humana, e sim de avaliar analiticamente o sinal do vídeo. Esse método possui um algoritmo que mensura a qualidade da imagem, baseado na comparação da imagem de origem com a imagem processada. Em muitas situações alguns métodos objetivos podem substituir com sucesso o uso de avaliação subjetiva.

Várias métricas de qualidade objetiva foram desenvolvidas para mensurar a qualidade da imagem, inclusive sem referencia do vídeo original. As métricas são caracterizadas de acordo com a quantidade de informações requeridas.

Full-reference: Requer todo o vídeo de referencia e pode não ser muito prático para aplicações em tempo real.

Reduced – reference: Usa uma versão reduzida do vídeo de referencia, enquanto métricas no-referece não utilizam referencias do vídeo original. A diferença entre a imagem de referencia e a imagem codificada fornece a métrica objetiva [Punchihewa *et al* 2003].

No Reference: Usa apenas o vídeo recebido para determinar a qualidade, dispensando assim, o vídeo original.

5.1.1 PSNR

O PSNR (*Peak-Signal-to-Noise-Ratio*) é uma métrica de avaliação de qualidade simples baseada no erro médio entre o vídeo original e o vídeo processado. O PSNR é derivado do MSE (*Mean Squared Error*) em relação ao máximo valor possível de luminância

[Wang 2006] [Changhoon and Bovik 2011]. A equação 5.1 representa o MSE e a equação 5.2 representa o PSNR:

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [f(i, j) - F(i, j)]^2}{M \cdot N} \quad (5.1)$$

$$PSNR = 20 \cdot \log_{10} \left(\frac{255}{\sqrt{MSE}} \right) \quad (5.2)$$

Onde $f(i, j)$ é o sinal original do pixel (i, j) , $F(i, j)$ é o sinal reconstruído e $M \times N$ é o tamanho da imagem. O resultado é um único número em *decibéis*. O PSNR considera somente componente de luminância e deixa de lado os componentes de crominância, que é importante para a percepção humana [Van der Auwera *et al.* 2008.]. Apesar de existirem várias métricas de qualidade objetiva, PSNR continua sendo a métrica mais popular.

5.1.2 VQM

O VQM foi desenvolvido pelo ITS (*The Institute for Telecommunication Science*) para fornecer uma avaliação objetiva da qualidade do vídeo. A avaliação do percentual de defeitos do vídeo inclui borrão, movimento brusco, ruído global, blocos distorcidos e cores distorcidas, combinados dentro de uma única métrica.

Resultados de testes mostram que o VQM possui alta correlação com a avaliação de qualidade de vídeo subjetiva e tem sido adotada pelo ANSI como um padrão de qualidade de vídeo objetiva [Wang 2006].

O VQM seleciona o vídeo original e o vídeo processado e o cálculo é feito considerando os seguintes processos:

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.

Extração de Recursos de Qualidade: Neste processo é extraído um conjunto de recursos de qualidade que caracterizam o percentual de mudanças espacial, temporal e de crominância nas propriedades das sub-regiões do vídeo usando funções matemáticas.

Parâmetros do cálculo de qualidade: Neste processo são calculados os parâmetros que descrevem o percentual das mudanças do vídeo comparando os parâmetros extraídos do vídeo processado com aqueles do vídeo original.

Cálculo do VQM: O VQM é calculado usando uma combinação linear de parâmetros calculados no processo anterior.

O modelo geral utiliza uma combinação linear de sete parâmetros, Quatro parâmetros são baseados em recursos de extração do gradiente do componente de luminância Y, dois parâmetros são baseados em recursos de extração do vetor formato e dois componentes de crominância (CB, CR) e um parâmetro baseado em contraste e recursos de informações temporais, ambos extraídos do componente de luminância Y.

O VQM também pode ser calculado usando vários modelos baseados em alguns critérios, estes modelos incluem televisão, Videoconferência, modelo geral.

Os valores do VQM variam de 0 até 1, onde 0 corresponde a correlação máxima entre o vídeo original e o processado e 1 corresponde a correlação mínima.

5.2 Métricas de qualidade de serviço

5.2.1 Atraso

Quando um pacote é transmitido, ao sair de sua origem até chegar e ao seu destino, ele passa por uma série de nós (roteadores), isso faz com que ele sofra diferentes tipos de atrasos ao longo do caminho. Dentre os atrasos sofridos tem-se o atraso de processamento, de fila, de transmissão e de propagação. O acúmulo de todos os tipos de atraso ao longo do caminho resulta no atraso fim- a- fim, que nada mais é que a soma de todos os tipos de atrasos envolvidos na transmissão.

O atraso de processamento está caracterizado pelo tempo necessário para verificação de erros de *bits* existentes no pacote, o tempo requerido para a verificação do cabeçalho e determinação do direcionamento do pacote.

O atraso de fila ocorrerá quando o pacote espera em *buffer* para ser transmitido no enlace. O atraso na fila dependerá da quantidade de outros pacotes que se encontram na fila esperando para serem transmitidos. Caso não haja nenhum pacote esperando na fila e nenhum pacote sendo transmitido naquele determinado momento, o atraso de fila será considerado zero.

O atraso de transmissão dependerá do tamanho do pacote em *bits* e a velocidade na qual será transmitido esse pacote do enlace ao roteador. Se o tamanho do pacote for

dividido pela velocidade de transmissão tem-se a quantidade de tempo necessária para transmitir todos os *bits* do pacote para o enlace.

O atraso de propagação de um enlace é a distância, entre dois roteadores, dividida pela velocidade de propagação. O tempo necessário para propagação do *bit* que vai desde o início do enlace até o roteador. A velocidade de propagação do *bit* é igual à velocidade de propagação do enlace, que pode ser igual à velocidade da luz.

5.2.2 Jitter

O *jitter* pode ser dado como a variação do atraso entre pacotes, decorrente de 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.

5.2.3 Perda de pacotes

A perda de pacotes é um dos indicadores de qualidade na transmissão de pacotes em redes IP e tem influência na qualidade de serviço, podendo ocorrer pelo estouro de *buffers* em roteadores e *switches*, 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

5.3 Qualidade de Experiência

Vários padrões de QoE (Qualidade de Experiência) têm sido propostos no ITU ou estão em desenvolvimento [Kuipers *et al* 2010]. QoE refere-se à percepção de aplicativos e serviços pelo usuário final. Muitos fatores podem influenciar nessa percepção, isso inclui o próprio usuário, o terminal, a rede, a infra-estrutura do serviço, entre outros. A QoE é idealmente avaliada com emprego de métrica subjetiva, com isso, os testes feitos para mensurá-la deveriam ser aplicados utilizando usuários reais, no entanto, esse processo é demorado e possui um custo elevado.

Os parâmetros que definem QoE podem ser diferenciados dependendo do serviço, isso faz com que o conceito de QoE seja mais que apenas métricas de QoV ou de QoS. A QoE depende de fatores tais como:

1. Qualidade do vídeo (QoV);

2. Qualidade de Serviço (QoS), que envolve a utilização da infraestrutura da rede para o serviço de "streaming" para fins de entrega do conteúdo (vídeo);
3. Percepção do usuário, que inclui ambiente, expectativas entre outros.

A qualidade do conteúdo está relacionada com o tipo de codificação usada, parâmetros de QoS que afetam o serviço de *streaming* como banda, atraso, *jitter* e perda de pacotes.

QoE pode ser avaliada com base em uma escala MOS (Mean Opinion Score), que envolve cinco valores possíveis (ITU-T P 800): 5 = excelente, 4 = bom, 3 = razoável (ou apropriada), 2 = pobre, 1 = ruim.

5.4 Ferramentas para aferição da qualidade de vídeo e redes e codificação de vídeo

5.4.1 JSVM

O JSVM (*Joint Scalable Video Model*) é um *software* de referência para codificação de vídeo escalável projetado pela JVT (*Joint Video Team*), pela ISO/IEC MPEG (*Moving Pictures Experts Group*) e pela ITU-T *Video Coding Experts Group* (VCEG), disponível em [JSVM 2010].

Ele é escrito em C++ e seu código fonte é aberto, além disso, no pacote de software do JSVM podem ser encontrados bibliotecas e arquivos executáveis.

As bibliotecas fornecidas pelo JSVM são:

H264AVCCommonLibStatic: Essa biblioteca fornece classes que são usadas tanto pelo codificador, quanto pelo decodificador.

H264AVCEncoderLibStatic: Essa biblioteca fornece classes que são usadas somente pelo codificador. Isso inclui as classes para estimativa de movimento, modo de decisão, e codificação de entropia.

H264AVCDecoderLibStatic: Essa biblioteca possui classes que são usadas somente pelo decodificador. Isso inclui as classes para decodificação de entropia e análise de *bit-stream*.

AvcRewriterLibStatic: Essa biblioteca compartilha os mesmos arquivos contidos em *H264AVCDecoderLibStatic*, mas compilado com um compilador definido, ***SHARP_AVC_REWRITE_OUTPUT***:. Ele fornece bibliotecas que são usadas somente pelo *AvcRewriter*.

H264AVCVideoIoLibStatic: Essa biblioteca fornece classes para leitura e escrita nas unidades NAL no formato de *byte-stream* bem como classes para leitura e escrita de dados brutos de vídeo.

Os arquivos executáveis fornecidos pelo JSVM são:

DownConvertStatic: Pode ser usado para aumentar ou diminuir a amostragem da sequência de vídeo.

H264AVCEncoderLibTestStatic: Pode ser usado para gerar camadas únicas AVC ou *bit-stream* SVC.

H264AVCDecoderLibTestStatic: Pode ser usado para decodificar AVC ou *bit-stream* SVC e reconstruir sequência bruta do vídeo.

AvcRewriter: pode ser usado para reescrever um *bit-stream* SVC para AVC quando o *avc_rewrite_flag* estiver presente na camada de reforço.

BitStreamExtractorStatic: Pode ser usado para extração de *sub-bitstream* com uma baixa resolução temporal/espacial e/ou taxas de *bits* do *bit-stream* escalável global.

QualityLevelAssignerStatic: Pode ser usado para geração de *bit-stream* com informações adicionais de qualidade na camada dado um determinado *bit-stream* escalável. Além disso, as informações adicionais de qualidade tanto de entrada quanto de saída são idênticas.

MCTFPreProcessor: Esta ferramenta é utilizada para pré filtragem das sequências de imagens.

PSNRStatic: Pode ser usado para mensurar o PSNR entre duas sequências brutas de vídeo. Além disso, pode ser usado para mensurar o fluxo de *bits* dado um *bit-stream*.

FixedQPEncoderStatic: Pode ser usado controlando a codificação e ajustando o fluxo de bits do *bit-stream* gerado.

SIPAnalyser: Pode ser usado para fazer a predição seletiva entre as camadas.

YUVCompareStatic: Pode ser usada para encontrar incompatibilidade na codificação/decodificação.

5.4.2 BVQM

O Instituto de Ciências e Telecomunicações (ITS - *The Institute for Telecommunication Sciences*) desenvolveu uma ferramenta de software de métrica de qualidade de vídeo chamada de (BVQM – *Batch Video Quality Metric*) para avaliar objetivamente a qualidade do vídeo , disponível em [BVQM 2010].

O BVQM fornece uma interface gráfica para o usuário selecionar os vídeos que serão processados em lote. O processo consiste em calibrar os vídeos processados e calcular a qualidade associada ao vídeo.

Os valores de calibragem podem ser definidos manualmente, ou pode ser feita escolhendo algumas das opções de calibragem automática. Algumas opções de modelo de qualidade de vídeo também estão disponíveis na ferramenta. Após o processamento, o BVQM fornece os resultados e gera relatórios em arquivos de texto. Todos os resultados podem ser exportados para uma planilha eletrônica.

5.4.3 FFMPEG

É uma ferramenta de conversão de áudio e vídeo não-gráfico que suporta diversos formatos de multimídia, através do comando **ffmpeg –formats** é possível se ter uma idéia de sua flexibilidade. Ele suporta uma imensa variedade de parâmetros, e uma lista completa de todos eles está disponível no manual do FFMPEG [Video Lan 2010].

5.4.4 VLC

O VLC é um *player* de mídia aberto e que suporta diversos formatos de codificadores. Ele pode ser utilizado como servidor de *streaming*, tanto de forma *unicast* quanto *multicast*. Ele toca *streaming* e fornece a funcionalidade de codificação/decodificação do *streaming* [Video Lan 2010].

Essa ferramenta foi utilizada em um dos cenários propostos neste trabalho para a transmissão e recepção de *streaming* codificado no padrão MPEG 4 - parte 2.

5.5 Formatos de vídeos

Para compressão de vídeo, pode ser considerada uma ampla variedade de formatos de vídeo. Os formatos mais comuns utilizados em terminais móveis são o QCIF (*Quarter*

Common Intermediate Format), que corresponde a uma resolução de 176 x 144 (Horizontal x vertical), CIF (*Common Intermediate Format*), que corresponde a uma resolução de 352 x 288.

O 4CIF (4xCIF), que corresponde a uma resolução de 704 x 576, é o formato padrão utilizado para televisão e vídeo DVD. Já os formatos CIF são utilizados para aplicações de vídeo conferência e aplicações multimídia, onde a resolução do display e a taxa de *bit* são limitadas.

5.5.1 Sequências de vídeo

Os vídeos utilizados foram codificados no formato H.264/SVC, com escalabilidade temporal, SNR e espacial, com tamanho do GOP 16, refinamento MGS, camada base em modo AVC, e no formato MPEG - 4, que foram armazenados no servidor antes do início do experimento.

A Figura 5.1 apresenta os primeiros quadros de cada vídeo utilizado no cenário de teste, enquanto a Tabela 5.1 apresenta os parâmetros de configuração dos vídeos codificados no padrão H.264/SVC. Os arquivos de configuração para codificação dos vídeos encontram-se no Apêndice II.



Figura 5.1: Quadros das sequências de vídeos utilizadas.

Parâmetros de configuração do vídeo	
Taxa de quadros	30
Nº de quadros a serem codificados	300
Tamanho GOP	16
Período Intra	16
Número de quadros de referência	1
Modo da camada Base	Compatível com AVC

Tabela 5.1: Parâmetros de configuração do vídeo.

5.5.2 Classificação das Sequências

Segundo [Ries *et al* 2008] foram definidas 5 classes de conteúdos utilizadas com mais frequência com diferentes impactos na percepção do usuário:

Classe 1: Notícias: Essa classe de conteúdo possui sequência com poucos movimentos de regiões de interesse e um fundo estático. As principais regiões de interesse são detectadas pelos movimentos dos olhos, mãos e face. A sequência de vídeo *Paris* pode ser citada como exemplo.

Classe 2: Futebol: Caracterizada por movimentos rápidos de objetos (bola e jogadores), plano de fundo com cores uniformes, na maioria das vezes verde. A sequência de vídeo *Football* pode ser citada como exemplo.

Classe 3: Desenhos: A movimentação dos objetos é dominante e o plano de fundo é geralmente estático. Os objetos não se movem naturalmente.

Classe 4: Paisagem: Movimento uniforme e única direção. As sequências *Flower* e *Waterfall* podem ser citadas como exemplo.

Classe 5: Outros: Possui uma grande quantidade de movimento global ou mudanças rápidas de cenas.

O *Video Quality Experts Group* –VQEG é um grupo de especialistas com vários *backgrounds* e afiliações, formado desde 1997, possuindo vários participantes reconhecidos internacionalmente trabalhando na avaliação da qualidade de vídeo.

Nos testes realizados pelo VQEG para avaliação da qualidade do vídeo, foram avaliadas algumas amostras de sequência de vídeo conforme alguns requisitos, dentre eles:

- Uma sequência na qual as cores são enfatizadas;
- Uma sequência sem muitos movimentos;
- Uma sequência na qual a luminância é enfatizada;

- Várias sequências de filme;
- Várias sequências contendo cenas curtas;
- Várias sequências contendo movimentos e detalhes espaciais;
- Uma sequência contendo texto.

Para o nosso trabalho foram adotadas as seguintes classificações, baseadas no VQEG [JSVM 2010]:

- Pelo menos uma sequência de paisagem.
- Pelo menos uma sequência com movimento.
- Pelo menos uma sequência característica da classe notícia.
- Pelo menos uma sequência característica da classe futebol.

6 Trabalhos Relacionados e Descrição dos Experimentos

6.1 Trabalhos Relacionados

A avaliação de *streaming* e aplicações multimídia tem se tornado um grande foco de pesquisa nos últimos anos, dessa forma alguns trabalhos podem ser encontrados na literatura tratando diretamente da avaliação de desempenho de *streaming* escalável utilizando diferentes tipos de escalabilidade em redes de 3ª geração.

No artigo [Alexiou *et al* 2005] foi projetado e implementado um esquema de decisão com *feedback*, baseado no RTP (*Real-Time Protocol*) e utilizando vídeo *streaming* sobre um sistema UMTS (*Universal Mobile Telecommunication System*), para sequências de vídeo codificadas em H.264/AVC. O artigo apresenta um esquema de decisão baseado em análises de *feedback*, que adapta a taxa de transmissão da sequência de vídeo de acordo com a banda disponível; a rede UMTS foi simulada e foi usado o UDP (*User Datagram Protocol*) para transmissão de dados multimídia e o TCP (*Transmission Control Protocol*) para transmissão de informações de controle. O método usado para avaliação da QoV foi baseado em PSNR.

No artigo [Singh, V. *et al* 2009] foram verificados diferentes mecanismos de adaptação de fluxo e redefinidos para redes 3G fazendo uso dos protocolos RTP, o RTCP e o TFRC (*TCP Friendly Rate Control*). O ambiente experimental foi simulado, onde foram transmitidas sequências de vídeo codificadas no padrão H.264/AVC em uma rede 3G utilizando dispositivos fixos. Para avaliação da QoV foi utilizada a métrica PSNR.

Em [Davies *et al* 2008] o método utilizado para avaliação da QoV foi MOS (*Mean Opinion Score*), foram utilizados dispositivos móveis e sequências de vídeo codificadas no padrão MPEG-4 parte 2. Para o estudo de caso foi feita uma simulação da rede 3G, utilizando comutação por pacotes e comutação por circuito. As sequências de vídeo foram recebidas utilizando um dispositivo móvel.

Em [Liu *et al* 2009] foi feita uma avaliação do *streaming* codificado no padrão H.264/SVC. Tanto a transmissão do *streaming*, quanto a rede GSM, WCDMA foram simuladas. A transmissão dos pacotes foi feita no formato RTP e não há um controle de fluxo nem *feedback* de entrega desses pacotes. A métrica utilizada para a avaliação do

streaming foi o PSNR. Para a recepção do *streaming*, os autores não especificaram os tipos de dispositivos.

Em [Diaz, A. *et al* 2007] foi feita uma análise de desempenho de serviço de *streaming* de vídeo, com o objetivo de fornecer um ambiente usando um serviço móvel real. Foram utilizados vídeos codificados em H.263 e MPEG-4 visual. Dois cenários foram propostos para o experimento utilizando rede GPRS (*General Packet Radio Service*) e rede UMTS, o primeiro em um ambiente onde o cliente é estático e o segundo em um ambiente onde o cliente está se movimentando em um veículo. A métrica utilizada para avaliação de do *streaming* foi PSNR.

No artigo [Fajardo *et al* 2009] a avaliação do *streaming* foi feita utilizando o MOS. A sequência de vídeo foi codificada no padrão H.264/AVC. O ambiente proposto para o teste, e a rede UMTS foram simulados. A transmissão dos pacotes foi feita no formato RTP, e o protocolo UDP também foi utilizado para transmissão da sequência de vídeo. A recepção do *streaming* foi feita utilizando dispositivos móveis.

6.2 Descrição dos Experimentos

6.2.1 Cenário proposto

O cenário proposto para os testes e a avaliação da qualidade do vídeo é composto por uma arquitetura cliente servidor. No servidor encontram-se os vídeos armazenados e no formato original YUV, o “aplicativo servidor” do esquema adaptativo (implementado na linguagem Java) e as informações de variação de atraso (*jitter*) e perda de pacotes que ocorrem durante a transmissão da sequência de vídeo. A Figura 6.1 ilustra o cenário utilizado para realização dos testes.

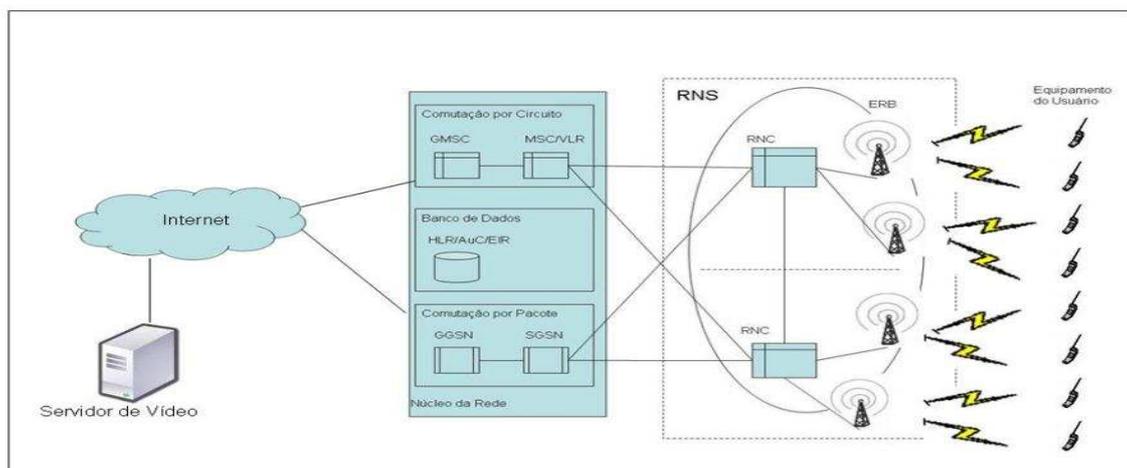


Figura 6.1: Cenário proposto para os experimentos.

O servidor é composto por um *desktop* configurado com o sistema operacional Windows 7, conectado à Internet. Nele está armazenado o aplicativo servidor do esquema adaptativo (implementado na linguagem Java) que recebe a mensagem de retorno (*feedback*) do cliente e faz a adaptação do conteúdo de acordo com as condições da rede. Nele também estão armazenadas todas as informações de variação de atraso (*jitter*) e perda de pacotes que ocorrem durante a transmissão da sequência de vídeo.

O cliente é composto por um *Notebook* marca HP, modelo Pavilion tx1070br, com processador AMD Turion 64, 1024 Kb de memória RAM e sistema operacional Windows Vista. Para ter acesso à rede 3G, o cliente está conectado via USB 1.0 a um modem ZTE modelo mf110, com acesso às redes GSM/GPRS/EDGE (850/900MHz), e HSUPA/HSDPA/UMTS (850/1.900/2.100 MHz). No cliente está armazenado o aplicativo cliente do esquema adaptativo, implementado na linguagem Java e é através dele que são feitas as requisições e a recepção do vídeo.

Os testes foram realizados utilizando a rede 3G de uma operadora de telecomunicações, considerando que a velocidade máxima de navegação é de 1Mbps. Ao atingir 100% (cem por cento) da franquia contratada, o cliente permanecerá conectado, porém, com velocidade máxima de 128 Kbps.

A velocidade de navegação também pode ser diminuída dependendo das condições topográficas e de relevo, velocidade de movimento e distância do cliente em relação à Estação Rádio Base (ERB), número de usuários que utilizam ao mesmo tempo a cobertura fornecida pela ERB, disponibilidade e intensidade do sinal no local em que se utiliza o serviço, condições climáticas e razões técnicas de cobertura móvel, interferência da transmissão do sinal da ERB na Estação Móvel. O endereço IP atribuído à rede 3G não é fixo, sendo fornecido um novo endereço a cada nova conexão. A rede alocada pela operadora oscila entre as redes HSPA, UMTS, GSM e EDGE. Os testes realizados neste trabalho foram feitos considerando estações fixas.

Neste trabalho foram utilizados vídeos codificados em H.264/SVC com diferentes combinações de camadas. Ao contrário do que foi utilizado em [Fajardo *et al* 2009], aqui, utilizaremos um ambiente real, com tráfego real na rede de uma operadora de telecomunicações móveis.

Considera-se, também, o seguinte mapeamento entre os valores de PSNR e a escala relativa à avaliação subjetiva (MOS – *Mean Opinion Score*), obtido de [Migliorini *et al* 2011] [67] e apresentado na Tabela 6.1:

PSNR[dB]	MOS	Grau deficiência da Imagem
> 37	5 (Excelente)	Imperceptível
31-37	4 (Boa)	Perceptível mas não incômodo
25-31	3 (Apropriada)	Pouco Incômodo
20-25	2 (Pobre)	Incômodo
< 20	1 (Ruim)	Muito Incômodo

Tabela 6.1: Relação entre PSNR, MOS e grau de deficiência.

6.2.2 Arquitetura

A arquitetura para o processo de avaliação está representada na Figura 6.2, que consiste das seguintes fases: Codificação, processamento e transmissão, recepção, decodificação e avaliação (aferição do PSNR e VQM).

Na fase de codificação o vídeo original que se encontra no formato bruto YUV é codificado no padrão H.264/SVC e com a ajuda da ferramenta JSVM (*Joint Scalable Video Model*) é produzido um arquivo de *trace* e um arquivo *bitrate*. O mesmo processo ocorre na fase de codificação no padrão MPEG-4, porém a ferramenta FFMpeg é utilizada para codificar o vídeo em MPEG-4. Após passar pelo processo de codificação, o vídeo no formato H.264/SVC é processado e transmitido via rede 3G. Nessa etapa ocorre uma divisão entre as sequências de vídeo que serão transmitidas com a ajuda do esquema adaptativo e as sequências de vídeo que serão transmitidas com a ajuda da ferramenta MPlayer modificada. Um servidor de vídeo Apache (Apêndice III) é configurado para armazenamento das sequências que serão transmitidas com a ajuda da ferramenta MPlayer. Para a transmissão das sequências de vídeo no formato MPEG-4 é usada a ferramenta VLC (*Video Lan Media Player*).

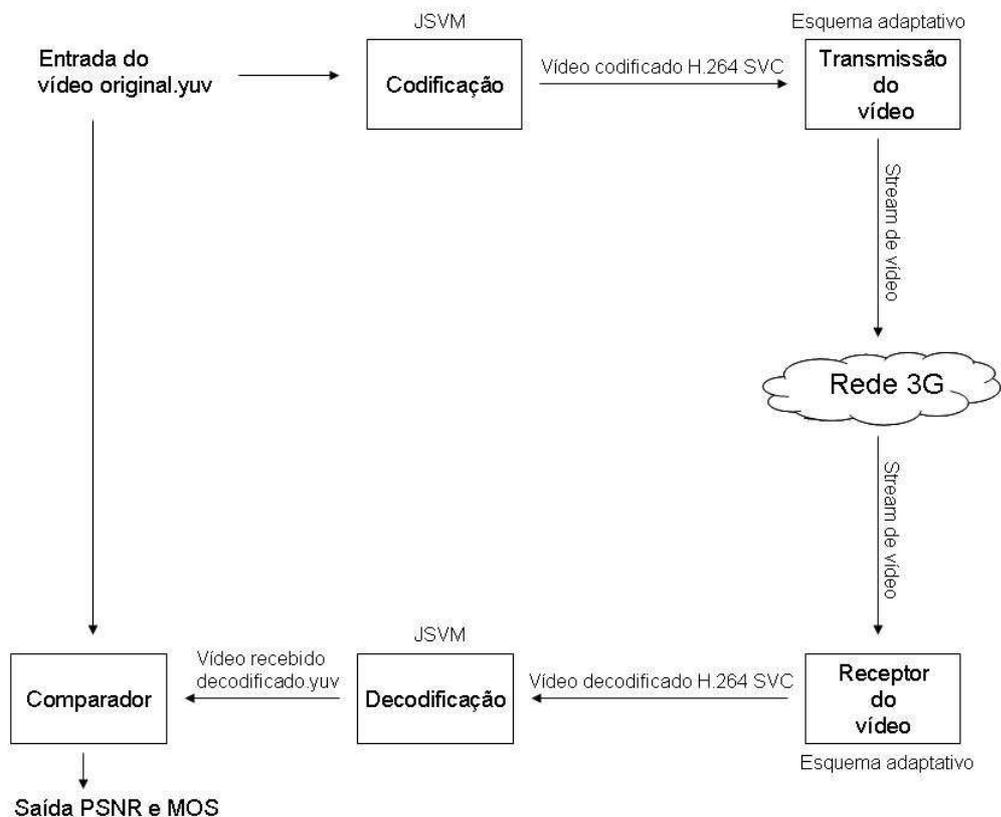


Figura 6.2: Fluxograma da arquitetura de avaliação

Na fase de recepção e decodificação, as sequências de vídeo são tratadas de formas semelhantes ao processo de codificação. No caso das sequências de vídeo codificadas em H.264/SVC a recepção é feita pelo esquema adaptativo implementado no lado cliente e pelo MPlayer configurado no receptor. Já para as sequências de vídeo codificadas em MPEG-4 a recepção é feita pelo VLC. A decodificação é feita utilizando a ferramenta JSVM, para sequências de vídeo codificadas em H.264/SVC, e decodificadas pela ferramenta FFMpeg para sequências de vídeo codificadas em MPEG-4.

A aferição do PSNR é feita após o processo de decodificação, onde o vídeo recebido é comparado com o original no formato bruto YUV.

7 Análise dos resultados

7.1 Introdução

Neste capítulo são apresentados resultados relativos às 700 (setecentas) transmissões realizadas, considerando duas métricas objetivas:

- Y-PSNR;
- VQM

São consideradas 3 (três) alternativas em termos de codificação e emprego de esquema adaptativo:

- MPEG-4 parte 2;
- H.264/SVC com o uso de esquema adaptativo;
- H.264/SVC sem o uso de esquema adaptativo.

Conforme o cenário apresentado neste trabalho, as transmissões se iniciam por iniciativa do cliente, após a sequência de vídeo ser transmitida a sessão é encerrada. A cada nova transmissão uma nova sessão é aberta.

Após serem transmitidos, os vídeos foram decodificados e armazenados nos receptores e a partir daí iniciou-se o processo de análise. Essa análise consiste na obtenção da média de valores Y-PSNR e VQMg de cada sequência de vídeo recebida, sendo tais valores relativos aos frames recebidos. Foram feitas 100 (cem) transmissões de cada sequência de vídeo.

7.2 Resultados com base em Y-PSNR

As Figuras de 7.1 a 7.7, respectivamente, apresentam os resultados obtidos nos testes considerando a métrica Y-PSNR.

Os gráficos apresentam no eixo das abscissas o número de transmissões realizadas para cada sequência de vídeo e no eixo das ordenadas as médias dos valores Y-PSNR dos frames recebidos a cada transmissão.

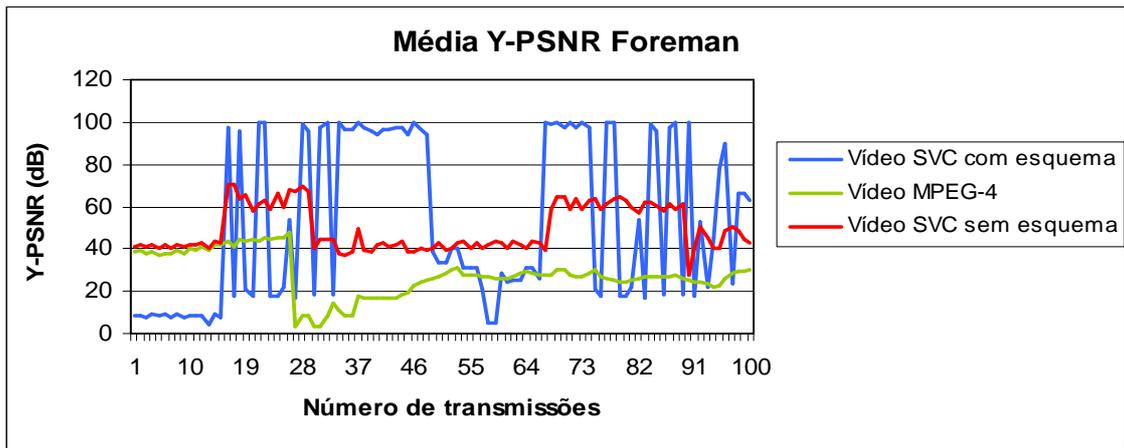


Figura 7.1: Média Y-PSNR da sequência de vídeo *Foreman*.

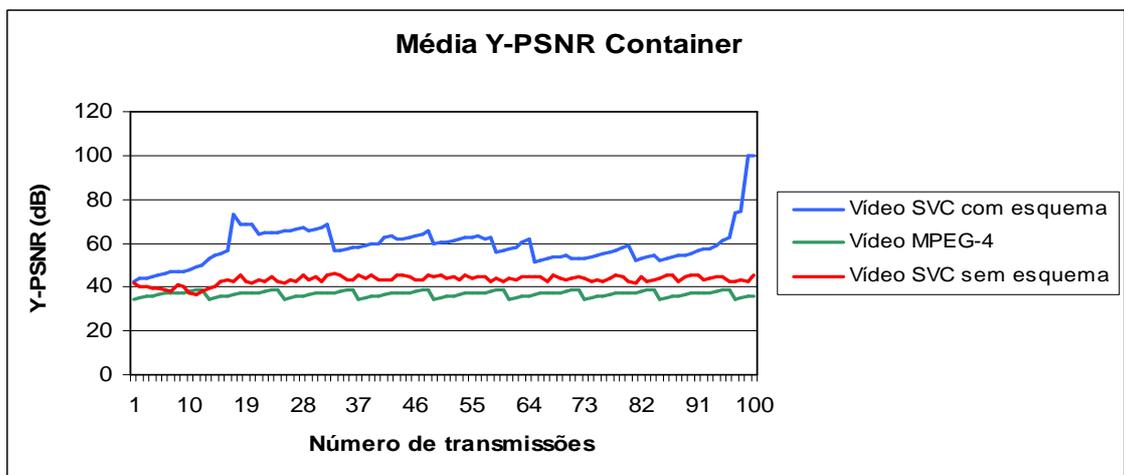


Figura 7.2: Média Y-PSNR da sequência de vídeo *Container*.

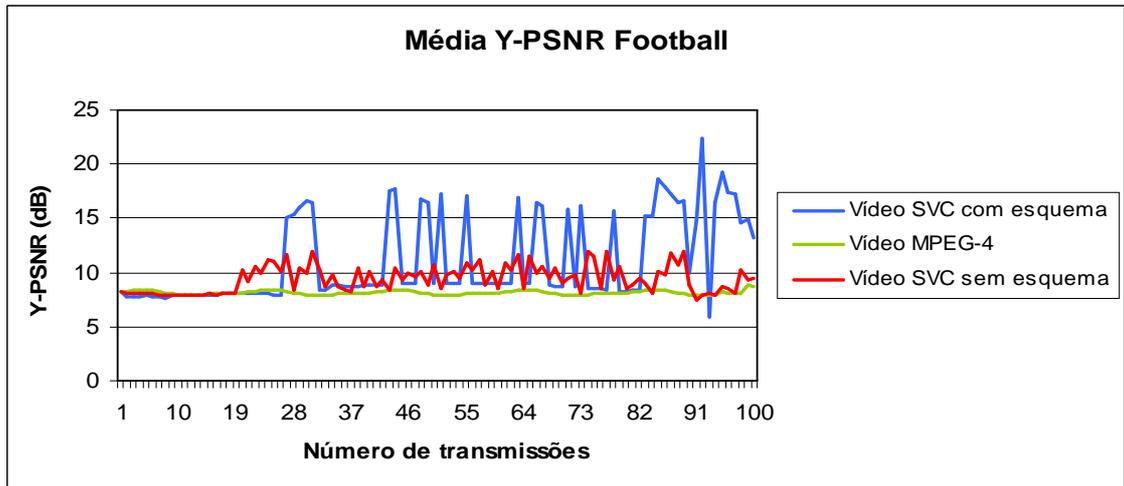


Figura 7.3: Média Y-PSNR da sequência de vídeo *Football*.

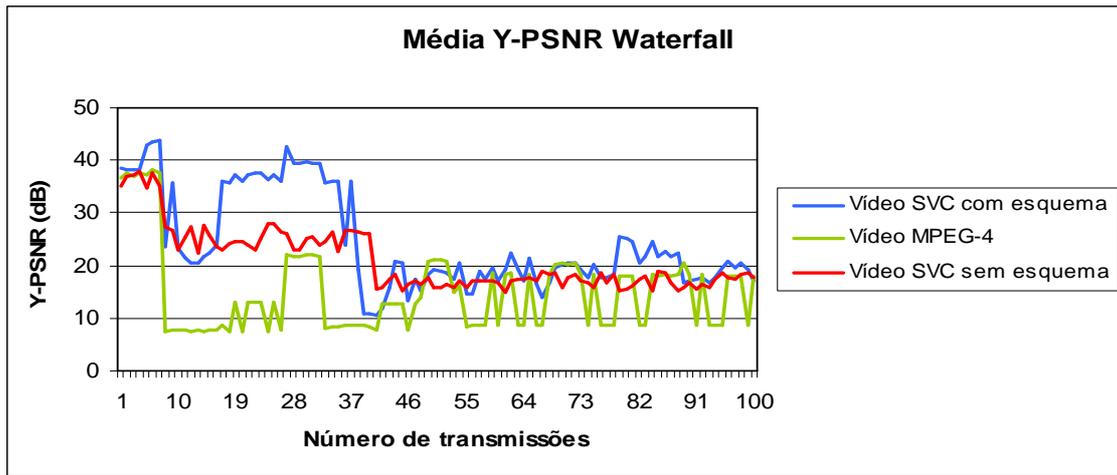


Figura 7.4: Média Y-PSNR da sequência de vídeo *Waterfall*.

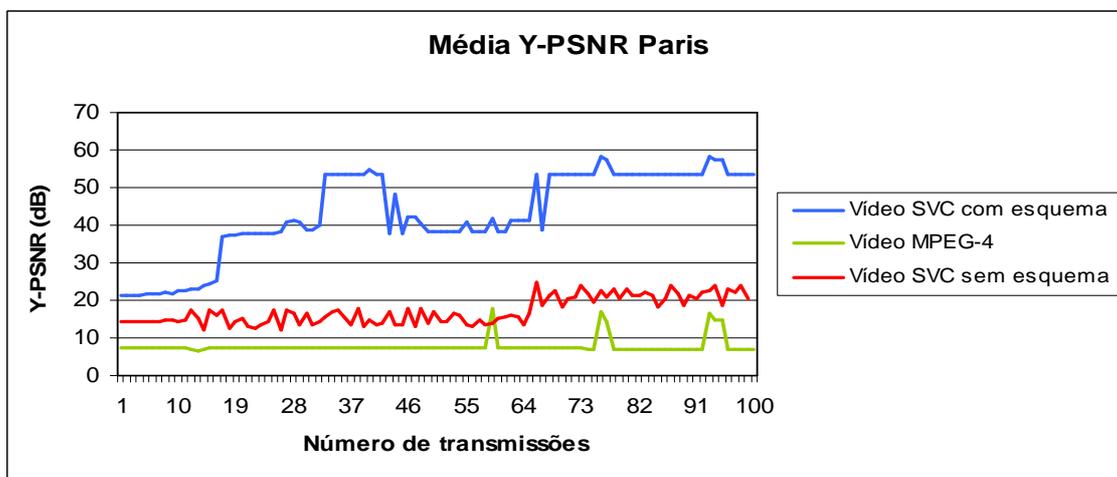


Figura 7.5: Média Y-PSNR da sequência de vídeo *Paris*.

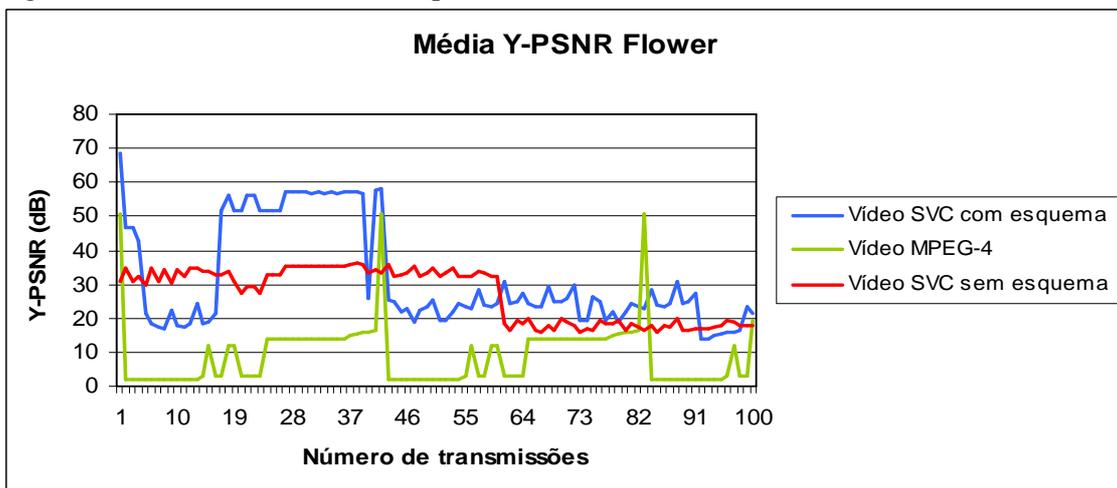


Figura 7.6: Média Y-PSNR da sequência de vídeo *Flower*.

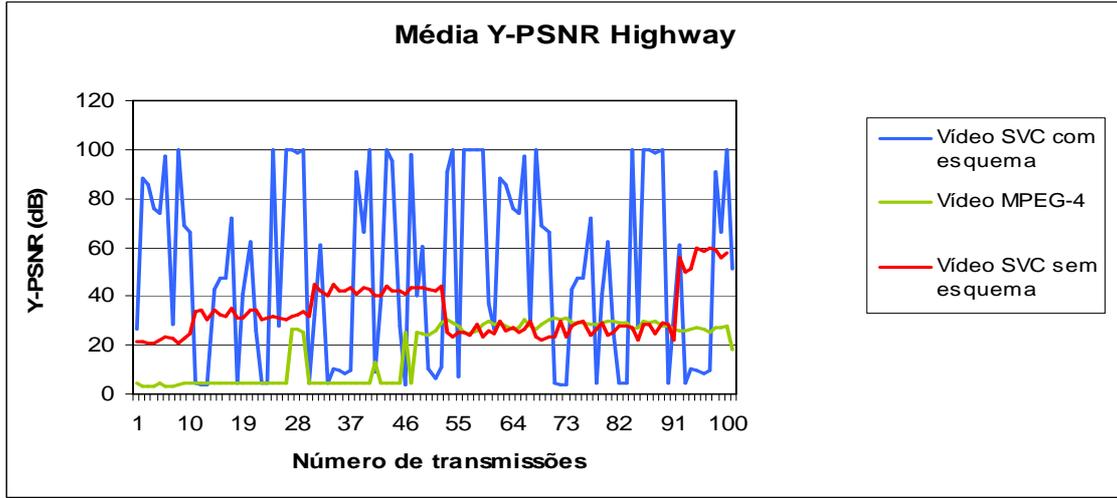


Figura 7.7: Média Y-PSNR da sequência de vídeo *Highway*.

O MPEG-4 é hoje o padrão de codificação mais usado para *streaming* em redes móveis, conforme os resultados obtidos foram observados que as sequências de vídeo codificadas com o padrão H.264/SVC mostraram-se superiores em valores médios de Y-PSNR em quase todas as transmissões.

O Y-PSNR também variou bastante em relação à classificação de cada sequência de vídeo. Foi observado que sequências de vídeo com poucas mudanças de cenas, poucas variações de cores e baixa compensação de movimento, obtiveram as melhores médias Y-PSNR, tanto as sequências de vídeo codificadas no padrão MPEG-4, quanto às codificadas no padrão H.264/SVC.

A equação (1) foi utilizada para o cálculo da diferença de porcentagem das sequências de vídeo transmitidas com a utilização do esquema adaptativo e o padrão H.264/SVC em relação às sequências de vídeo transmitidas no padrão MPEG-4 e H.264/SVC sem a utilização do esquema adaptativo.

$$Dif(s) = \left(\frac{\sum_{i=1}^{NT(s)} YPSNR_{\gamma i} - \sum_{j=1}^{NT(s)} YPSNR_{\alpha j}}{NT(s)} \right) \% \quad (1)$$

Onde $Dif(s)$ é a diferença média do desempenho geral do padrão H.264/SVC com o esquema, em relação ao padrão H.264/SVC sem o esquema ou ao padrão MPEG-4. $YPSNR_{\gamma i}$ é a média do Y-PSNR da sequência de vídeo s , codificada no padrão H.264/SVC com o esquema, na transmissão i , $YPSNR_{\alpha j}$ é a média do Y-PSNR da sequência de vídeo s codificada no outro padrão (MPEG-4 ou H.264/SVC sem o esquema) na transmissão j , e $NT(s)$ é a quantidade de vezes em que a sequência de vídeo s foi transmitida.

Com a ajuda da equação (1) foi possível comparar e confirmar a melhora introduzida pelo uso do padrão H.264/SVC combinado com o esquema adaptativo, em relação ao uso do padrão H.264/SVC sem o esquema e ao uso do padrão MPEG-4, na qualidade da sequência de vídeo recebida.

Na sequência de vídeo *Foreman* codificada no padrão H.264/SVC e com o uso do esquema adaptativo foi obtida uma melhora, em valores Y-PSNR, de 3,52 % em relação à mesma sequência codificada no padrão H.264/SVC sem o esquema. Em relação à sequência de vídeo codificada no padrão MPEG-4 a melhora foi de 24,99 %.

Na sequência de vídeo *Container* a melhora com o esquema foi de 15,65 % em relação à mesma sequência de vídeo codificada em H.264/SVC sem o esquema adaptativo e de 22,08 ,08 % em relação ao MPEG-4. Na sequência de vídeo *Football* a melhora com o esquema foi de 1,72% em relação à mesma sequência codificada em H.264 sem o esquema e de 2,99 % em relação à sequência codificada no padrão MPEG-4.

Na sequência de vídeo *Waterfall* a melhora com o esquema foi de 3,38 % em relação ao vídeo codificado em H.264/SVC sem o esquema e de 9,41 % em relação ao codificado em MPEG-4. Na sequência de vídeo *Paris* a melhora com o esquema em relação ao vídeo codificado em H.264/SVC sem o esquema foi de 25,87% e em relação ao codificado em MPEG-4 foi de 35,33%.

Na sequência de vídeo *Flower* a melhora com o esquema em relação ao vídeo codificado em H.264/SVC sem o esquema foi de 4,76 % e em relação ao codificado em MPEG-4 foi de 13,83 %. Na sequência de vídeo *Highway* a melhora com o esquema em relação ao vídeo codificado em H.264/SVC sem o esquema foi de 18,07 % e em relação ao codificado em MPEG-4 foi de 33,48 %.

A Tabela 7.1 apresenta o mapeamento correspondente aos valores de Y-PSNR e MOS obtidos com os resultados dos testes.

Tabela 7.1: Mapeamento entre PSNR e MOS.

Sequência de vídeo	Nº de Frames	Mapeamento entre Y-PSNR e MOS					
		MPEG-4		H.264/SVC com esquema adaptativo		H.264/SVC sem esquema adaptativo	
		Y-PSNR	MOS	Y-PSNR	MOS	Y-PSNR	MOS
<i>Highway</i>	2000	17,86	1	51,35	5	33,29	4
<i>Foreman</i>	300	27,88	3	52,88	5	49,36	5
<i>Paris</i>	1065	7,78	1	43,11	5	17,24	1
<i>Container</i>	300	36,80	4	58,90	5	43,24	5
<i>Flower</i>	250	8,89	1	31,79	4	27,09	3
<i>Football</i>	90	8,12	1	11,11	1	9,38	1
<i>Waterfall</i>	260	15,00	1	24,42	2	21,04	2

De acordo com [Ke *et al* 2008] e seguindo o mapeamento feito na Tabela 7.1 foi observado que a maioria das sequências de vídeo transmitidas no padrão MPEG-4 pode ser avaliada, conforme escala MOS, ruim com grau de deficiência da imagem muito incômodo. Já nas sequências de vídeo transmitidas no padrão H.264/SVC com uso do esquema pode-se considerar, conforme escala MOS, que as sequências de vídeo foram recebidas com um grau de deficiência da imagem imperceptível ou perceptível, mas não incomodo. Por fim, nas sequências de vídeo codificadas e transmitidas utilizando o padrão H.264/SVC sem uso do esquema o grau de deficiência da imagem pode ser considerado pouco incomodo (apropriado), conforme escala MOS.

É importante ressaltar que a utilização do PSNR para avaliar a qualidade de vídeo de forma objetiva não é suficiente para garantir uma boa avaliação subjetiva por parte do usuário, apesar do mapeamento realizado conforme as Tabelas 6.1 e 7.1. Dessa forma, é possível a ocorrência de situações em que valores elevados de PSNR (por exemplo, entre 31 dB e 37 dB, ou superiores a 37dB) não levem, necessariamente, a uma qualidade de vídeo considerada boa ou excelente, respectivamente, por parte do usuário final. Assim, uma avaliação subjetiva seria a ideal, apesar das dificuldades de implementação (incluindo a impossibilidade de sua automatização e a necessidade de maior tempo e de recursos humanos necessários).

As Figuras 7.8 a 7.14 apresentam os valores da média e do desvio padrão obtidos de acordo com a sequência de vídeo transmitida, onde SVC-CE representa o vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema, SVC-SE representa o vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema e MPEG-4 representa o vídeo transmitido e codificado no padrão MPEG-4.

A Figura 7.8 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Foreman*.

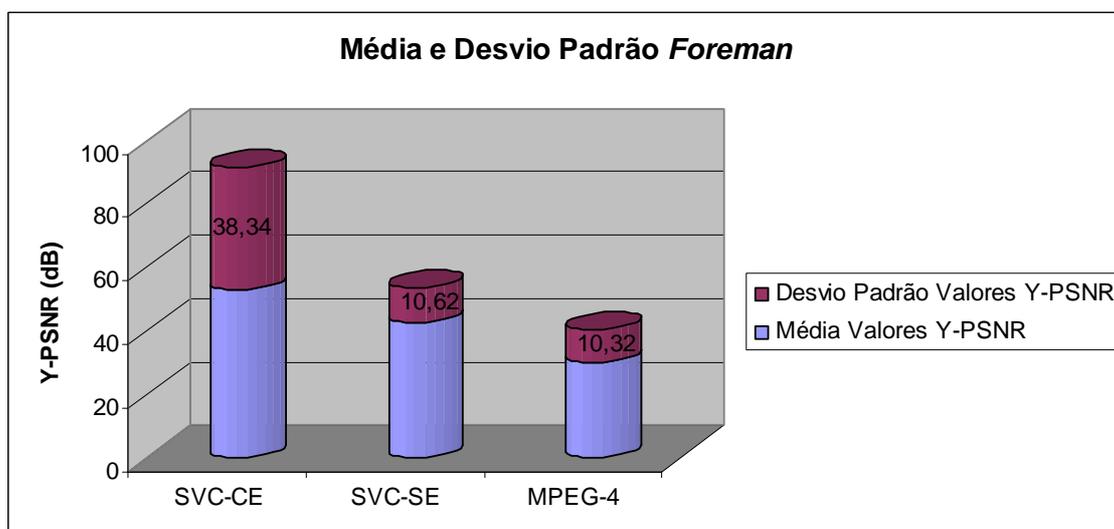


Figura 7.8: Média e desvio padrão *Foreman*.

O maior valor do desvio padrão pode ser observado na transmissão do vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema onde o valor foi de 38,34 dB. No vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema o valor do desvio padrão foi de 10,62 dB. O menor valor do desvio padrão foi no vídeo transmitido e codificado em MPEG-4, que foi de 10,32 dB.

A Figura 7.9 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Container*.

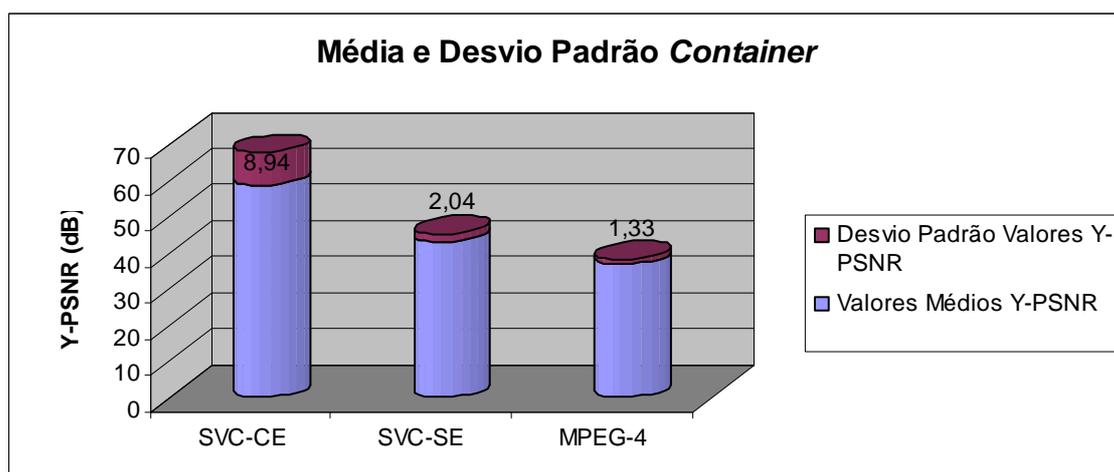


Figura 7.9: Média e desvio padrão *Container*.

Ao contrário do que acontece com a sequência de vídeo *Foreman*, os valores do desvio padrão encontrados para essa sequência foram bem menores. Sendo, o maior valor do desvio padrão pode ser observado na transmissão do vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema onde o valor foi de 8,94

dB. No vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema o valor do desvio padrão foi de 2,04 dB. O menor valor do desvio padrão foi no vídeo transmitido e codificado em MPEG-4, que foi de 1,33 dB.

A Figura 7.10 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Football*.

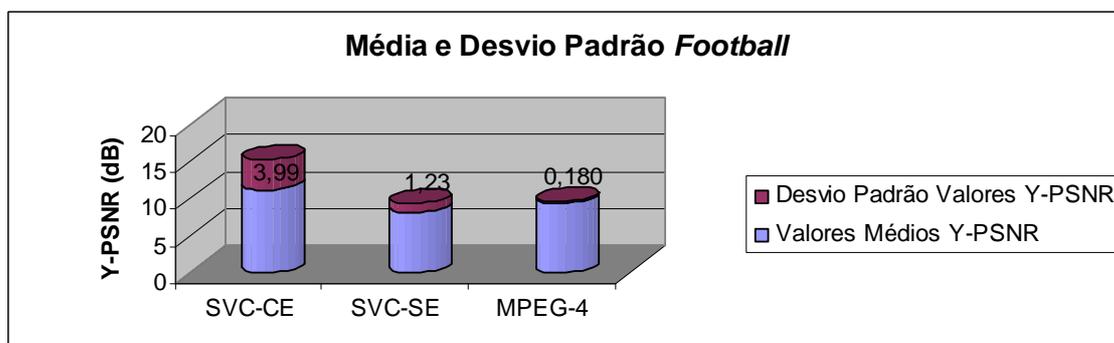


Figura 7.10: Média e desvio padrão *Football*.

Conforme apresentado na Figura 7.10 a sequência de vídeo *Football*, obteve valores do desvio padrão de 3,99 dB para a sequência de vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema. Já no vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema o valor do desvio padrão foi de 1,23 dB e o menor valor do desvio padrão foi no vídeo transmitido e codificado em MPEG-4, que foi de 0,180 dB.

A Figura 7.11 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Waterfall*.

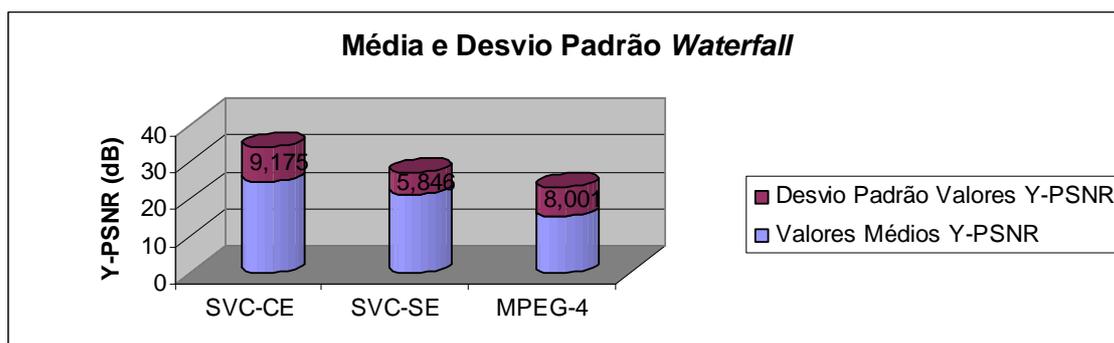


Figura 7.11: Média e desvio padrão *Waterfall*.

De acordo com a Figura 7.11 apresentada, a sequência de vídeo *Waterfall*, obteve valores do desvio padrão de 9,17 dB para a sequência de vídeo codificada no padrão H.264/SVC e transmitida com a utilização do esquema. Já na sequência de vídeo codificada em H.264/SVC e transmitida sem a utilização do esquema o valor do desvio padrão foi de 5,84 dB sendo o menor valor do desvio padrão calculado e de 8,00 dB para a sequência de vídeo transmitida e codificada em MPEG-4.

A Figura 7.12 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Paris*.

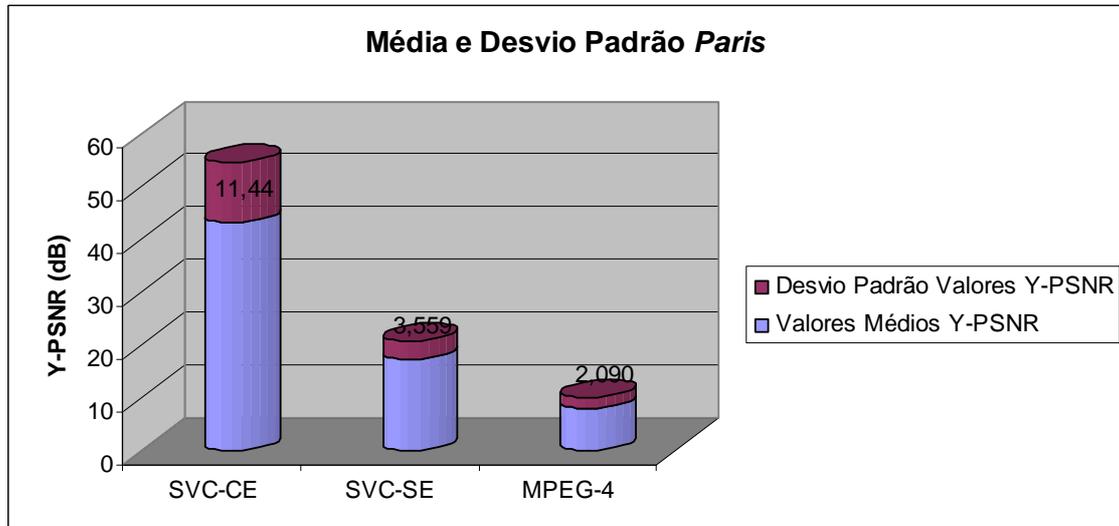


Figura 7.12: Média e desvio padrão *Paris*.

Conforme a Figura 7.12 apresentada, o maior valor do desvio padrão pode ser observado na transmissão do vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema onde o valor foi de 11,44 dB. No vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema o valor do desvio padrão foi de 3,55 dB e o menor valor do desvio padrão foi no vídeo transmitido e codificado em MPEG-4, que foi de 2,090 dB.

A Figura 7.13 apresenta os valores da média do desvio padrão calculados para a sequência de vídeo *Flower*.

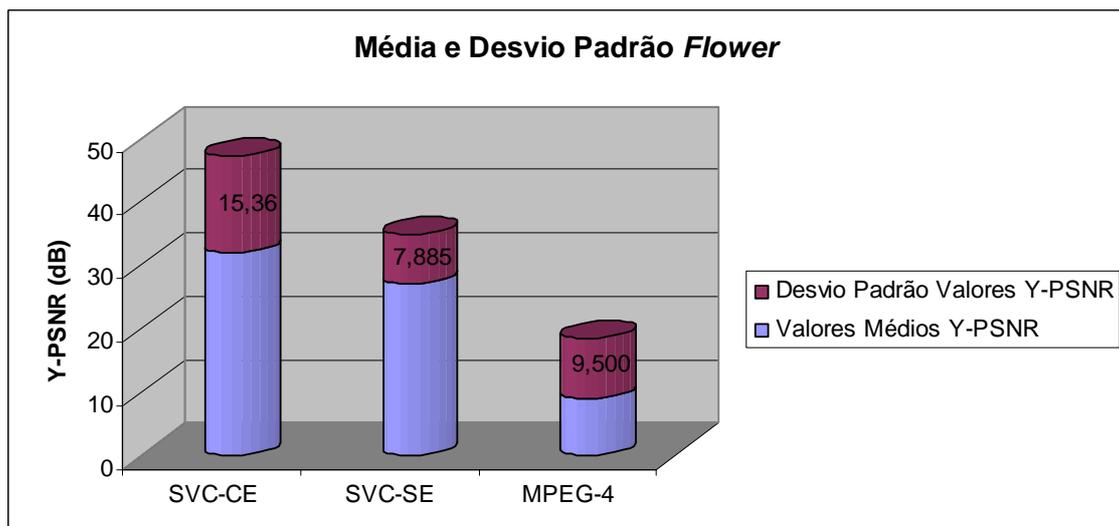


Figura 7.13: Média e desvio padrão *Flower*.

De acordo com a Figura 7.13 apresentada, a sequência de vídeo *Flower*, obteve valores do desvio padrão de 15,36 dB para a sequência de vídeo codificada no padrão

H.264/SVC e transmitida com a utilização do esquema. Já na sequência de vídeo codificada em H.264/SVC e transmitida sem a utilização do esquema o valor do desvio padrão foi de 7,88 dB sendo o menor valor do desvio padrão calculado e de 9,50 dB para a sequência de vídeo transmitida e codificada em MPEG-4.

A Figura 7.14 apresenta os valores da média e do desvio padrão calculados para a sequência de vídeo *Highway*.

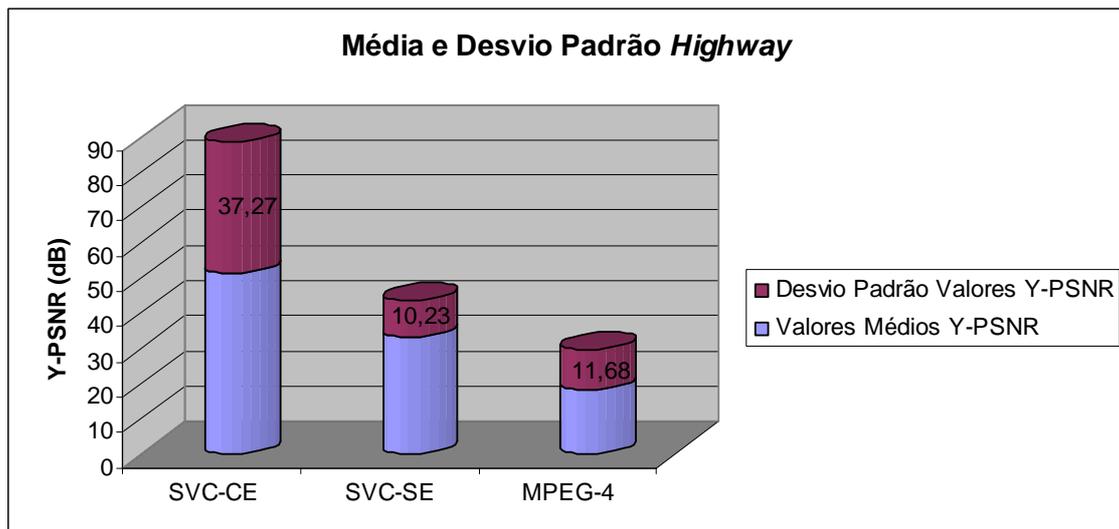


Figura 7.14: Média e desvio padrão *Highway*.

O maior valor do desvio padrão, conforme Figura 7.14 pode ser observado na transmissão do vídeo codificado no padrão H.264/SVC e transmitido com a utilização do esquema onde o valor foi de 37,27 dB. O menor valor do desvio padrão foi no vídeo codificado em H.264/SVC e transmitido sem a utilização do esquema o valor do desvio padrão foi de 10,23 dB e o valor do desvio padrão para o vídeo transmitido e codificado em MPEG-4, foi de 11,68 dB.

7.3 Resultados com base em VQMg

Após experimentos realizados utilizando o esquema adaptativo proposto por [Mulinari e Gondim 2009], e resultados obtidos nos experimentos realizados com o cenário apresentado neste trabalho, foram obtidos os valores de VQMg.

Os valores de VQMg variam de 0 até 1, onde 0 corresponde à correlação máxima entre o vídeo original e o processado e 1 corresponde à correlação mínima. A Figura 7.15 apresenta os valores da média e do desvio padrão VQMg obtidos para cada sequência de vídeo transmitida com o uso do esquema.

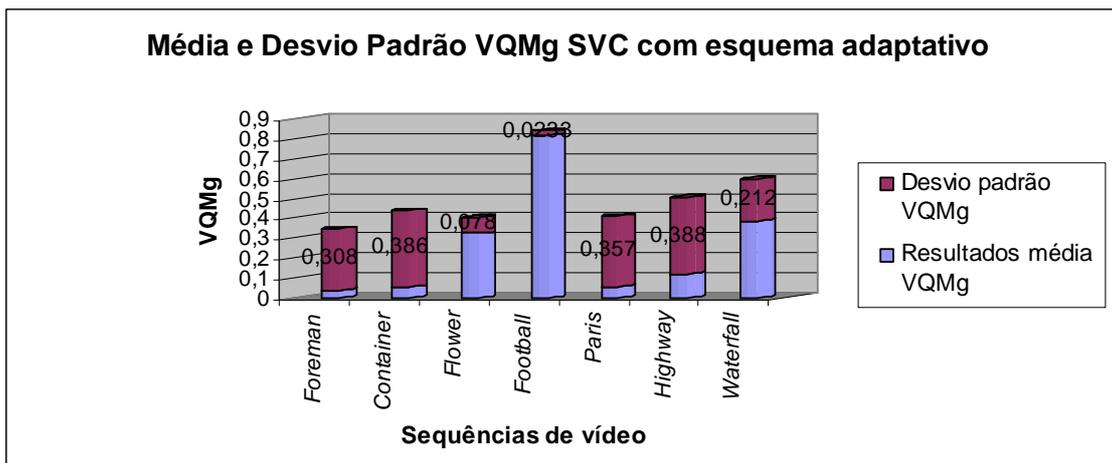


Figura 7.15: Resultados média e desvio padrão VQMg com o uso do esquema.

De acordo com a Figura 7.15, os valores de VQMg medidos para os vídeos transmitidos pelo esquema ficaram próximos de 0,1 a 0,4, que pode ser considerado como de qualidade ótima. Com exceção da sequência de vídeo *Football*, cujo valor médio ficou muito próximo de 0,8, considerado ruim. O maior desvio padrão foi obtido na sequência de vídeo *Highway*, cujo valor ficou em 0,388 e o menor valor do desvio padrão foi obtido na sequência de vídeo *Football*.

A Figura 7.16 apresenta os valores da média e do desvio padrão VQMg obtidos para a transmissão das sequências de vídeo codificadas no padrão MPEG-4.

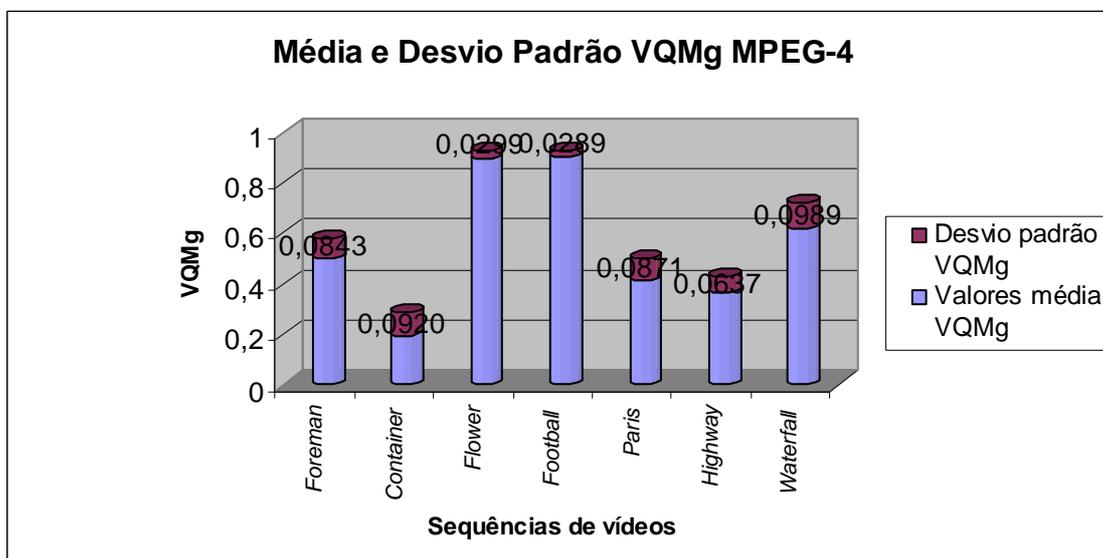


Figura 7.16: Resultados média e desvio padrão VQMg para os vídeos MPEG-4.

Conforme apresentado na Figura 7.16, os valores de VQMg medidos para os vídeos *Flower* e *Football* foram muito próximos de 0,9, considerado de média a má qualidade, enquanto que os vídeos *Foreman*, *Paris* e *Highway*, foram medidos entre 0,3 e 0,5, considerado de boa qualidade. O valor medido de VQMg para o vídeo *Waterfall* ficou próximo de 0,6, considerado de qualidade média. O desvio padrão obtido para

todas as sequências de vídeo ficou abaixo de 0,1. O maior desvio padrão obtido foi para sequência de vídeo *Waterfall* que foi de 0,098 e o menor desvio padrão obtido foi de 0,0289 que para a sequência de vídeo *Football*.

A Figura 7.17 apresenta os valores da média e desvio padrão VQMg obtidos para a transmissão das sequências de vídeo codificadas no padrão H.264/SVC sem a utilização do esquema adaptativo.

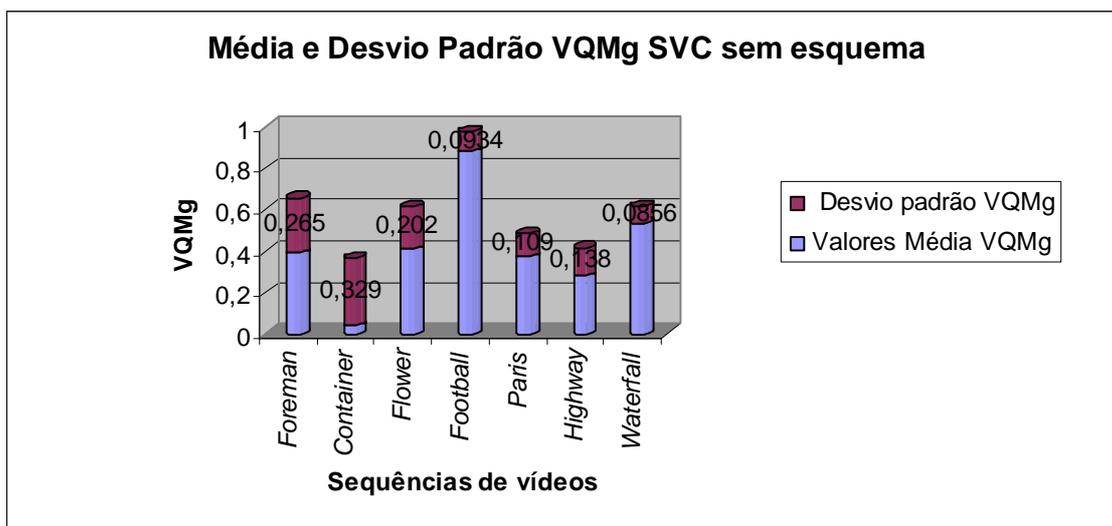


Figura 7.17: Resultados média e desvio padrão VQMg dos vídeos SVC sem o esquema.

Conforme apresentado na Figura 7.17, os vídeos transmitidos codificados no padrão H.264/SVC sem a utilização do esquema adaptativo obtiveram valores de VQMg mais baixos, quando comparados aos valores médios VQMg obtidos no padrão MPEG-4 e mais altos, quando comparados aos resultados dos valores médios VQMg obtidos com os vídeos codificados no padrão H.264/SVC utilizando o esquema. O valor VQMg obtido para a sequência de vídeo *Foreman* foi de 0,3.

Para a sequência de vídeo *Flower* foi de 0,4, para a sequência de vídeo *Highway* foi de 0,2 e para a sequência de vídeo *Waterfall* foi de 0,5, todos considerados de boa qualidade, dando destaque para a sequência de vídeo *Container* que ficou bem próximo de 0, considerada de ótima qualidade. O maior desvio padrão obtido foi para a sequência de vídeo *Container*, cujo valor foi de 0,329, e o menor desvio padrão obtido foi para a sequência de vídeo *Waterfall*, cujo valor foi de 0,0856.

De acordo com os resultados apresentados para o desvio padrão VQMg, podemos concluir que em algumas transmissões, para algumas sequências de vídeo, houve valores, cuja a correlação entre o vídeo original e o transmitido ficaram próximos a correlação máxima, mesmo quando os valores médios apresentaram, em alguns casos, valores que correspondem a uma sequência de vídeo com qualidade ruim.

8 Conclusões e Trabalhos Futuros

8.1 Conclusões

O trabalho descrito nesta dissertação visou à análise e comparação entre dois padrões de codificação de vídeo, o MPEG-4 e o H.264/SVC, com base em sequências transmitidas em rede sem fio de 3ª geração. O processo de avaliação foi feito com base nos valores obtidos para as métricas objetivas *Y-PSNR* e *VQM*, tanto dos vídeos transmitidos com a ajuda do esquema adaptativo, quanto dos vídeos transmitidos com a ajuda de outros codificadores/decodificadores.

Os resultados obtidos mostram que a qualidade das sequências de vídeo sofre impacto das características do vídeo, das características dos padrões de codificação utilizados, além dos parâmetros de QoS que também afetam a qualidade do vídeo.

A análise dos resultados mostrou que, no contexto considerado, o padrão H.264/SVC possibilitou uma melhor qualidade dos vídeos recebidos, dando destaque aos resultados obtidos com a utilização do esquema adaptativo, onde os resultados se mostram, em alguns casos, bem superiores aos resultados obtidos com o padrão MPEG-4.

Foi realizado um mapeamento dos resultados obtidos com a métrica Y-PSNR na métrica MOS, a partir do qual verificou-se que os melhores resultados foram obtidos com as sequências de vídeo codificadas no padrão H.264/SVC, adotando o esquema adaptativo para prover adaptabilidade da infraestrutura de rede.

Foi proposta uma equação (equação 1) para quantificar a melhora das sequências de vídeo transmitidas com o esquema adaptativo em relação às sequências de vídeo que foram transmitidas em MPEG-4 e em H.264/SVC.

Com a utilização do esquema adaptativo, foi possível adaptar a taxa de transmissão do *streaming* de vídeo de acordo com as mudanças e oscilações ocorridas na rede.

Verificou-se, ainda, os reflexos dessa adaptação de taxa de transmissão sobre a qualidade do vídeo entregue ao usuário final, por meio de métricas objetivas, demonstrando a adequação do esquema adaptativo anteriormente proposto em [Mulinari and Gondim, 2009].

8.2 Trabalhos Futuros

Para trabalhos futuros podem ser propostas:

1. Priorização de tráfego multimídia, com base em políticas de QoS e sua influência na qualidade de experiência;
2. Análise de desempenho utilizando outras combinações de escalabilidade em redes de 3ª geração;
3. Avaliação do desempenho da escalabilidade de vídeo e do esquema adaptativo utilizado neste trabalho em redes de 4ª geração LTE;
4. Avaliação objetiva utilizando a métrica SSIM (*Structural Similarity*);
5. Avaliação da qualidade de vídeo e de experiência em redes heterogêneas.

Referências

[3GPP 2010] 3GPP TSG, “*Overview of 3GPP Release 5*”, ETSI Mobile Competence Centre, ETSI 2003. Disponível em: http://www.3gpp.org/ftp/tsg_ran/tsg_ran/TSGR_20/Docs/PDF/RP-030375.pdf. Acessado em: 01-11-2010.

[3GPP TS 2010] 3GPP TS 26.244 “*Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS)*”; 3GPP file format (3GP) (Release 9) V9.2.0.

[Abdallah *et al* 2011] Abdallah, H.A. and Hadhoud, M.M. and Shaalan, A.A. and Elsamie, F.E.A. (2011), “*Blind Wavelet-Based Image Watermarking*”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 4.

[Alexiou *et al* 2005] Alexiou, AG and Bouras, CJ and Igglesis, VG. (2005), “*A decision feedback scheme for multimedia transmission over 3G mobile networks*”. Second IFIP International Conference on Wireless and Optical Communications Networks. WOCN, pages 357-361.

[Antoniou *et al* 2007] Antoniou, P. and Vassiliou, V. and Pitsillides, A. (2007), “*ADIVIS: A Novel Adaptive Algorithm for Video Streaming over the Internet*”. IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, pages 1-5.

[Apache Server 2010] Apache - <http://www.apachelounge.com/download/>. Acessado em Outubro 2010.

[Apache] Apache Server : <http://h264.code-shop.com/trac/wiki/Mod-H264-Streaming-Apache-Version2>. Acessado em outubro, 2010.

[Blestel and Raulet 2010] Blestel, M. and Raulet, M. (2010), “*Open SVC Decoder: a Flexible SVC Library*”. Proceedings of the international conference on Multimedia, pages 1463-1466.

[Bouras and Gkamas 2005] Bouras, C. and Gkamas, A. (2005), “*Performance of Adaptive Multimedia Transmission: The case of Unicast Technique*”, 5th Int'l Network Conference INC'05, Proc.

[BVQM 2010] *BVQM – Batch Video Quality Metric*. Disponível em <http://www.its.bldrdoc.gov/vqm/>. Acessado em Outubro, 2010.

[Chan *et al* 2008] Chan, Y.S. and Cosman, P.C. and Milstein, L.B. (2008), “*A Multiple Description Coding and Delivery Scheme for Motion-Compensated Fine Granularity Scalable Video*”, IEEE Transactions on Image Processing, Vol. 17, number 8, pages 1353-1367.

[Changhoon and Bovik 2011] Changhoon Y.,and Bovik C. A.(2011), “*Quality Assessment of Deblocked Images*”, IEEE Transactions on image processing, Vol. 20, Janeiro 2011.

[Chuang *et al* 2010] Chuang, T.D. and Chen, Y.J. and Chen, Y.H. and Chien, S.Y. and Chen, L.G. (2010), “*Architecture Design of Fine Grain Quality Scalable Encoder with CABAC for H.264/AVC Scalable Extension*”, Journal of Signal Processing Systems, Springer, Vol. 60, number 3, pages 363-375.

[Davies *et al* 2008] Davies, M. and Dantcheva, A. and, Fröhlich, P. (2008), “*Comparing Access Methods and Quality of 3G Mobile Video Streaming Services*”. CHI'08 extended abstracts on Human factors in computing systems, page 2817-28[Wiegand et al 2003] .

[Diaz *et al* 2007] Diaz, A. and Merino, P. and Panizo, L. and Recio, A.M. (2007), “*Evaluating Video Streaming Over GPRS/UMTS Networks: A Practical Case*”, The 65th IEEE vehicular technology conference, VTC , Spring, pages 624-628.

[Dziri *et al.* 2008] Dziri, A. and Diallo, A. and Kieffer, M. and Duhamel, P.(2008), “*P-Picture based H.264 AVC to H.264 SVC Temporal Transcoding*”, Wireless Communications and Mobile Computing Conference. IWCMC'08. International, pages 425-430.

[Fajardo *et al* 2009] Fajardo, J.O. and Liberal, F. and Bilbao, N.(2009),“*Impact of the Video Slice Size on the Visual Quality for H.264 over 3G UMTS Services*”. Sixth International Conference on Broadband Communications, Networks, and Systems, BROADNETS, page 1.

[FFMPEG 2010] FFmpeg - <http://www.ffmpeg.org/>. Acessado em Janeiro, 2010.

[Frojdth *et al* 2006] Frojdth, P. and Horn, U. and Kampmann, M. and Nohlgren, A. and Westerlund, M. (2006), “*Adaptive streaming within the 3GPP packet-switched streaming service*”, IEEE Network, 2006, Vol. 20, pages 34-40.

[Hartung *et al* 2007] Hartung, F. and Horn, U. and Huschke, J. and Kampmann, M. and Lohmar, T. and Lundevall, M. (2007), “*Delivery of Broadcast Services in 3G Networks*”, IEEE Transactions on Broadcasting, Vol. 53, number 1, pages. 188-199.

[Hussain 2009] Hussain, S.Z.(2009), “*Performance Evaluation of H.264/AVC encoded Video over TETRA Enhanced Data Service (TEDS)*”, Department Communications and Networking.

[Janevski 2003] Janevski, T. (2003), “*Traffic Analysis and Design of Wireless IP Networks*”, Artech House Publishers, London.

[JSVM 2010] JSVM Software Manual.<<http://ube.ege.edu.tr/~boztok/JSVM/SoftwareManual.pdf>. Acessado em outubro, 2010.

[Kang *et al* 2007] Kang, K. and Cho, Y. and Cho, J. and Shin, H. (2007), “*Scheduling Scalable Multimedia Streams for 3G Cellular Broadcast and Multicast Services*” IEEE Transactions on Vehicular Technology, Vol. 56, number 5, pages 2655-2672.

[Kang *et al.* 2009] Kang, J. and Goh, G. and Chung, K.(2009), “A New Inter-Layer Prediction Scheme for Spatial Scalability with Different Frame Rates”, Proceedings of the 2009 ACM Symposium on Applied Computing, pages 1779-1783.

[Ke *et al* 2008] Ke, C.H. and Shieh, C.K. and Hwang, W.S. and Ziviani, A. (2008), “An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission” Journal of information science and engineering, Vol. 24, pages 425-440.

[Kuipers *et al* 2010] Kuipers, F. and Kooij, R. and De Vleeschauwer, D. and Brunnstrom K. (2010), “Techniques for Measuring Quality of Experience”, Wired/Wireless Internet Communications WWIC, Springer-Verla, pages.216-[Wiegand *et al* 2003] 7.

[Kurose *et al* 2006] Kurose J. F., Ross K. W. (2006), “Redes de Computadores e a Internet – Uma Abordagem top down”, Person Addison Wesley, 3ª edição.

[Li *et al* 2011] Li, X. and Amon, P. and Hutter, A. and Kaup, A. (2011), “Performance Analysis of Inter-Layer Prediction in Scalable Video Coding Extension of H.264/AVC”, IEEE Transactions on Broadcasting, number 99, pages 1-1.

[Liu *et al* 2009] Liu, Y. and Zhang, S. and Xu, S. and Zhang, Y. (2009), Yin Z., “Research on H.264/SVC Compressed Video Communication in 3G”. 4th International Conference on Computer Science e Education, ICCSE'09, pg 327-332.

[Ma *et al* 2011] Ma, K.J. and Bartos, R. and Bhatia, S. (2011), “A Survey of Schemes for Internet-Based Video Delivery”. Journal of Network and Computer Applications, Elsevier.

[Martinez-Yelmo *et al* 2010] Martinez-Yelmo, I. and Seoane, I. and Guerrero, C. (2010), “Fair Quality of Experience (QoE) Measurement Related with Networking Technologies”, Wired/Wireless Internet Communications, pages [Wiegand *et al* 2003] 8-239.

[McFarland *et al* 2007] McFarland, M.A. and Pinson, M.H. and Wolf, S (2007), “Batch Video Quality Metric (BVQM) User’s Manual”, Insitute for Telecommunication Sciences, Boulder, Colorado, USA, Department of Commerce, Nacional Telecommunications and Information Administration, NTIA Handbook HB-08-441b.

[Megrich 2009] Megrich, A.M. (2009), “Televisão Digital Princípios e Técnicas”, 1ª Edição, São Paulo, Editora Érica Ltda.

[Migliorini *et al* 2011] Migliorini, D. and Mingozi, E. and Vallati, C. (2011), “Performance Evaluation of H.264/SVC Video Streaming Over Mobile Wimax”, Computer Networks, Elsevier.

[Mplayer 2010] Mplayer - <http://sourceforge.net/projects/opensvcdecoder/>. Acessado em Outubro, 2010.

[Mulinari e Gondim 2009] Mulinari R., Gondim P. R. L. (2009), “Esquemas Adaptativos para Distribuição de Vídeo na Internet”. Dissertação de Mestrado – Departamento de Engenharia Elétrica. UnB.

[Narvekar *et al* 2009] Narvekar, N.D. and Konnanath, B. and Mehta, S. and Chintalapati, S. and AlKamal, I. and Chakrabarti, C. and Karam, L.J. (2009), “*An H.264/SVC Memory Architecture Supporting Spacial and Course-Grained Quality Scalabilities*”, 16th IEEE International Conference on Image Processing –ICIP, pages 2661-2664.

[Nunes 2009] Nunes R. P. (2009), “Codificação Escalável de Vídeo para Recepção Fixa no Sistema Brasileiro de Televisão Digital”, Dissertação de Mestrado, Universidade de São Paulo, 2009.

[Park *et al* 2009] Park, C.S. and Baek, S.J. and Yoon, M.S. and Kim, H.K. and Ko, S.J. (2009), “*Selective Inter-layer Residual Prediction for SVC based Video Streaming*”, IEEE Transactions on Consumer Electronics, Vol. 55, number 1, pages. 235-239.

[Pinson and Wolf 2004] Pinson, M.H. and Wolf, S. (2004), “*A New Standardized Method for Objectively Measuring Video Quality*”, Transactions On Broadcasting, IEEE, Vol. 50, number 3, pages 312-3[Wiegand *et al* 2003] .

[Pliakas *et al* 2007] Pliakas, T. and Kormentzas, G. and Skianis, C. (2007), “*Scalable Video Streaming Traffic Delivery in IP/UMTS Networking Environments*”, Journal of Multimedia, Vol. 2, number 2, pages 37-46.

[Punchihewa *et al* 2003] Punchihewa, A. and Bailey, D.G. and Hodgson, RM. (2003), “*A Survey Of Coded Image and Video Quality Assessment*”, Proceedings of Image and Vision Computing New Zealand, pages 326-331.

[Ramos 2004] Ramos, M. C. (2004), “*Desempenho da Modulação Adaptativa em Enlaces WCDMA/HSDPA em Presença de Multipercursos*”, Dissertação de Mestrado, Pontifícia Universidade Católica do Rio de Janeiro.

[RFC 3550 2003] RFC 3550, RTP (2003): “*A Transport Protocol for Real-Time Applications*”.

[RFC 768 1980] RFC 768, UDP – *User Datagram Protocol*.

[Richardson 2003] Richardson Iain E. G. (2003), “*H.264 and MPEG-4 Video Compression Video Coding for Next-generation Multimedia*”, WILEY.

[Rieckh and Rupp 2008] Rieckh, J. and Rupp, M.(2008), “*Scalable Video for Peer-to-Peer Streaming*”, Institute of Communications and Radio-Frequency Engineering Technical University of Vienna Summer, Outlook.

[Ries *et al* 2008] Ries, M. and Nemethova, O. and Rupp, M. (2008), “*Video Quality Estimation for Mobile H.264/AVC Video Streaming*”. Journal of Communications. Vol. 3, number 1, pages 41-50.

[Rojagopalan 2010] Rojagopalan R. (2010), “*Video Quality Measurements for Mobile Networks, Openware Media Optimizer*”, Providing Best-in-Class Video with Superior Quality and User Experience, Product Marketing Manager – Media Optimization, Openware Systems Inc.

[Schierl *et al* 2005] Schierl, T. and Wiegand, T. and Kampmann, M. (2005), “*3GPP compliant adaptive wireless video streaming using H.264/AVC*”, IEEE International Conference on Image Processing, Vol. 3 pages III-696, ICIP 2005.

[Schierl *et al* 2007] Schierl, T. and Stockhammer, T. and Wiegand, T. (2007), “*Mobile Video Transmission Using Scalable Video Coding*”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, number 9, pages 1204-1217.

[Schwarz and Wiegand 2008] Schwarz, H. and Wiegand T. (2008), “*The Scalable Video Coding Amendment of H 264/AVC Standard*”, Heinrich-Hertz-Institut.

[Schwarz and Wien 2008] Schwarz, H. and Wien, M.(2008), “*The Scalable Video Coding Extension of the H.264/AVC Standard*, IEEE Signal Processing Magazine, Institute of Electrical and Electronic Engineers, Vol. 25, number 2, page 135.

[Schwarz *et al.* 2007] Schwarz, H. and Marpe, D. and Wiegand, T. (2007), “*Overview of the Scalable Video Coding Extension of the H.264/AVC Standard*”, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 17, numero 9, pages 1103--1120.

[Seshadrinathan *et al* 2010] Seshadrinathan, K. and Soundararajan, R. and Bovik, A.C. and Cormack, L.K. (2010), “*Study of Subjective and Objective Quality Assessment of Video*”. IEEE Transactions on Image Processing, Vol. 19, number 6, pages 1427-1441.

[Shahid *et al.* 2009] Shahid, Z. and Chaumont, M. and Puech, W. (2009), “*An Adaptive Scan of High Frequency Subbands for Dyadic Intra Frame in MPEG4-AVC/H.264 Scalable Video Coding*”, Electronic Imaging, Visual Communications and Image Processing, SPIE, IS&T.

[Shoaib and Waheed 2008] Shoaib, M. and Waheed, M. (2008), “*Streaming Video in Cellular Networks Using Scalable Video Coding Extension of H.264-AVC*”, 4th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM'08, pages 1-4.

[Silva 2008] Silva, E.P.F. (2008), “*Transcodificador De Vídeo Wyner-Ziv/H.263 Para Comunicação Entre Dispositivos Móveis*”, Dissertação de Mestrado em Engenharia Elétrica, Universidade de Brasília - UnB.

[Singh *et al* 2009] Singh, V. and Ott, J. and Curcio, I.D.D.(2009), “*Rate Adaptation for Conversational 3G Video*”, IEEE INFOCOM Workshops 2009, pages 1-7.

[Sohn *et al* 2010] Sohn, H. and Yoo, H. and De Neve, W. and Kim, C.S. and Ro, Y.M. (2010), “*Full-Reference Video Quality Metric for Fully Scalable and Mobile SVC Content*”, IEEE Transactions on Broadcasting, Vol. 56, number 3, pages 269-280.

[Stretch 2011] Stretch, M. O. (2011), Tutorial: “*The H.264 Scalable Video Codec (SVC)*”. Disponível em: <http://www.eetimes.com/design/signal-processing->

dsp/4017613/Tutorial-The-H-264-Scalable-Video-Codec-SVC. Acessado em 02/02/2012.

[Uitto and Vehkaperä 2009] Uitto, M. and Vehkaperä, J. (2009), “*Spatial Enhancement Layer Utilisation for SVC in Base Layer Error Concealment*”, Proceedings of the 5th International ICST Mobile Multimedia Communications Conference, Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering – ICST, page 10.

[Van der Auwera *et al.* 2008.] Van der Auwera, G. and David, P.T. and Reisslein, M.(2008), “*Traffic and quality characterization of single-layer video streams encoded with the H. 264/MPEG-4 advanced video coding standard and scalable video coding extension*”, IEEE Transactions on Broadcasting, Vol. 54, number 3, pages 698-718.

[Vídeo Lan 2010] Video Lan - <http://www.videolan.org/>. Acessado em Junho, 2010.

[Vídeo] Vídeo: <http://www.tkn.tu-berlin.de/research/evalvid/cif.html>. Acessado em Junho, 2011.

[VQEG 2011] VQEG - <http://www.its.bldrdoc.gov/vqeg/projects/projects.php>. Acessado em Agosto, 2011.

[Wang 2006] Wang, Y. (2006), “*Survey of Objective Video Quality Measurements*”, EMC Corporation Hopkinton, MA, Vol. 1748.

[Wiegand *et al* 2003] Wiegand, T. and Sullivan, G.J. and Bjontegaard, G. and Luthra, A. (2003), “*Overview of the H.264/AVC Video Coding Standard*”, IEEE Transactions on Circuits and Systems for video technology, Vol. 13, number 7, pages 560-576.

[Wien *et al* 2007] Wien, M. and Schwarz, H. and Oelbaum, T.(2007), “*Performance Analysis of SVC*”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, number 9, pages 1194-1203.

[Xiaoqing *et al.*2008] Xiaoqing Zhua, Thomas Schierlb, Thomas Wiegandb and Bernd Giroda (2008), “*Video Multicast over Wireless Mesh Networks with Scalable Video Coding (SVC)*”, NSF Grant CCR-0325639.

[Yannuzzi *et al* 2010] Yannuzzi, M. and Monteiro, E. and Masip-Bruin, X. (2010), “*Na Overview of Quality of Experience Measurement Challenges for Video Applications in IP Networks*”, 8th International Conference Wired/Wireless Internet Communications: WWIC, pages 252-263.

[Zamek and Zamek 2005] Zamek, I. and Zamek, S. (2005), “*Definitions of Jitter Measurement Terms and Relationships*”. IEEE International Test Conference, Proceedings, ITC, pages 10—pp.

APÊNDICE I

I.1 – Artigo publicado no evento COMTEL 2011

Streaming de Vídeo Escalável em Rede Celular de 3ª Geração

Vanice Canuto Cunha¹ Paulo Roberto de Lira Gondim¹ Rodrigo Mulinari¹

¹Universidade de Brasília

Faculdade de Tecnologia

vanicecunha@gmail.com, pgondim@unb.br, rmulinari@gmail.com

Resumo

Este artigo apresenta um estudo da avaliação objetiva da qualidade do vídeo codificado no padrão H.264/SVC (Scalable Video Coding), transmitido por meio de redes de comunicação móvel de 3ª Geração (3G). O padrão H.264 permite o ajuste do bit rate de acordo com as condições da rede e pode ser utilizado em diferentes cenários, onde múltiplos fluxos de bits são enviados a partir de uma mesma fonte, com diferenciação na taxa de quadros e de bits por segundo. Com a ajuda de um esquema auto-adaptativo que considera feedbacks dos receptores, o conteúdo codificado na origem é então transmitido de acordo com as condições de rede e capacidade do terminal, permitindo avaliar o desempenho do esquema adaptativo para streaming de vídeo escalável por meio de redes 3G.

Abstract

This article presents a study of objective evaluation of the quality of video encoded in H.264/SVC (Scalable Video Coding) standard, transmitted via 3rd Generation (3G) mobile communication networks. The H.264 standard allows adjustment of the bit rate according to network condition and can be used in different scenarios, where multiple streams of bits are sent from one source, with differentiation in frame rate and bit per second. With the help of a self-adaptive scheme which considers feedback from receivers, the encoded content is then transmitted according to network conditions and terminal capacities, allowing to evaluate the performance of the adaptive scheme for scalable video streaming through 3G.

Palavras-chaves: Vídeo escalável, redes 3G, QoV.

1. Introdução

A transmissão de vídeo em tempo real gera uma série de novos desafios. Alguns desses desafios incluem garantir qualidade de serviços para os diferentes receptores, utilizar ao máximo a capacidade de cada receptor, prover o melhor aproveitamento de banda permanecendo em níveis aceitáveis de qualidade, evitar ou reagir rapidamente a congestionamentos na rede e atender a variações repentinas do número de usuários.

Nestas circunstâncias, as aplicações de transmissão de vídeo precisam se adaptar às mudanças das condições da rede e às diferentes demandas. Assim, elas podem adotar técnicas escaláveis e adaptativas tanto em termos de conteúdo quanto em termos de taxas de transmissão para poderem lidar com as variações de tráfego da rede e heterogeneidades de dispositivos.

O objetivo deste trabalho é avaliar a qualidade do vídeo transmitido em uma rede 3G, utilizando um esquema adaptativo de transmissão de vídeo que faz uso de um codificador de vídeo escalável, baseado no padrão H.264/SVC.

O presente trabalho está dividido da seguinte forma: na seção 2 são apresentados os trabalhos relacionados, juntamente com a contribuição desses para o desenvolvimento deste artigo. Na seção 3 são apresentados os aspectos da transmissão de vídeo escalável, bem como o padrão utilizado para a transmissão de *streaming* de vídeo em redes de 3ª geração. Na seção 4 é apresentado o esquema adaptativo utilizado na transmissão do *streaming* de vídeo, com detalhamento dos componentes que compõem o servidor e o cliente do esquema. Na seção 5 são apresentados o ambiente de teste e a análise dos resultados obtidos nos experimentos. Na seção 6 são apresentadas conclusões e indicados trabalhos futuros.

2. Trabalhos relacionados

No artigo [Alexiou *et al.*, 2005] foi projetado e implementado um esquema de decisão com *feedback*, baseado no RTP (*Real-Time Protocol*) e utilizando vídeo *streaming* sobre sistema UMTS (*Universal Mobile Telecommunication System*), para sequência de vídeo codificada em H.264/AVC.

A qualidade de experiência de serviço de streaming de vídeo *unicast* em redes 3G móvel e os métodos de acesso foram avaliados em [Davies *et al.*, 2008]. Para o estudo de caso foi utilizado comutação de pacotes padrão, comutação por circuito padrão e código de acesso curto. O objetivo foi avaliar se streaming por comutação por circuito pode concorrer com serviços de comutação por pacote, dando ênfase a métodos de acesso particular. O método utilizado para avaliação da qualidade de experiência foi o MOS (*Mean Opinion Score*), foram utilizados dispositivos móveis e três diferentes sequências de vídeo. Foi verificado que método de acesso particular é um fator que exerce influência na recepção de serviço de vídeo móvel.

O foco do artigo em [Hartung *et al.*, 2007] foi em modelos de tráfego, sendo discutido o limite da capacidade das redes 3G para distribuição *unicast/broadcast* de TV móvel. Foram apresentadas a tecnologia do 3GPP chamada de MBMS (*Multimedia Broadcast Multicast Service*) e uma arquitetura de entrega de conteúdo *unicast/broadcast* híbrida para transporte de mídia.

Em [Liu *et al.*, 2009] foi avaliada a transmissão de vídeo SVC em redes 3G, com base em simulação por software e transferência de pacotes no formato RTP. A transmissão foi feita via simulação utilizando canais de rede 3G e serviço de multimídia *broadcast/multicast*. Diferentemente do que foi proposto por [Liu *et al.*, 2009], o esquema adaptativo utilizado neste artigo, além de transferir pacotes via RTP, trata pacotes de controles com *feedback* do cliente para o servidor transmitidos via RTCP (*Real Time Control Protocol*) e a forma de transmissão é via *unicast*.

O artigo [Fajardo *et al.*, 2007] concentra-se na transmissão de vídeo H.264/AVC sobre uma rede UMTS. A principal contribuição do artigo está em apresentar resultados utilizando diferentes tamanhos de *slices* (fatias da imagem) durante o processo de codificação, sendo o ambiente proposto simulado no OPNET.

No artigo [Singh *et al.*, 2009] foram verificados diferentes mecanismos de adaptação de fluxo e redefinidos para redes 3GPP reutilizando o RTCP existente e o TFRC (*TCP Friendly Rate Control*). No simulador NS-2, foi feita uma simulação do núcleo da rede 3G e de uma interface Nokia. O esquema utilizado neste artigo possui um controle de fluxo baseado no protocolo RTCP e a transmissão é feita via protocolo UDP em um ambiente real, utilizando a rede 3G de uma operadora.

Uma análise de desempenho de serviço de *streaming* de vídeo foi feita em [Diaz *et al.*, 2007], com o objetivo de fornecer um ambiente usando um serviço móvel real. Foram utilizados vídeos codificados em H.263 e MPEG-4 visual. Dois cenários foram propostos para o experimento utilizando rede GPRS (*General Packet Radio Service*) e rede UMTS.

O escopo do artigo [Ries *et al.*, 2008] foi estimar a qualidade subjetiva do vídeo para sequências de vídeos de baixa resolução, que é característico de *streaming* para dispositivos móveis. Foram apresentados dois métodos para determinar a qualidade, o primeiro método é baseado em parâmetros adaptativos de conteúdo, permitindo a estimação da qualidade dependendo do conteúdo. O segundo método é dividido em dois passos: primeiro há uma classificação do conteúdo; em seguida, com base na classe do conteúdo, taxa de quadros e *bitrate* é estimada a qualidade do vídeo.

Uma arquitetura foi proposta e validada em [Thomas and George, 2007], que explora a priorização de pacotes de vídeo escalável juntamente com um mapeamento do tráfego da rede UMTS. O cenário proposto, com base em *DiffServ* (para o gerenciamento do buffer *DiffServ* na rede foi utilizado o WRED - *Weighted Random Early Detection*) em um único *link* sem fio de 512 Kbps e uma única célula UMTS de 1Mbps, foi simulado utilizando o NS-2 (EURANE), oito sequências de vídeo diferentes com 300 a 2000 frames e sequências de vídeo no formato YUV (QCIF-176x144) codificada em MPEG-4 FGS e em H.264/SVC. Foi avaliada a qualidade do vídeo escalável e não escalável, com e sem priorização de tráfego.

As referências [Diaz *et al.*, 2007], [Fajardo *et al.*, 2007], [Kang *et al.*, 2007], [Thomas and George, 2007], [Liu *et al.*, 2009], [Mulinari e Gondim, 2009] adotam o PSNR (*Peak Signal-to-Noise Ratio*), que é uma métrica objetiva de qualidade de vídeo, para avaliação do vídeo. Neste artigo foi utilizado o Y-PSNR, combinado com as métricas de qualidade de serviço (QoS – *jitter* e perda de pacotes) obtidas no esquema adaptativo descrito na seção 4.

3. Aspectos da transmissão de vídeo escalável em redes 3G

3.1 Vídeo SVC

A escalabilidade de vídeo é fornecida com base em uma estrutura de codificação de vídeo em camadas, a partir da qual uma camada base (*base layer*) pode ser acrescida de uma ou mais camadas de refinamento ou reforço (*enhancement layers*).

Dentre os padrões de vídeo escalável, destaca-se o padrão H.264 SVC (*Scalability Video Coding*), que é uma extensão do padrão H.264/AVC (*Advanced Video Coding*) em que, a partir de uma camada base, é possível obter uma ou mais camadas de reforço, aumentando a resolução temporal, a resolução espacial e/ou a qualidade do conteúdo de vídeo.

Os tipos de escalabilidade de vídeo mais usuais são a escalabilidade temporal, a escalabilidade espacial e a escalabilidade SNR (*Signal-to-Noise Ratio*) ou de qualidade, conforme descrito a seguir [Thomas e George, 2007]:

- Escalabilidade temporal: a taxa de quadros da camada de reforço é maior se comparada com as taxas de quadros das camadas inferiores, inclusive com relação à camada base;
- Escalabilidade espacial: as camadas de reforço possuem uma resolução maior ou igual a da camada base ou se comparada com outras camadas inferiores: A escalabilidade espacial corresponde à possibilidade de se transmitir imagens com diferentes resoluções;
- Escalabilidade de qualidade (SNR): as camadas com resolução espacial e temporal permanecem as mesmas e apenas as camadas de qualidade são reforçadas. Quanto maior o SNR, maior será a qualidade da imagem produzida.

A escalabilidade de qualidade é dividida em três classes, conforme a granularidade, em grossa (*Coarse-Grain Quality Scalable Coding* - CGS), média (*Medium-Grain Quality Scalability* - MGS) e fina (*Fine-Grain Quality Scalable* - FGS):

- CGS: Os dados comprimidos são embutidos em um único *bitstream* para permitir que diferentes taxas de bits possam ser obtidas;
- MGS: É uma variação da CGS e fornece um aumento na flexibilidade de adaptação do fluxo de *bits*;
- FGS: o *bitstream* pode ser decodificado com um conjunto pequeno de *bits* e pré-estabelecido durante a codificação.

Para a codificação do vídeo utilizado neste artigo, foi utilizada a escalabilidade de granularidade média (MGS), bem como as escalabilidades temporal e espacial de forma combinada.

3.2 Serviço de *Streaming* no 3GPP

O padrão de *Streaming* comutado por pacotes (PSS – *Packet-Switched Streaming Standard*), desenvolvido pelo 3GPP (3rd *Generation Partnership Project*), é hoje o padrão mais usado para *streaming* baseado em *unicast* em redes celulares. PSS fornece um completo *framework* de *streaming* e *download* fim a fim para redes móveis que incluem desde servidores de arquivos (arquivos do servidor) para armazenamento de sessões de *streaming*, servidores de *streaming* para entrega, até clientes de *streaming* para recepção [Elsen *et al.*, 2001].

O principal escopo do PSS é definir uma aplicação que forneça *streaming* sincronizado, incluindo áudio, vídeo e texto. O PSS também padroniza um mecanismo opcional de proteção de confidencialidade e integridade para o conteúdo por meio de encriptação do *stream* e de arquivos. Outros destaques do padrão incluem seleção de mídia, onde um cliente pode escolher diferentes alternativas de taxas de bits e idiomas e reporte de qualidade de experiência (QoE), o qual dá ao fornecedor do serviço medidas para avaliar a experiência do usuário final. A Figura 1 apresenta a pilha de protocolos para *streaming* 3GPP.

Vídeo Áudio Texto sincronizado com outras mídias	Descrição do cenário Imagens Gráfico bitmap Vetor gráfico Texto	Apresentação da descrição
Formato do Payload	HTTP	RTSP
RTP		
UDP	TCP	UDP
IP		

Figura 1: Pilha de protocolos para streaming 3GPP.

Conforme apresentado na Figura 1, fazem parte da pilha do 3GPP o protocolo de streaming de tempo real (RTSP- *Real Time Streaming Protocol*), o protocolo de descrição de sessão (SDP – *Session Description Protocol*) para ferramentas de sessão e controle e o protocolo de transferência de tempo real (RTP – *Real Time Protocol*) para transporte de mídia em tempo real, tais como vídeo, voz (fala) e áudio. No esquema adaptativo apresentado neste artigo foram utilizados os protocolos RTP, RTCP e o UDP. Além da definição de mecanismos de transporte, o escopo do PSS também envolve o conjunto de codificadores de mídia.

O PSS inclui características, que levam ao melhoramento na qualidade relativa ao *streaming* sobre redes móveis. Uma dessas características é chamada de *streaming* adaptativo, que permite que os fluxos a serem transmitidos sejam adaptados as variações das condições da rede. Embora as redes de melhor esforço possam, eventualmente, prover a taxa de bits requerida, não há garantia da disponibilidade dessa taxa de bits requerida durante o tempo da sessão, especialmente em enlaces móveis, geralmente caracterizados por uma variação na vazão devido à natureza do canal sem fio [Hartung *et al.*, 2007].

O 3GPP fornece uma estrutura para integração entre os diferentes formatos de codificadores, incluindo o H.263, MPEG-4 Visual, H.264/AVC [3GPP TS 26.244, 2010], entretanto o 3GPP não fornece uma estrutura para integração entre o H.264/SVC e os demais codificadores, por isso fez-se necessário a utilização de um esquema adaptativo.

4. Esquema adaptativo

O esquema adaptativo proposto em [Mulinari e Gondim, 2009], foi implementado e comparado com outros dois esquemas ADIVIS [Antoniou *et al.*, 2007] e “Bouras e Gkamas” [Bouras *et al.*, 2003] com base, inicialmente, em redes WLAN. Os vídeos utilizados foram codificados no padrão H.264/SVC e o modelo de transmissão foi *unicast*. O mesmo esquema foi utilizado para os experimentos realizados em redes de 3ª Geração neste artigo e será detalhado nesta seção.

O esquema baseia-se na incorporação da lógica *fuzzy* e em *feedbacks* do cliente com informações das condições da rede para tomar decisões de ajuste na taxa de bits do vídeo transmitido. O conteúdo de vídeo e os pacotes de controles com os *feedbacks* do cliente para o servidor são transmitidos via protocolos RTP / RTCP. O papel do mecanismo de *feedback* e dos componentes de adaptação é o de associar a qualidade de vídeo desejada às condições de rede. A Figura 2 apresenta a arquitetura do esquema adaptativo.

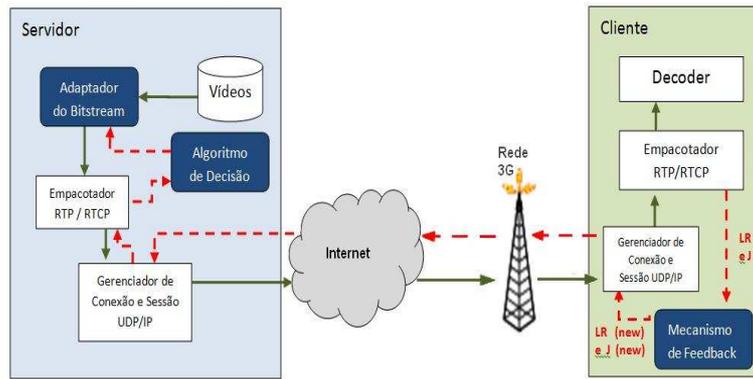


Figura 2: Arquitetura do esquema adaptativo.

O mecanismo para processamento da adaptação de rede utiliza um algoritmo *fuzzy* para decisão. O mecanismo de *feedback*, para sensoriamento, combina informações de perda de pacotes e *jitter* enviadas pelo receptor sobre a qualidade percebida, a fim de avaliar a situação atual da rede.

Os valores das variáveis que indicam a qualidade percebida na rede são introduzidos no algoritmo de decisão, que decide com base na melhora ou piora das condições de rede, conforme segue abaixo:

- Taxa de perda (LR- *Loss Rate*): O primeiro valor de *feedback*, a taxa de perda de pacotes, é calculado pela variação da medição atual de perda de pacote e a medição anterior LR;
- Atraso médio: O segundo parâmetro de *feedback* é calculado utilizando o valor gerado pelo *interval jitter* do pacote RR (*Receiver-Report*);
- Os dois parâmetros coletados pelo mecanismo de *feedback* são enviados para o servidor em pacotes do tipo RR do protocolo RTCP.

4.1 Componentes do Servidor

O servidor é composto de:

- Adaptador de *bitstream*: responsável pela adaptação de conteúdo, incluindo o gerenciamento da quantidade de camadas de vídeo que devem ser enviadas a partir da banda disponível;
- Algoritmo de decisão: recebe os *feedbacks* dos clientes (receptores), enviados em pacotes RTCP e decide pela manutenção, diminuição ou incremento da banda disponível;
- Empacotador RTP/UDP (*User Datagram Protocol*) / Desempacotador RTCP/UDP: encapsula os pacotes de vídeo, ou unidade NAL em pacotes no protocolo RTP/UDP; desencapsula os pacotes RTCP/UDP recebidos pelo gerenciador de conexão e sessão e encaminha para o algoritmo de decisão;
- Gerenciador de conexão e sessão: gerencia as conexões e sessões e encaminha os pacotes RTP/UDP e RTCP/UDP.

4.2 Componentes do Cliente

O cliente é composto de:

- Gerenciador de conexão e sessão: solicita a abertura de sessão com o servidor, envia pacotes de controle e recebe pacotes de vídeo;
- Mecanismo de *feedback*: recebe os pacotes e faz o sensoriamento das condições da rede.

Contabiliza os pacotes perdidos e o atraso dos pacotes durante o intervalo medido;

- Empacotador RTCP/Desempacotador RTP: encapsula os pacotes de controle e encaminhá-lo ao gerenciador de conexão e sessão; desencapsular os pacotes recebidos e armazenar os pacotes.

O intervalo de medição para computar e gerar um relatório com os valores coletados das condições de rede (*feedback*) é de 0,5 segundo e o envio de pacotes RRs é de 1 pacote a cada 0,5 segundo. O mecanismo de decisão informa ao adaptador de *bitstream* a nova largura de banda disponível e identifica as camadas que podem ser transmitidas na largura de banda disponível. O esquema faz uso de dois arquivos extraídos com a ajuda do aplicativo JSVM 9.17. O arquivo *video.bitrate* é utilizado para informar quais camadas podem ser transportadas com a banda disponível. Ele identifica a taxa de bits necessária para transportar cada camada. O arquivo de *trace* é utilizado para identificar os pacotes NAL (*Network Adaptation Layer*) que pertencem à camada escolhida ou a camadas inferiores.

5. Ambiente de teste e análise dos resultados

5.1 Cenário proposto

O cenário é composto de uma arquitetura cliente - servidor, implantada no laboratório de Televisão Digital Interativa da Universidade de Brasília (UnB), com um servidor conectado à Internet que armazena vídeo pré-codificado no formato H.264 SVC e um cliente conectado à internet através de um modem 3G. A Figura 3 apresenta a ilustração do cenário proposto para os experimentos.

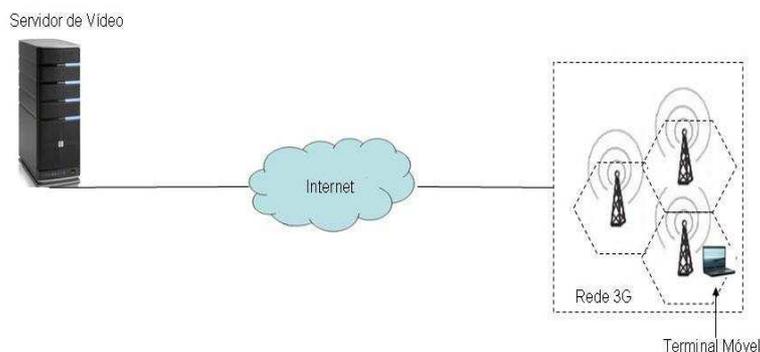


Figura 3: Cenário proposto para os experimentos.

O servidor é composto por um *desktop* configurado com o sistema operacional Windows 7, conectado à Internet. Nele está armazenado o aplicativo servidor do esquema adaptativo (implementado na linguagem Java) que recebe a mensagem de retorno (*feedback*) do cliente e faz a adaptação do conteúdo de acordo com as condições da rede. Nele também estão armazenadas todas as informações de variação de atraso (*jitter*) e perda de pacotes que ocorrem durante a transmissão da sequência de vídeo.

O cliente é composto por um *Notebook* marca HP, modelo Pavilion tx1070br, com processador AMD Turion 64, 1024 Kb de memória RAM e sistema operacional Windows Vista. Para ter acesso à rede 3G, o cliente está conectado via USB 1.0 a um modem ZTE modelo mf110, com acesso às redes GSM/GPRS/EDGE (850/900MHz), e HSUPA/HSDPA/UMTS (850/1.900/2.100 MHz). No cliente está armazenado o aplicativo

cliente do esquema adaptativo, implementado na linguagem Java e é através dele que são feitas as requisições e a recepção do vídeo.

Os testes foram realizados utilizando a rede 3G de uma operadora de telecomunicações, considerando que a velocidade máxima de navegação é de 1Mbps e, ao atingir 100% (cem por cento) da franquia contratada, que é de 1GB, o cliente permanecerá conectado, porém, com velocidade máxima de 128 Kbps.

A velocidade de navegação também pode ser diminuída dependendo das condições topográficas e de relevo, velocidade de movimento e distância do cliente em relação à Estação Rádio Base (ERB), número de usuários que utilizam ao mesmo tempo a cobertura fornecida pela ERB, disponibilidade e intensidade do sinal no local em que se utiliza o serviço, condições climáticas e razões técnicas de cobertura móvel, interferência da transmissão do sinal da ERB na Estação Móvel. O endereço IP atribuído à rede 3G não é fixo, sendo fornecido um novo endereço a cada nova conexão. A rede alocada pela operadora oscila entre as redes HSPA, UMTS, GSM e EDGE. Os testes realizados neste trabalho foram feitos considerando estações fixas.

Para o esquema apresentado neste trabalho não foram utilizados *slices* de tamanhos diferentes, porém foi utilizado, em uma rede 3G UMTS, o vídeo codificado em H.264/SVC com diferentes combinações de camadas. Ao contrário do que foi utilizado em [Fajardo *et al.*, 2007], aqui, utilizaremos um ambiente real, com tráfego real na rede de uma operadora de telecomunicações móveis.

5.2 Vídeo

O vídeo utilizado para o experimento é o *Highway* com resolução CIF (352x288), com taxa de 30 quadros por segundo, com um total de 2.000 *frames* (quadros). O vídeo foi codificado e decodificado utilizando o padrão H.264/SVC com escalabilidade temporal, espacial e SNR (qualidade). Para codificação do vídeo foram utilizadas cinco camadas de reforço temporais, três camadas de reforço SNR (qualidade) e duas camadas de reforço espaciais (resolução). A ferramenta JSVM foi utilizada para a codificação e decodificação do vídeo.

5.3 Métrica para aferição da QoV

Após recebido, o vídeo é decodificado com o auxílio da ferramenta JSVM no formato YUV para ser comparado com o vídeo original e assim ser obtida sua QoV (*Quality of Video*) com base no valor de Y-PSNR em dB.

Considera-se o seguinte mapeamento entre os valores de PSNR e a escala relativa à avaliação subjetiva (MOS – *Mean Opinion Score*), obtida de [Chih-Heng Ke *et al.*, 2008] apresentada na Tabela 1:

PSNR[dB]	MOS	Grau deficiência da Imagem
> 37	5 (Excelente)	Imperceptível
31-37	4 (Boa)	Perceptível mas não incômodo
25-31	3 (Apropriada)	Pouco Incômodo
20-25	2 (Pobre)	Incômodo
< 20	1 (Ruim)	Muito Incômodo

Tabela 1: Relação entre PSNR e MOS.

5.4 Experimentos e análise dos resultados

As transmissões foram realizadas sempre por iniciativa do cliente (vídeo sob demanda). O vídeo foi transmitido 24 (vinte e quatro) horas por dia, durante um período de 15 (quinze) dias. Cada transmissão possui uma duração de quase 25 minutos, até serem transmitidos todos os pacotes ao cliente e encerrada a sessão.

Após as transmissões, os vídeos decodificados no cliente (formato YUV) foram coletados para análise. Esta análise consiste no cálculo da média Y-PSNR, bem como no cálculo do desvio padrão do Y-PSNR medido. O gráfico 1 apresenta a média dos resultados obtidos durante os experimentos:

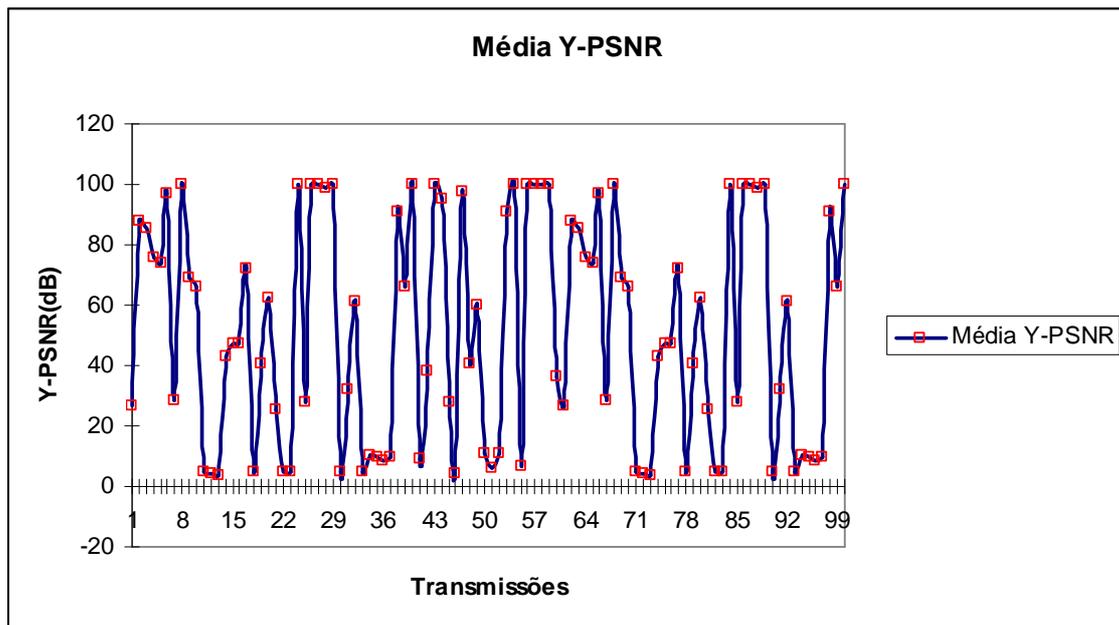


Gráfico 1: Média Y-PSNR dos vídeos transmitidos.

O gráfico 1 apresenta no eixo das abscissas as 100 (cem) transmissões realizadas, enquanto no eixo das ordenadas apresenta as médias Y-PSNR obtidas durante as transmissões de cada vídeo. Foi verificado que durante as transmissões dos primeiros vídeos o Y-PSNR chegou a ficar bem acima de 37 dB, que conforme Tabela 1, é considerado, na escala MOS, excelente levando a um grau imperceptível de deficiência da imagem. No decorrer das transmissões foi observada uma queda do Y-PSNR ficando entre 20 dB a 10 dB devido às várias oscilações da rede. Entretanto os resultados mostraram-se satisfatórios, pois foi verificado que o Y-PSNR esteve acima de 37 dB na maior parte das transmissões. O gráfico 2 apresenta a média do Y-PSNR e o desvio padrão de acordo com o horário de cada transmissão.

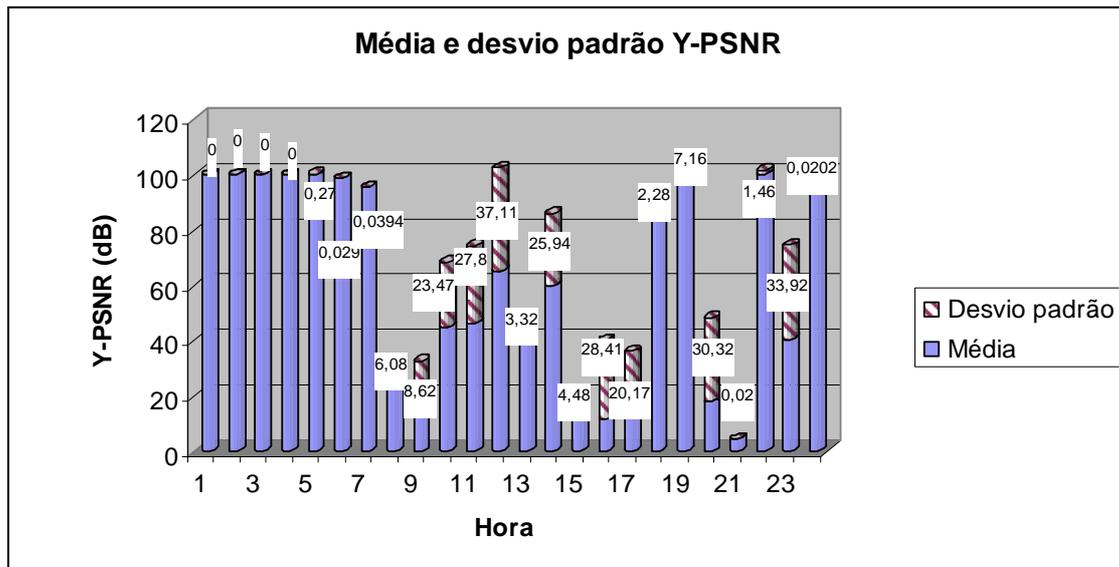


Gráfico 2: Média e desvio padrão Y-PSNR dos vídeos transmitidos.

O gráfico 2 apresenta no eixo das abscissas os horários de realização das transmissões das sequências de vídeo, enquanto no eixo das ordenadas apresenta as médias Y-PSNR e o desvio padrão obtidos durante as transmissões de cada vídeo. Os melhores resultados foram obtidos no período entre 1h da manhã às 7h da manhã, quando o Y-PSNR chegou a 99,99 dB e não houve grandes variações na medida de qualidade de vídeo adotada, isso ocorreu devido às poucas oscilações e variações no tráfego da rede. As maiores variações na média do Y-PSNR ocorreram entre 8h da manhã às 18h. Foi verificado que os menores valores obtidos do Y-PSNR foram às 15h com média de 15,97 dB, às 16h com média de 11,80 dB, às 17h com média de 16,06 dB e às 21h com média de 4,61 dB. De acordo com [Chih-Heng Ke *et al.*, 2008] esses últimos são considerados, na escala MOS, ruim (com grau de deficiência da imagem muito incômoda).

Os valores em destaque no gráfico correspondem aos valores do desvio padrão obtidos no experimento. De acordo com o cenário proposto para os experimentos, foi observado que o valor do desvio padrão variou em quase todas as transmissões realizadas. Os menores valores do desvio padrão foram às 5h, 6h, 7h da manhã, 15h, 18h, 21h, [Wiegand et al 2003] h e 24h que foram, respectivamente, de 0,27 dB, 0,029 dB, 0,0394 dB, 2,28 dB, 0,02 dB, 1,46 dB e 0,020 dB. Os maiores valores do desvio padrão foram às 10h, 11h, 12h, 14h, 16h, 20h e 23h, sendo, respectivamente, de 23,47 dB, 27,80 dB, 37,11 dB, 25,94 dB, 28,41 dB, 30,32 dB e 33,92 dB. Não houve variações nos valores do desvio padrão entre 1h e 4h da manhã. Pode-se notar que o maior desvio padrão foi às 12h, porém a média do Y-PSNR ficou acima de 37 dB, o que é considerado excelente segundo [Chih-Heng Ke *et al.*, 2008].

6. Conclusão

Neste trabalho foram apresentados os resultados obtidos com a transmissão de vídeo escalável em redes de 3ª geração. Com a utilização do esquema adaptativo, foi possível adaptar a taxa de transmissão do streaming de vídeo de acordo com as mudanças e oscilações ocorridas na rede e avaliar o desempenho do vídeo em redes sem fio de 3ª geração, garantindo assim uma qualidade de vídeo aceitável para o usuário, mesmo em horários considerados de pico e grande tráfego na rede. Com a análise feita dos resultados deste trabalho, pode-se concluir que o esquema adaptativo funciona de maneira eficiente e se adapta às oscilações e mudanças das condições de rede.

Para trabalhos futuros, deve ser criada uma quantidade maior de sequências de vídeo, tanto em classificação e características de cenas quanto em taxas de quadros e diferentes combinações de escalabilidade. Outras sugestões são: a avaliação da qualidade subjetiva de vídeo utilizando a métrica subjetiva MOS (*Mean Opinion Score*) e a recepção de vídeo utilizando diferentes receptores, tais como: PDA's e *smartphones*.

Referências

[3GPP TS 26.244, 2010] 3gpp ts 26.244 (2010-06) “*Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS)*”; 3GPP file format (3GP) (Release 9) V9.2.0.

[Alexiou *et al.*, 2005] Alexiou, A.G. Bouras, C.J. Igglesis, V.G. (2005) “A decision feedback scheme for multimedia transmission over 3G mobile networks”. Second IFIP International Conference on Wireless and Optical Communications Networks. WOCN.pg 357-361.

[Antoniou *et al.*, 2007] Antoniou P., Vassiliou V., Pitsilides A. (2007) “ADIVIS: A Novel Adaptive Algorithm for Video Streaming over the Internet”. 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.

[Bouras *et al.*, 2003] Bouras C., Gkamas A. (2003) “Performance of Adaptive Multimedia Transmission: The case of Unicast Technique”, Research Academic Computer Technology Institute, Greece.

[Chih-Heng Ke *et al.*, 2008] Chih-Heng Ke, Ce-Kuen Shieh, Wen-Shyang Hwang, Artur Ziviani (2008) “An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission” *Journal of Information Science And Engineering* 24, 425-440.

[Davies *et al.*, 2008] Davies, M., Dantcheva, A., Fröhlich, P. (2008) “Comparing Access Methods and Quality of 3G Mobile Video Streaming Services”. *Proceeding of the twenty-sixth annual CHI conference extended abstracts on Human factors in computing systems - CHI '08*, pg 2817.

[Diaz *et al.*, 2007] Diaz A., Merino P., Panizo L., Recio, A. M. (2007) “Evaluating Video Streaming over GPRS / UMTS networks: A practical case”. *IEEE 65th Vehicular Technology Conference - VTC2007-Spring*. 2007. pg 624-628.

[Elsen *et al.*, 2001] Elsen, I., Hartung, F., Horn, U., Kampmann, M., Peters, L. (2001) “Streaming Technology in 3G Mobile Communication Systems”. *IEEE Computer Society Press. Computer Journal*. Vol. 34. Pg. 46-52. September.

[Fajardo *et al.*, 2007] Fajardo J. O., Liberal F., Bilbao N. (2007) “Impact of the video slice size on the visual quality for H.264 over 3G UMTS services”. *Methodology*.

[Hartung *et al.*, 2007] Hartung, F., Horn, U., Huschke J., Kampmann M., Lohmar T., Lundevall M., (2007) “Delivery of Broadcast Services in 3G Networks”. *IEEE Transactions on Broadcasting*. Vol. 53. Pg 188-199. March.

[Iain Richardson, 2003] Iain E. G. Richardson (2003) “H.264 and MPEG-4 Video Compression Video Coding for Next-generation Multimedia”, WILEY.

[Kang *et al.*, 2007] Kang, K., Cho, Y., Cho, J., Shin, H. (2007) “Scheduling Scalable Multimedia Streams for 3G Cellular Broadcast and Multicast Services” IEEE Transactions on Vehicular Technology, Vol. 56. Pg 2655-2672.

[Liu *et al.*, 2009] Liu, Y., Zhang S., Xu S., Yin Z. (2009) “Research on H.264/SVC Compressed Video Communication in 3G.” 4th International Conference on Computer Science e Education, IEEE. pg 327-332.

[Mulinari e Gondim, 2009] Mulinari R., Gondim P. R. L. (2009) “Esquemas Adaptativos para Distribuição de Vídeo na Internet”. Dissertação de Mestrado – Departamento de Engenharia Elétrica. UnB.

[Ries *et al.*, 2008] Ries, M., Nemethova, O., Rupp, M. (2008) “Video Quality Estimation for Mobile H.264/AVC Video Streaming”. Journal of Communications. Vol. 3. pg 41-50.

[Schwarz and Wien, 2008] Schwarz, H., and Wien, M. (2008) “The Scalable Video Coding Extension of the H.264/AVC Standard”. IEEE Signal Processing Magazine p.135 March.

[Singh *et al.*, 2009] Singh, V., Ott, J., Curcio, I.D.D (2009)“Rate Adaptation for Conversational 3G Video”. IEEE Infocom Workshops. pg 1-7.

[Thomas and George, 2007] Thomas, P., George K. (2007)“Scalable Video Streaming Traffic Delivery in IP/UMTS Networking Environments”, Journal of Multimedia, Vol. 2, Nr. 2, Abril.

APÊNDICE II

II.1– Arquivos de configuração dos vídeos - cfg

Este apêndice apresenta os arquivos de configurações utilizados para a codificação de cada vídeo, juntamente com suas respectivas camadas.

Arquivo de configuração do vídeo

```
# Scalable H.264/AVC Extension Configuration File
```

```
# apostila 2 3 camadas
```

```
##### GENERAL #####
OutputFile          SVCContainer.264  # Bitstream file
FrameRate           30              # Maximum frame rate [Hz]
FramesToBeEncoded   300            # Number of frames (at input frame
rate)
MaxDelay            1200.0          # Maximum structural delay [ms]
NonRequiredEnable   0              # NonRequiredSEI enable (0:disable,
1:enable)
CgsSnrRefinement    1              # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures   1              # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl          1              # ME/MC for non-key pictures in MGS
layers
                                # (0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize             16              # GOP Size (at maximum
frame rate)
IntraPeriod         16              # Intra Period
NumberReferenceFrames 1            # Number of reference
pictures
BaseLayerMode       1              # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

##### MOTION SEARCH #####
SearchMode          4              # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel   0              # Search function full
pel
                                # (0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)
SearchFuncSubPel    0              # Search function sub
pel
                                # (0:SAD, 1:SSE,
2:HADAMARD)
SearchRange         16              # Search range (Full
Pel)
BiPredIter          2              # Max iterations for
bi-pred search
IterSearchRange     2              # Search range for
iterations (0: normal)
```

```

#===== LOOP FILTER
=====
LoopFilterDisable      0          # Loop filter idc (0:
on, 1: off, 2:          #   on except for slice
                           boundaries)
LoopFilterAlphaC0Offset 0          # AlphaOffset(-6..+6):
valid range              #
LoopFilterBetaOffset   0          # BetaOffset (-6..+6):
valid range              #

#===== LAYER DEFINITION
=====
NumLayers              6          # Number of layers
LayerCfg               SVCContainerQCIFlayer0.cfg # Layer
configuration file
LayerCfg               SVCContainerQCIFlayer1.cfg # Layer
configuration file
LayerCfg               SVCContainerQCIFlayer2.cfg # Layer
configuration file
LayerCfg               SVCContainerCIFlayer3.cfg # Layer
configuration file
LayerCfg               SVCContainerCIFlayer4.cfg # Layer
configuration file
LayerCfg               SVCContainerCIFlayer5.cfg # Layer
configuration file

PreAndSuffixUnitEnable 1          # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable        1          # MMCO for base
representation (0: off, 1: on)
TLNestingFlag         1          # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable 1          # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD          0
EnableVclHRD          0

# Scalable H.264/AVC Extension Configuration File

# apostila 2 3 camadas

#===== GENERAL =====
OutputFile            SVCFlower.264 # Bitstream file
FrameRate             30          # Maximum frame rate [Hz]
FramesToBeEncoded     250        # Number of frames (at input frame
rate)
MaxDelay              1200.0     # Maximum structural delay [ms]
NonRequiredEnable     0          # NonRequiredSEI enable (0:disable,
1:enable)

```

```

CgsSnrRefinement          1          # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures         1          # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl                1          # ME/MC for non-key pictures in MGS
layers
                                # (0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize                   16          # GOP Size (at maximum
frame rate)
IntraPeriod               16          # Intra Period
NumberReferenceFrames     1          # Number of reference
pictures
BaseLayerMode             1          # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

#===== MOTION SEARCH
#=====
SearchMode                4          # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel        0          # Search function full
pel
                                # (0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)
SearchFuncSubPel         0          # Search function sub
pel
                                # (0:SAD, 1:SSE,
2:HADAMARD)
SearchRange               16          # Search range (Full
Pel)
BiPredIter                2          # Max iterations for
bi-pred search
IterSearchRange           2          # Search range for
iterations (0: normal)

#===== LOOP FILTER
#=====
LoopFilterDisable         0          # Loop filter idc (0:
on, 1: off, 2:
boundaries)
                                # on except for slice
LoopFilterAlphaC0Offset  0          # AlphaOffset(-6..+6):
valid range
LoopFilterBetaOffset      0          # BetaOffset (-6..+6):
valid range

#===== LAYER DEFINITION
#=====
NumLayers                 6          # Number of layers
LayerCfg                  SVCFlowerQCIFlayer0.cfg # Layer configuration
file
LayerCfg                  SVCFlowerQCIFlayer1.cfg # Layer configuration
file
LayerCfg                  SVCFlowerQCIFlayer2.cfg # Layer configuration
file
LayerCfg                  SVCFlowerCIFlayer3.cfg # Layer configuration
file
LayerCfg                  SVCFlowerCIFlayer4.cfg # Layer configuration
file

```

```

LayerCfg          SVCFlowerCIFlayer5.cfg # Layer configuration
file

PreAndSuffixUnitEnable 1 # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable 1 # MMCO for base
representation (0: off, 1: on)
TLNestingFlag 1 # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable 1 # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD 0
EnableVclHRD 0

# Scalable H.264/AVC Extension Configuration File

# apostila 2 3 camadas

#===== GENERAL =====
OutputFile SVCFootball.264 # Bitstream file
FrameRate 30 # Maximum frame rate [Hz]
FramesToBeEncoded 90 # Number of frames (at input frame
rate)
MaxDelay 1200.0 # Maximum structural delay [ms]
NonRequiredEnable 0 # NonRequiredSEI enable (0:disable,
1:enable)
CgsSnrRefinement 1 # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures 1 # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl 1 # ME/MC for non-key pictures in MGS
layers
# (0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize 16 # GOP Size (at maximum
frame rate)
IntraPeriod 16 # Intra Period
NumberReferenceFrames 1 # Number of reference
pictures
BaseLayerMode 1 # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

#===== MOTION SEARCH
=====
SearchMode 4 # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel 0 # Search function full
pel
# (0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)

```

```

SearchFuncSubPel      0          # Search function sub
pel
                                # (0:SAD, 1:SSE,
                                # 2:HADAMARD)
SearchRange           16         # Search range (Full
Pel)
BiPredIter            2          # Max iterations for
bi-pred search
IterSearchRange       2          # Search range for
iterations (0: normal)

#===== LOOP FILTER
=====
LoopFilterDisable     0          # Loop filter idc (0:
on, 1: off, 2:
                                # on except for slice
boundaries)
LoopFilterAlphaC0Offset 0        # AlphaOffset(-6..+6):
valid range
LoopFilterBetaOffset  0          # BetaOffset (-6..+6):
valid range

#===== LAYER DEFINITION
=====
NumLayers              6          # Number of layers
LayerCfg               SVCFootballQCIFlayer0.cfg # Layer
configuration file
LayerCfg               SVCFootballQCIFlayer1.cfg # Layer
configuration file
LayerCfg               SVCFootballQCIFlayer2.cfg # Layer
configuration file
LayerCfg               SVCFootballCIFlayer3.cfg # Layer configuration
file
LayerCfg               SVCFootballCIFlayer4.cfg # Layer configuration
file
LayerCfg               SVCFootballCIFlayer5.cfg # Layer configuration
file

PreAndSuffixUnitEnable 1          # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable         1          # MMCO for base
representation (0: off, 1: on)
TLNestingFlag          1          # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable  1          # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD           0
EnableVclHRD           0

# Scalable H.264/AVC Extension Configuration File

# apostila 2 3 camadas

```

```

#===== GENERAL =====
OutputFile          Videoreult/SVCHighway.264  # Bitstream file
FrameRate          30                        # Maximum frame rate [Hz]
FramesToBeEncoded  2000                      # Number of frames (at input frame
rate)
MaxDelay           1200.0                    # Maximum structural delay [ms]
NonRequiredEnable  0                        # NonRequiredSEI enable (0:disable,
1:enable)
CgsSnrRefinement  1                        # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures 1                        # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl         1                        # ME/MC for non-key pictures in MGS
layers
# (0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize            16                        # GOP Size (at maximum
frame rate)
IntraPeriod        16                        # Intra Period
NumberReferenceFrames 1                    # Number of reference
pictures
BaseLayerMode      1                        # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

#===== MOTION SEARCH
=====
SearchMode         4                        # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel  0                        # Search function full
pel
# (0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)
SearchFuncSubPel   0                        # Search function sub
pel
# (0:SAD, 1:SSE,
2:HADAMARD)
SearchRange        16                        # Search range (Full
Pel)
BiPredIter         2                        # Max iterations for
bi-pred search
IterSearchRange    2                        # Search range for
iterations (0: normal)

#===== LOOP FILTER
=====
LoopFilterDisable  0                        # Loop filter idc (0:
on, 1: off, 2:
# on except for slice
boundaries)
LoopFilterAlphaC0Offset 0                    # AlphaOffset(-6..+6):
valid range
LoopFilterBetaOffset 0                      # BetaOffset (-6..+6):
valid range

#===== LAYER DEFINITION
=====
NumLayers          6                        # Number of layers
LayerCfg           cfg/SVCHighwayQCIFlayer0.cfg # Layer
configuration file

```

```

LayerCfg          cfg/SVCHighwayQCIFlayer1.cfg # Layer
configuration file
LayerCfg          cfg/SVCHighwayQCIFlayer2.cfg # Layer
configuration file
LayerCfg          cfg/SVCHighwayCIFlayer3.cfg # Layer
configuration file
LayerCfg          cfg/SVCHighwayCIFlayer4.cfg # Layer
configuration file
LayerCfg          cfg/SVCHighwayCIFlayer5.cfg # Layer
configuration file

PreAndSuffixUnitEnable 1 # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable 1 # MMCO for base
representation (0: off, 1: on)
TLNestingFlag 1 # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable 1 # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD 0
EnableVclHRD 0

# Scalable H.264/AVC Extension Configuration File

# Apostila2
# 3 camadas de quallidade
# com camadas espaciais CIF e QCIF

#===== GENERAL =====
OutputFile          Videoresult/SVCParis.264 # Bitstream file
FrameRate          30 # Maximum frame rate [Hz]
FramesToBeEncoded  1065 # Number of frames (at input frame
rate)
MaxDelay           1200.0 # Maximum structural delay [ms]
NonRequiredEnable  0 # NonRequiredSEI enable (0:disable,
1:enable)
CgsSnrRefinement   1 # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures  1 # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl         1 # ME/MC for non-key pictures in MGS
layers
# (0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize            16 # GOP Size (at maximum
frame rate)
IntraPeriod        16 # Intra Period
NumberReferenceFrames 1 # Number of reference
pictures
BaseLayerMode      1 # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

```

```

#===== MOTION SEARCH
=====
SearchMode          4          # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel  0          # Search function full
pel
# (0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)
SearchFuncSubPel   0          # Search function sub
pel
# (0:SAD, 1:SSE,
2:HADAMARD)
SearchRange        16         # Search range (Full
Pel)
BiPredIter         2          # Max iterations for
bi-pred search
IterSearchRange    2          # Search range for
iterations (0: normal)

#===== LOOP FILTER
=====
LoopFilterDisable  0          # Loop filter idc (0:
on, 1: off, 2:
boundaries)
# on except for slice
LoopFilterAlphaC0Offset 0      # AlphaOffset(-6..+6):
valid range
LoopFilterBetaOffset 0        # BetaOffset (-6..+6):
valid range

#===== LAYER DEFINITION
=====
NumLayers          6          # Number of layers
LayerCfg           cfg/SVCParisQCIFlayer0.cfg # Layer
configuration file
LayerCfg           cfg/SVCParisQCIFlayer1.cfg # Layer
configuration file
LayerCfg           cfg/SVCParisQCIFlayer2.cfg # Layer
configuration file
LayerCfg           cfg/SVCParisCIFlayer3.cfg # Layer
configuration file
LayerCfg           cfg/SVCParisCIFlayer4.cfg # Layer
configuration file
LayerCfg           cfg/SVCParisCIFlayer5.cfg # Layer
configuration file

PreAndSuffixUnitEnable 1      # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable     1          # MMCO for base
representation (0: off, 1: on)
TLNestingFlag      1          # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable 1       # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD       0
EnableVclHRD       0

```

Scalable H.264/AVC Extension Configuration File

apostila 2 3 camadas

#===== GENERAL =====

OutputFile SVCWaterfall.264 # Bitstream file
FrameRate 30 # Maximum frame rate [Hz]
FramesToBeEncoded 260 # Number of frames (at input frame
rate)
MaxDelay 1200.0 # Maximum structural delay [ms]
NonRequiredEnable 0 # NonRequiredSEI enable (0:disable,
1:enable)
CgsSnrRefinement 1 # SNR refinement as 1: MGS; 0: CGS
EncodeKeyPictures 1 # Key pics at T=0 (0:none, 1:MGS,
2:all)
MGSControl 1 # ME/MC for non-key pictures in MGS
layers
(0:std, 1:ME with EL, 2:ME+MC
with EL)

GOPSize 16 # GOP Size (at maximum
frame rate)
IntraPeriod 16 # Intra Period
NumberReferenceFrames 1 # Number of reference
pictures
BaseLayerMode 1 # Base layer mode
(0,1:AVC compatible, 2:AVC w subseq SEI)

#===== MOTION SEARCH =====

SearchMode 4 # Search mode
(0:BlockSearch, 4:FastSearch)
SearchFuncFullPel 0 # Search function full
pel
(0:SAD, 1:SSE,
2:HADAMARD, 3:SAD-YUV)
SearchFuncSubPel 0 # Search function sub
pel
(0:SAD, 1:SSE,
2:HADAMARD)
SearchRange 16 # Search range (Full
Pel)
BiPredIter 2 # Max iterations for
bi-pred search
IterSearchRange 2 # Search range for
iterations (0: normal)

#===== LOOP FILTER =====

LoopFilterDisable 0 # Loop filter idc (0:
on, 1: off, 2:
on except for slice
boundaries)
LoopFilterAlphaC0Offset 0 # AlphaOffset(-6..+6):
valid range

```

LoopFilterBetaOffset    0                # BetaOffset (-6..+6):
valid range

#===== LAYER DEFINITION
=====
NumLayers                6                # Number of layers
LayerCfg                 SVCWaterfallQCIFlayer0.cfg # Layer
configuration file
LayerCfg                 SVCWaterfallQCIFlayer1.cfg # Layer
configuration file
LayerCfg                 SVCWaterfallQCIFlayer2.cfg # Layer
configuration file
LayerCfg                 SVCWaterfallCIFlayer3.cfg # Layer
configuration file
LayerCfg                 SVCWaterfallCIFlayer4.cfg # Layer
configuration file
LayerCfg                 SVCWaterfallCIFlayer5.cfg # Layer
configuration file

PreAndSuffixUnitEnable  1                # Add prefix and suffix
unit (0: off, 1: on) shall always be on in SVC contexts (i.e. when
there are FGS/CGS/spatial enhancement layers)
MMCOBaseEnable          1                # MMCO for base
representation (0: off, 1: on)
TLNestingFlag           1                # Sets the temporal
level nesting flag (0: off, 1: on)
TL0DepRepIdxSeiEnable  1                # Sends dependency
representation index sei (0: off, 1: on)

#===== HRD
=====
EnableNalHRD            0
EnableVclHRD            0

```

APÊNDICE III

III.1 - Arquivos de configuração do Apache

Este apêndice apresenta o arquivo de configuração utilizado pelo servidor apache já com o módulo H.264 habilitado.

Configuração do arquivos httpd.conf

```
#
# Dynamic Shared Object (DSO) Support
#
# To be able to use the functionality of a module which was built as a DSO you
# have to place corresponding `LoadModule' lines at this location so the
# directives contained in it are actually available _before_ they are used.
# Statically compiled modules (those listed by `httpd -l') do not need
# to be loaded here.
#
# Example:
# LoadModule foo_module modules/mod_foo.so
#
LoadModuleh264_streaming_module
/usr/lib/apache2/modules/mod_h264_streaming.so
LoadModule actions_module modules/mod_actions.so
LoadModule alias_module modules/mod_alias.so
LoadModule asis_module modules/mod_asis.so
LoadModule auth_basic_module modules/mod_auth_basic.so
#LoadModule auth_digest_module modules/mod_auth_digest.so
#LoadModule authn_anon_module modules/mod_authn_anon.so
#LoadModule authn_dbm_module modules/mod_authn_dbm.so
LoadModule authn_default_module modules/mod_authn_default.so
LoadModule authn_file_module modules/mod_authn_file.so
#LoadModule authz_dbm_module modules/mod_authz_dbm.so
LoadModule authz_default_module modules/mod_authz_default.so
LoadModule authz_groupfile_module modules/mod_authz_groupfile.so
LoadModule authz_host_module modules/mod_authz_host.so
LoadModule authz_user_module modules/mod_authz_user.so
LoadModule autoindex_module modules/mod_autoindex.so
#LoadModule cern_meta_module modules/mod_cern_meta.so
LoadModule cgi_module modules/mod_cgi.so
#LoadModule dav_module modules/mod_dav.so
#LoadModule dav_fs_module modules/mod_dav_fs.so
#LoadModule deflate_module modules/mod_deflate.so
LoadModule dir_module modules/mod_dir.so
LoadModule env_module modules/mod_env.so
#LoadModule expires_module modules/mod_expires.so
#LoadModule file_cache_module modules/mod_file_cache.so
#LoadModule headers_module modules/mod_headers.so
```

```

LoadModule imagemap_module modules/mod_imagemap.so
LoadModule include_module modules/mod_include.so
#LoadModule info_module modules/mod_info.so
LoadModule isapi_module modules/mod_isapi.so
LoadModule log_config_module modules/mod_log_config.so
LoadModule mime_module modules/mod_mime.so
#LoadModule mime_magic_module modules/mod_mime_magic.so
#LoadModule proxy_module modules/mod_proxy.so
#LoadModule proxy_ajp_module modules/mod_proxy_ajp.so
#LoadModule proxy_balancer_module modules/mod_proxy_balancer.so
#LoadModule proxy_connect_module modules/mod_proxy_connect.so
#LoadModule proxy_http_module modules/mod_proxy_http.so
#LoadModule proxy_ftp_module modules/mod_proxy_ftp.so
LoadModule negotiation_module modules/mod_negotiation.so
#LoadModule rewrite_module modules/mod_rewrite.so
LoadModule setenvif_module modules/mod_setenvif.so
#LoadModule speling_module modules/mod_speling.so
#LoadModule status_module modules/mod_status.so
#LoadModule unique_id_module modules/mod_unique_id.so
LoadModule userdir_module modules/mod_userdir.so
#LoadModule usertrack_module modules/mod_usertrack.so
#LoadModule vhost_alias_module modules/mod_vhost_alias.so
#LoadModule ssl_module modules/mod_ssl.so

# 'Main' server configuration
#
# The directives in this section set up the values used by the 'main'
# server, which responds to any requests that aren't handled by a
# <VirtualHost> definition. These values also provide defaults for
# any <VirtualHost> containers you may define later in the file.
#
# All of these directives may appear inside <VirtualHost> containers,
# in which case these default settings will be overridden for the
# virtual host being defined.
#
#
# ServerAdmin: Your address, where problems with the server should be
# e-mailed. This address appears on some server-generated pages, such
# as error documents. e.g. admin@your-domain.com
#
ServerAdmin vanicecunha@gmail.com

#
# ServerName gives the name and port that the server uses to identify itself.
# This can often be determined automatically, but we recommend you specify
# it explicitly to prevent problems during startup.
#
# If your host doesn't have a registered DNS name, enter its IP address here.
#

```

ServerName localhost:80

#

DocumentRoot: The directory out of which you will serve your
documents. By default, all requests are taken from this directory, but
symbolic links and aliases may be used to point to other locations.

#

DocumentRoot "C:/Arquivos de programas/Apache Software
Foundation/Apache2.2/htdocs"

#

Each directory to which Apache has access can be configured with respect
to which services and features are allowed and/or disabled in that
directory (and its subdirectories).

#

First, we configure the "default" to be a very restrictive set of
features.