

UNIVERSIDADE DO EXTREMO SUL CATARINENSE – UNESC

CURSO CIÊNCIA DA COMPUTAÇÃO

MÁRLON HENRIQUE DE SOUZA

**AVALIAÇÃO DE DESEMPENHO DA ESCALABILIDADE ESPACIAL NA
RECOMENDAÇÃO H.264/SVC PARA CODIFICAÇÃO DE VÍDEO.**

CRICIÚMA, DEZEMBRO 2011

MÁRLON HENRIQUE DE SOUZA

**AVALIAÇÃO DE DESEMPENHO DA ESCALABILIDADE ESPACIAL NA
RECOMENDAÇÃO H.264/SVC PARA CODIFICAÇÃO DE VÍDEO.**

Trabalho de Conclusão de Curso apresentado para obtenção do Grau de bacharel em Ciência da Computação da Universidade do Extremo Sul Catarinense.


Orientador: Prof. M.Eng. Evânio Ramos Nicoleit

CRICIÚMA, DEZEMBRO 2011

MÁRLON HENRIQUE DE SOUZA

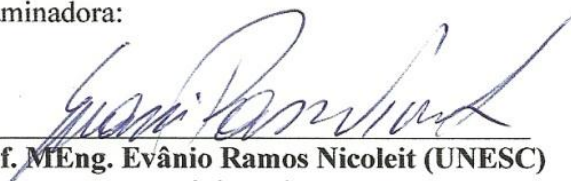
**Avaliação do Desempenho da Escalabilidade Espacial, na Recomendação
H.264/SVC para Codificação de Vídeo.**

Submetido ao corpo docente do Curso de Ciência da Computação da
Universidade do Extremo Sul Catarinense como um dos requisitos para obtenção do grau
de Bacharel em Ciência da Computação.



Profa. MSc. Ana Claudia Garcia Barbosa
Coordenadora do Curso de Ciência da Computação


Banca Examinadora:



Prof. MEng. Evânio Ramos Nicoleit (UNESC)
Orientador



Prof. MSc. Gustavo Bisognin (UNESC)



Prof. MSc. Paracelso de Oliveira Caldas (UNESC)

Dedico este trabalho primeiramente a Deus;
aos meus pais José Paulo e Marlei; a meu
irmão Maicon; e a minha namorada Kétlym;
pela compreensão e esforço para que eu
alcançasse este objetivo.

AGRADECIMENTOS

Agradeço a Deus, que me deu saúde para prosseguir nesta caminhada e enfrentar as dificuldades. Aos meus pais e familiares, pelo o apoio incondicional e compreensão nos momentos difíceis, a força de vocês me fez persistir neste sonho.

Ao meu orientador Prof. MEng. Evânio Ramos Nicoleit pelo apoio e disposição no decorrer desta pesquisa. Aos meus colegas de classe, por compartilharem informações e pelo convívio fraternal e familiar.

A Todos, o meu, **MUITO OBRIGADO!**

“O único lugar onde sucesso vem
antes do trabalho é no dicionário”
(Albert-Einstein)

RESUMO

Com o crescimento da utilização e manipulação de vídeos digitais em alta resolução, o uso de padrões de codificação que atendam as necessidades atuais é fundamental. O padrão H.264/AVC é o estado da arte atualmente com relação à codificação de vídeo, conseguindo aumentar em duas vezes a taxa de compressão em relação MPEG-2. A diversidade de dispositivos que permitem à visualização de vídeos digitais, aliada a evolução dos padrões de codificação resulta em um aumento na dificuldade de execução em dispositivos de baixa resolução, sem contar no desperdício de banda para transmissão a estes dispositivos. O padrão H.264/SVC incorpora escalabilidade ao padrão AVC, permitindo a manipulação da escala do vídeo seja ela espacial, temporal ou de qualidade. Consequentemente pode se extrair sequências com escalares inferiores ao vídeo original, para facilitar utilização do padrão H.264 em dispositivos com baixa resolução. Este projeto discute o desempenho do H.264/SVC, em especial a sintonia de quantização para melhor explorar as condições de escala espacial. A quantização de vídeos digitais varia de 0 a 51, as simulações realizadas visaram avaliar a influência da quantização neste intervalo, de acordo com métricas como taxa de bits, PSNR e complexidade computacional. Os resultados comprovaram que as condições de quantização da camada base de um vídeo com escalabilidade espacial, influenciam diretamente na taxa de bits e no tempo de codificação do vídeo resultante.

Palavras-chave: H.264. Escalabilidade. Quantização. Codificação de vídeo.

ABSTRACT

With the growing use and manipulation of digital video in high resolution, the use of coding standards that meet current needs is crucial. The H.264/AVC standard is the current state of the art with respect to video encoding, thus increasing at twice the rate of compression ratio MPEG-2. The diversity of devices that allow viewing of digital video, coupled with changing patterns of coding results in an increase in the difficulty of running on devices with low resolution, not to mention the waste of bandwidth for transmission to these devices. The H.264/SVC standard incorporates scalability to the H.264/AVC standard, allowing the manipulation of the scale of the video whether spatial, temporal or quality. Result can be extracted sequences with scalar below the original video, to facilitate use of the H.264 standard on devices with low resolution. This project discusses the performance of H.264/SVC, especially tuning quantization to better explore the conditions of spatial scale. The quantization of digital video ranges from 0 to 51, the simulations aimed to evaluate the influence of this quantization interval, according to metrics such as bit rate, PSNR and computational complexity. The results showed that the conditions of quantization of a base layer of video spatial scalability, directly influence the bit rate and encoding time of the resulting video.

Keywords: H.264. Scability. Quantization. Video Coding.

LISTA DE ILUSTRAÇÕES

Figura 1. Processo de codificação escalar.	24
Figura 2. Codificação escalar espacial.....	26
Figura 3. Progressão escalar espacial	26
Figura 4. Camada de enriquecimento na escalabilidade temporal.	27
Figura 5. Progressão escalar temporal	28
Figura 6. Progressão escalar SNR	29
Figura 7. Processo de subamostragem.....	30
Figura 8. Projeção para camada de enriquecimento.	31
Figura 9. Progressão escalar por meio da codificação piramidal.	32
Figura 10. Filtros de sub-bandas.....	32
Figura 11. Evolução dos padrões de codificação de vídeo.....	36
Figura 12. Codificador H.264/AVC	38
Figura 13. Estrutura do codificador SVC	40
Figura 14. Comportamento da escalabilidade espacial no H.264/SVC.....	41
Figura 15. Comportamento da escalabilidade temporal no H.264/SVC.	42
Figura 16. Comando básico do codificador JSVM.....	48
Figura 17 Saída de dados em tempo de codificação.....	51
Figura 18 Resultados apresentados ao final da codificação.	52
Figura 19. Comando básico do extrator JSVM.	52
Figura 20. Comando básico do decodificador JSVM.....	53
Figura 21. Informações das sequências utilizadas.....	55
Figura 22. Taxa de bits em função da quantização.....	59
Figura 23. Interseção das curvas de taxa de bits.....	60

Figura 24. PSNR em função da quantização.	61
Figura 25. Tempo em função da quantização.	62
Figura 26. PSNR em função da taxa de bits.	63
Figura 27. Interseção das curvas de PSNR.	64
Figura 28. Tempo de codificação em função da PSNR.	65
Figura 29. Tempo de codificação em função da taxa de bits.	66
Figura 30. Comparativo de frames da camada base.	67
Figura 31. Camada de Incremento - Coastguard.	68
Figura 32. Camada de Incremento - Foreman.	69
Figura 33. Camada de Incremento - Walk.	69

LISTA DE TABELAS

Tabela 1. Bloco 8x8 de resíduos.....	22
Tabela 2. Tabela de quantização JPEG.....	22
Tabela 3. Bloco Transformado	22
Tabela 4. Bloco transformado e quantizado.	22
Tabela 5. Arquivo de ConFiguração "Main".....	49
Tabela 6. Arquivo de conFiguração da primeira camada "layer0".....	50
Tabela 7. Arquivo de conFiguração da segunda camada "layer1".....	50
Tabela 8. Tabela saída de dados JSVM.....	51

LISTA DE SIGLAS E ABREVIATURAS

ATSC	<i>Advanced Television Systems Committee</i>
AVC	<i>Advanced Video Coding</i>
CGS	<i>Coars-Grain Quality Scalability</i>
CODEC	<i>Coder-Decoder</i>
DCT	<i>Discrete Cosine Transform</i>
DVB	<i>Digital Video Broadcasting</i>
DVD	<i>Digital Versatile Disk</i>
HDTV	<i>High Definition Digital Television</i>
IEC	<i>International Electrotechnical Commission</i>
ISDB	<i>Integrated Services Digital Broadcasting</i>
ISO	<i>International Organization for Standardization</i>
ITU-T	<i>International Telecommunication Union - Telephony</i>
JPEG	<i>Joint Photography Experts Group</i>
JSVM	<i>Joint Scalable Video Model</i>
JVT	<i>Joint Video Team</i>
MB	<i>Macro Bloco</i>
MPEG	<i>Moving Pictures Experts Group</i>
MSE	<i>Mean Squared Error</i>
PSNR	<i>Peak Signal to Noise Ratio</i>
SBTVD	<i>Sistema Brasileiro de TV Digital</i>
SDTV	<i>Standard Definition TV</i>
SNR	<i>Signal-Noise Ratio</i>
SVC	<i>Scalable Video Coding</i>

TV	<i>Televisão</i>
VCEG	<i>Video Coding Experts Group</i>
VCL	<i>Video Coding Layer</i>

SUMÁRIO

1 INTRODUÇÃO	9
1.1 OBJETIVO GERAL.....	11
1.2 OBJETIVOS ESPECÍFICOS	11
1.3 JUSTIFICATIVA	12
1.4 ESTRUTURA DO PROJETO.....	13
2 VÍDEOS DIGITAIS	15
2.1 EVOLUÇÃO DOS VÍDEOS DIGITAIS	15
2.2 CODIFICAÇÃO E DIGITALIZAÇÃO DE VÍDEOS DIGITAIS	16
2.3 TIPOS DE REDUNDÂNCIA	17
2.3.1 Redundância Espacial	17
2.3.2 Redundância Temporal.....	18
2.3.3 Redundância Psico-Visual	18
2.4 TÉCNICAS DE COMPRESSÃO DE VÍDEO	19
2.4.1 Predição Espacial.....	20
2.4.2 Transformada	20
2.4.3 Quantização.....	21
2.4.4 Codificação por Entropia.....	23
3 CODIFICAÇÃO ESCALÁVEL.....	24
3.1 CLASSIFICAÇÃO DAS TÉCNICAS DE ESCALABILIDADE	25
3.1.1 Escalabilidade Espacial.....	25

3.1.2 Escalabilidade Temporal	27
3.1.3 Escalabilidade Qualitativa (SNR)	28
3.2 TÉCNICAS DE CODIFICAÇÃO ESCALÁVEL.....	29
3.2.1 Codificação Piramidal.....	30
3.2.2 Codificação em Sub-bandas.....	32
3.2.3 Transformada DCT.....	33
4 PADRÕES DE CODIFICAÇÃO	35
4.1 EVOLUÇÃO DOS PADRÕES DE CODIFICAÇÃO	35
4.2 PADRÃO H.264.....	38
4.3 EXTENSÃO ESCALÁVEL DO PADRÃO H.264.....	39
4.3.1 Escalabilidade Espacial.....	41
4.3.2 Escalabilidade Temporal	42
4.3.3 Escalabilidade por Qualidade (SNR).....	43
5 TRABALHOS CORRELATOS	44
5.1 OVERVIEW OF THE SVC EXTENSION H.264/AVC	44
5.2 CODIFICAÇÃO ESCALÁVEL DE VIDEO PARA RECEPÇÃO FIXA DE SBTV	45
5.3 AVALIAÇÃO SUBJETIVA DE QUALIDADE APLICADA A CODIFICAÇÃO DE VIDEO ESCALÁVEL	45
6 AMBIENTE DE SIMULAÇÃO	46
6.1 H.264/SVC JSVM REFERENCE 9.19	46
6.2 CODIFICADOR “H264AVCENCODERLIBTESTSTATIC”	47
6.2.1 ConFiguração do Arquivo “main”.....	48

6.2.2 ConFiguração das Camadas (Layers)	49
6.2.3 Formato de Saída	50
6.3 APLICATIVO DE AMOSTRAGEM “BITSTREAMEXTRACTORSTATIC”	52
6.4 DECODIFICADOR “H264AVCDECODERLIBTESTSTATIC”	53
6.5 PROCEDIMENTO EXPERIMENTAL	54
6.5.1 Sequências de Testes	54
6.5.2 Métricas de Comparação	55
6.6 ANÁLISE DAS MÉTRICAS EM FUNÇÃO DA QUANTIZAÇÃO	56
6.6.1 Dados Utilizados na Análise	57
7 RESULTADOS E ANÁLISE	58
7.1 TAXA DE BITS EM FUNÇÃO DA QUANTIZAÇÃO.....	58
7.2 PSNR EM FUNÇÃO DA QUANTIZAÇÃO	60
7.3 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA QUANTIZAÇÃO.....	62
7.4 PSNR EM FUNÇÃO DA TAXA DE BITS	63
7.5 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DE PSNR	65
7.6 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA TAXA DE BITS.....	66
7.7 ANÁLISE DAS SEQUÊNCIAS RESULTANTES	67
7.7.1 Camada Base dos Vídeos Simulados	67
7.7.2 Camada de Incremento dos Vídeos Simulados	68
CONCLUSÃO	70
REFERÊNCIAS	72

1 INTRODUÇÃO

À medida que os tempos avançam, a sociedade sente necessidade de sistemas de comunicação cada vez mais rápidos, eficientes e de baixo custo. Atualmente o Brasil passa por uma fase de transição histórica: do sistema de TV aberta analógica para o sistema de TV aberta digital.

A chave para este recente sistema de TV são as técnicas de compressão. Nos últimos anos, essas técnicas passaram por uma grande evolução. Frequentemente surgem novas técnicas e padrões que melhoram a qualidade do vídeo digital para uma dada taxa de bits, ou reduzem a taxa de *bits* mantendo a qualidade. Paralelamente a esta evolução, avanços tecnológicos importantes permitiram a implementação em tempo real e a um custo razoável de eficientes algoritmos de compressão. Ainda assim, deve ser sempre considerada a complexidade dos algoritmos envolvidos, que é computacionalmente tão importante quanto a qualidade visual obtida.

Dentre as aplicações de vídeo digital, destacam-se, portanto: armazenamento e distribuição de filmes, difusão de TV digital aberta e por assinatura e videoconferência, videofonia e webconferência.

Nos últimos anos, tem se verificado uma grande evolução nos padrões internacionais para compressão e codificação de vídeo. Uma consequência desta padronização é o crescente interesse em técnicas de compressão mais complexas e eficientes, que combinem altas taxas de compressão com alta qualidade visual. Diversos esquemas de codificação digital de vídeo têm sido propostos procurando atender às mais diferentes aplicações, classificadas de acordo com a relação entre taxa de bits e qualidade requerida.

O padrão H.264/AVC - Advanced Video Coding) é atualmente o estado-da-arte em codificação de vídeo. Ele alcança taxas de compressão superiores aos padrões anteriores, tais como o H.263 (ITU-T, 2005).

O padrão H.264 possibilita que vídeos digitais sejam manipulados por diferentes tipos de dispositivos, tais como dispositivos móveis, aparelhos celulares, com baixas resoluções, e aparelhos de TV de alta definição. É requerida, portanto a transmissão de vídeo em diferentes resoluções de vídeo (escalabilidade espacial e escalabilidade temporal) compatíveis com estes tipos de dispositivos.

Escalabilidade, de forma geral, refere-se à capacidade de extração de partes do *bitstream* de um vídeo codificado para adaptá-lo às diferentes necessidades e/ou preferências do usuário, assim também como às diferentes capacidades de receptores e/ou a condições das redes de comunicação (NUNES, 2009).

O padrão H.264/AVC possui uma extensão desenvolvida que visa o suporte a escalabilidade denominado H.264/SVC – Scalable Video Coding (WIEGAND ET AL., 2007, tradução nossa). Esta extensão possibilita a adição ou retirada de camadas dentro do fluxo de bits, manipularem-se a escala de acordo com a aplicação. O H.264/SVC provê suporte a múltiplas resoluções de exibição com um único *bitstream*. Adicionalmente, o H.264/SVC também suporta escalabilidade temporal, que está relacionada com a taxa de quadros, ou seja, alguns quadros podem ser descartados; e a escalabilidade de qualidade, que é o enriquecimento na fidelidade dos quadros mantendo-se a resolução entre as camadas. O H.264/SVC também permite a combinação dessas técnicas (SEGALL; SULLIVAN, 2007, tradução nossa).

Diante desta variedade de cenários para atender às necessidades de diferentes usuários, equipamentos e redes, esta pesquisa busca avaliar o desempenho em termos de taxa

de *bits*, qualidade e complexidade computacional para diferentes escalas do padrão H.264/SVC.

A qualidade do vídeo codificado/decodificado pelo codec do H.264/SVC será comparada com o vídeo original nas diferentes escalas em diferentes taxas de *bits*, segundo medidas (métricas) objetivas, tais como Razão Sinal Ruído de Pico (PSNR) e tempo de codificação sobre sequências de vídeo de referência.

1.1 OBJETIVO GERAL

Avaliar o desempenho da escalabilidade na Recomendação H.264/SVC para codificação de vídeo em diferentes taxas de *bits*.

1.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos do trabalho proposto seguem abaixo:

- a) compreender e descrever os fundamentos de codificação de vídeo;
- b) compreender e descrever escalabilidade espacial, temporal e de qualidade;
- c) compreender e descrever o processo de codificação e decodificação de vídeo baseado na Recomendação H.264/SVC;
- d) descrever a implementação de referência JSVM do sistema de codificação e decodificação de vídeo baseado na recomendação H.264/SVC para a realização das avaliações;

- e) comparar a qualidade do vídeo reconstruído com o vídeo original com respeito às métricas PSNR, MSE e Complexidade Computacional nas diferentes escalas espaciais;

1.3 JUSTIFICATIVA

Devido à evolução das aplicações com vídeo, a compreensão e o aprimoramento das técnicas de compressão se tornam fundamentais. A busca por formas de elevar a qualidade dos vídeos, reduzindo a sua taxa de *bits* e a sua complexidade computacional tem tornado os codificadores cada vez mais eficientes.

O padrão H.264 é atualmente o estado da arte em se tratando de codificação de vídeo, por ter alcançado uma grande taxa de compressão, muito superior aos padrões já existentes (RICHARDSON, 2003). O sistema de codificação de vídeo depende basicamente de três características fundamentais: da qualidade requerida das imagens do vídeo resultante, da taxa de *bits* associada e da complexidade computacional do codificador.

Uma alternativa eficiente para adequar a disponibilidade de banda de transmissão às disponibilidades das redes dos usuários é o uso da escalabilidade, que se baseia no conceito de camadas de enriquecimento, onde a primeira camada, chamada de camada base, possui menos informação do que as camadas superiores, chamadas de camadas de aprimoramento/enriquecimento (*enhancement layers*).

As camadas de enriquecimento são chamadas assim, pois cada nova camada vai adicionando novas informações para prover mais detalhes, como por exemplo, informações para uma resolução mais elevada. Como existe apenas um sinal sendo transmitido, de forma escalável, o dispositivo/rede do usuário pode selecionar até qual camada de resolução lhe é compatível, processando apenas as informações associadas.

Esta pesquisa visa com uso da escalabilidade, avaliar a qualidade do vídeo na decodificação de vídeos no padrão H.264/SVC, em diferentes resoluções comparativamente a métricas objetivas tais como PSNR e MSE e Complexidade Computacional.

1.4 ESTRUTURA DO PROJETO

O Capítulo 1 deste projeto traz a definição do problema bem como o objetivo geral, objetivos específicos e a sua justificativa.

O Capítulo 2 traz conceitos sobre vídeos digitais a sua evolução e disseminação na sociedade atual. Além disso, trata o objetivo da compressão e digitalização dos vídeos, as redundâncias reduzidas com isso e as métricas ou características que permitem a comparação do desempenho dos padrões de codificação de vídeo.

O Capítulo 3 entra no contexto da codificação escalável, classificando as escalas de vídeo e buscando as principais técnicas utilizadas para codificação escalável nos últimos anos.

O Capítulo 4 aborda um histórico dos padrões de codificação dos órgãos regulamentadores ITU-T e MPEG. Posteriormente é conceituado o padrão de H.264 de codificação de vídeo, bem como sua extensão SVC principal objeto de estudo desta pesquisa. Esta extensão é conceituada ao longo do capítulo por meio das três classificações escalares de vídeo: espacial, temporal e de qualidade. Ao final do capítulo são citados e analisados os principais simuladores para o padrão H.264.

O Capítulo 5 traz uma abordagem sobre os trabalhos já realizados sobre escalabilidade com o padrão de codificação de vídeo H.264/SVC.

O Capítulo 6 inicia uma descrição sobre o ambiente de simulação utilizado JSVM e seus aplicativos principais, além das sequências utilizadas nos testes. Este capítulo também

apresenta as métricas utilizadas na pesquisa e o método utilizado para se verificar o desempenho do padrão H.264/SVC, em função de seu parâmetro de quantização.

O Capítulo 7 apresenta todos os resultados e uma análise sobre os mesmos, correlacionando taxa de bits, PSNR-Y, tempo de codificação com o parâmetro de quantização.

2 VÍDEOS DIGITAIS

Os vídeos digitais são constituídos por uma cadeia de quadros sequenciados, cada quadro compreende a uma imagem com um determinado número de *pixels* (SILVA, 2006).

Um vídeo digital é classificado com uma mídia contínua, ou seja, possui modificação de acordo com o tempo. Contudo um vídeo provém de uma sucessão de imagens consideradas discretas, informações estáticas. Esta relação das imagens nos vídeos vem sendo tratada nos últimos anos, nos algoritmos de codificação de vídeo (BRUNO, 2003).

Dentre as aplicações de vídeo digital, destacam-se, portanto: armazenamento e distribuição de filmes, difusão de TV digital aberta e por assinatura e videoconferência, videofonia e webconferência. Estas aplicações vêm utilizando um sinal progressivo, o mais utilizado atualmente.

2.1 EVOLUÇÃO DOS VÍDEOS DIGITAIS

No mundo atual, muitas mídias fazem parte do cotidiano das pessoas, a digitalização das imagens, vídeos e sons tornou mais fácil o acesso a estes tipos de dados. Estamos em período da história aonde a utilização das mídias digitais vem sendo aplicadas fortemente na disseminação de informação.

Ouve-se muito falar atualmente no advento da TV digital, o sinal de TV digital já esta sendo disponibilizado em muitas regiões do Brasil. Vamos verificar a seguir que sem a evolução dos vídeos digitais não seria possível o advento da TV digital (NUNES, 2009).

Segundo Teraoka (2003, p. 14), o sistema utilizado em TV digital seria inviável há poucos anos atrás sem os avanços das técnicas de codificação de imagens. Sem estas técnicas, um sinal de vídeo digital na mesma resolução utilizada pelo sistema brasileiro de

televisão necessitaria de aproximadamente 270 Mbits/s (270 Mbps) no padrão ITU-R-BT-601, o que é impossível de se transmitir em um único canal de seis MHz utilizado para a transmissão destes sinais.

Estamos em um momento da história em que a tecnologia evolui exponencialmente, os vídeos vem crescendo em nível de aplicação e qualidade, diminuindo o tamanho de armazenamento.

2.2 CODIFICAÇÃO E DIGITALIZAÇÃO DE VÍDEOS DIGITAIS

A busca pela compactação de um vídeo digital consiste na eliminação de dados redundantes, visando otimizar o espaço de armazenamento e a transmissão do vídeo nas diferentes redes existentes. A compressão de vídeos é fundamental para que se torne possível à utilização em grandes escala, pois a não codificação de um vídeo digital traria taxas de *bits* elevadíssimas dificultando manipulação dos mesmos (ASCENSO, 2003).

Como já descrito, um vídeo é sequência de imagens ao longo do tempo. Contudo, se cada *pixel* de cada imagem for digitalizado um arquivo imenso seria gerado. Por exemplo, uma tela de uma televisão comum de 640x480 *pixels* em um total de 307200. Em um vídeo apenas em preto e branco, com 256 níveis de cinza e 10 segundos de duração, a uma taxa de 30 quadros (*frames*) por segundo (fps), resultaria em um arquivo de vídeo com 737 Mbits, ou seja, mais de 97 Mbytes. As técnicas de compressão buscam justamente eliminar esta condição, trabalhando principalmente as redundâncias geradas entre os quadros na estrutura de um vídeo digital.

A codificação de vídeos digitais busca manter a maior fidelidade possível com relação ao vídeo original para uma taxa de bits determinada, ou busca-se a menor taxa de *bits* se mantendo a qualidade do vídeo original (SULLIVAN; WIEGAND, 2005, tradução nossa).

A representação digital nada mais é do que a representação de informação por meio de símbolos pré-definidos. Um exemplo muito conhecido é o código Morse utilizado para comunicação à distância, criado em 1835 por Jovan Costa (WIKIPEDIA, 2011).

A codificação de sistemas de multimídia é tratada pelos codificadores e decodificadores denominados *codec's*. Os *codec's* são algoritmos desenvolvidos para compressão de áudio e vídeo digital, alcançando reduções taxas de redução sem perdas entre 2:1 e 5:1 (NUNES, 2009).

2.3 TIPOS DE REDUNDÂNCIA

Como já descrito a codificação de um vídeo digital e sua consequente compressão visa à eliminação das redundâncias em sinal. Se tais redundâncias fossem mantidas, não seria possível em tal nível de usabilidade para vídeos digitais como temos atualmente (DARONCO, 2009).

Levando em consideração os objetivos desta pesquisa, serão explicados os seguintes tipos de redundâncias: redundância espacial, redundância temporal e redundância psico-visual.

2.3.1 Redundância Espacial

A redundância espacial é também chamada de redundância intraquadro ou redundância *interpixel* e advém da correlação existente entre os pixels espacialmente distribuídos em um quadro (AGOSTINI, 2007).

Na redundância espacial são observados pixels vizinhos em um quadro que tendem a possuir valores semelhantes. Este tipo de redundância pode ser reduzida em uma

operação chamada de codificação intraquadro, presente em alguns padrões de codificação de vídeo atuais (RICHARDSON, 2003).

2.3.2 Redundância Temporal

A redundância temporal, também chamada de redundância interquadros (GHANBARI, 2003), é causada pela não modificação de *pixels* existente em quadros temporalmente próximos no decorrer de um vídeo digital. Na verdade, a redundância temporal poderia ser classificada como apenas mais uma dimensão da redundância espacial, como em (GONZALEZ, 2003). Blocos de *pixels* um vídeo digital não possuem variação temporalmente, pois todo vídeo digital possui cenas com quadros muito parecidos, temos inúmeros exemplos de vídeos que um determinado fundo permanece sem alteração ou alterações mínimas de luminosidade.

Além disso, existem exemplos de vídeos onde um determinado bloco de *pixels* se desloque de um frame para o outro, um exemplo é a movimentação de um objeto qualquer em uma cena. Os padrões de codificação atuais alcançam bons resultados focando na redundância temporal, no padrão de codificação adotado na pesquisa H.264/AVC se caracteriza por grandes ganhos de compressão por trabalhar com macro blocos (blocos de pixels) reduzindo consideravelmente a redundância temporal.

2.3.3 Redundância Psico-Visual

A redundância psico-visual vale-se da sensibilidade do sistema visual humano para descartar algumas informações que podem ser removidas sem afetar significativamente a percepção da qualidade visual (RICHARDSON, 2003). Por exemplo, o olho humano é muito

mais sensível às mudanças da luminância do que crominância, assim um sistema pode eliminar algumas informações de cores sem que as pessoas percebam ou, caso isso ocorra, notem pouca diferença. Para explorar este tipo de redundância, parte da informação original da imagem é eliminada de forma irreversível pelo codificador.

Existem dois tipos principais de processos utilizados para este fim. O primeiro, chamado de subamostragem, é utilizado na entrada do vídeo e elimina parte das informações de cores. O segundo, conhecido por quantização, normalmente é aplicado no domínio das frequências e elimina ou atenua as frequências de menor importância para o sistema visual humano (SOLER, 2006).

É importante destacar que a eliminação das redundâncias psico-visuais contribui para que o codificador atinja elevadas taxas de compressão, com pequeno impacto na qualidade visual da imagem.

2.4 TÉCNICAS DE COMPRESSÃO DE VÍDEO

Existem dois tipos de técnicas de compressão de vídeo, com perdas e sem perdas. As técnicas de compressão sem perdas visam reproduzir o vídeo exatamente igual o vídeo original, muito utilizado onde a riqueza de detalhes é extremamente necessária, como por exemplo, em arquivos médicos onde todo detalhe influencia no diagnóstico. Nas técnicas de codificação com perdas, o vídeo codificado perde algumas características em relação ao vídeo original, um exemplo disso é o caso de vídeos utilizados na web onde não há a necessidade de um vídeo totalmente fiel ao original para que a informação seja repassada (SILVA, 2006).

Dentro desses dois tipos de compressão existem técnicas utilizadas para atingir o objetivo de cada um, a seguir algumas técnicas serão apresentadas que contenham correlação a resolução espacial de vídeos digitais, foco desta pesquisa.

2.4.1 Predição Espacial

A predição espacial também conhecida como predição intra usa como referencia amostras do quadro atual que estão decodificadas para predizer as amostras que estão sendo codificadas (FERREIRA R., 2009).

A predição intra-quadro é realizada através da interpolação de valores dos pixels das bordas de blocos espacialmente vizinhos que foram anteriormente codificados. Esta interpolação é realizada direcionalmente, em múltiplos modos, cada um deles empregado em uma direção espacial diferente (PURI; CHEN; LUTHRA, 2004).

2.4.2 Transformada

A transformada consiste em tratar os valores das amostras de um bloco, reduzindo as redundâncias espaciais. Além disso, este processo transforma o sinal para o domínio da frequência espacial e direcionando o sinal, concentrando a energia em alguns coeficientes (MANOEL, 2007).

O processo de transformação consiste em converter os valores dos resíduos que estão no domínio espacial, para o outro domínio de frequência. Devido ao fato que as amostras no domínio espacial variam muito ao longo do tempo, elas não são comprimidas muito bem pelas técnicas de entropia. Já quando as amostras são convertidas para o domínio de frequência, elas se tornam muito mais adequadas para compressão por entropia (HUNG, 2007).

2.4.3 Quantização

O processo de quantização mapeia um sinal com intervalo de valores X para um sinal quantizado com um intervalo reduzido de valores X_q . Permitindo representar um sinal quantizado com taxa de *bits* inferiores ao sinal original (FERREIRA R., 2009).

Por meio do processo de quantização acontecem as maiores perdas de qualidade e também as maiores taxas de compressão. A redução causada é decorrente do processo de aproximação onde um ou mais valores são substituídos por índices de uma escala finita (BRUNO, 2003).

A quantização tem por objetivo reduzir a amplitude do sinal dividindo-o e posteriormente o arredondando. Com isso pode se transmitir a informação com menos *bits*, e quando recebidos aplica-se o processo inverso multiplicando os coeficientes para reconstruir o sinal original (FERREIRA R., 2009).

$$X_q = \text{round} \left(\frac{x}{Q_p} \right) \quad (1)$$

$$\hat{X} = X_q \cdot Q_p \quad (2)$$

As Equações 1 e 2 se referem ao processo de quantização linear onde X_q é o valor quantizado, \hat{X} é o valor recuperado. O valor recuperado é normalmente considerado como valor desquantizado, contudo o processo de quantização não permite a recuperação exata do valor original (FERREIRA R., 2009).

As Tabelas 1,2 demonstram os valores de um bloco de tamanho 8x8 pixels do resíduo de uma sequência de frames e a tabela de quantização do padrão JPEG. As Tabelas 3 e 4 demonstram o processo de quantização sendo que a Tabela 3 traz o valores após a transformada e a Tabela 4 os valores já quantizados, onde pode se perceber a redução da amplitude dos valores.

Tabela 1. Bloco 8x8 de resíduos.

-2	-2	-1	6	18	12	-1	-2
-1	0	-1	-1	1	6	5	-1
0	-1	0	2	-7	-6	1	-3
-1	1	-2	-3	-1	2	3	2
-1	1	2	-1	2	1	-2	-1
0	2	1	0	0	0	1	-1
-1	1	2	1	2	2	1	1
1	1	1	1	-1	-3	-2	-2

Fonte: (FERREIRA R., 2009).

Tabela 2. Tabela de quantização JPEG.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Fonte: (FERREIRA R., 2009).

Tabela 3. Bloco Transformado

57,8750	34,0422	-30,8251	-1,8822	3,6250	-8,4951	-5,4972	4,2405
-38,2364	1,1645	-41,5549	22,1580	-27,7356	8,9246	-7,5870	-1,4612
-9,2204	28,3195	-41,8711	33,5663	-34,4742	6,5632	-0,8436	-5,8853
-1,8892	41,9590	-7,5757	25,4611	-1,9353	-2,6480	0,7227	-3,4208
-31,8750	52,5784	1,1200	-11,4830	24,6250	-7,0802	2,3773	2,3093
-4,9106	6,4469	2,1081	-9,7191	4,1768	4,6382	2,3321	-1,9163
-0,3751	-9,9092	-5,3436	7,4830	-7,7741	-1,0881	4,6211	-2,4698
11,2908	-8,5322	-8,7200	6,9507	-5,5660	-4,8681	7,8368	-0,7637

Fonte: (FERREIRA R., 2009).

Tabela 4. Bloco transformado e quantizado.

-4	3	-3	0	0	0	0	0
-3	0	-3	1	-1	0	0	0
-1	2	-3	1	-1	0	0	0
0	2	0	1	0	0	0	0
-2	2	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Fonte: (FERREIRA R., 2009).

2.4.4 Codificação por Entropia

A codificação por entropia converte uma série de símbolos, que representam os elementos da sequência dos vídeos, em um fluxo de *bits* compactados apropriados para transmissão ou armazenamento (AGOSTINI, 2007).

Segundo Manoel (2007, p. 31), a codificação por entropia atribui sequências binárias de tamanhos diferenciados para determinados símbolos (parâmetros, coeficientes, etc.) com probabilidades diferentes de aparecerem, com intuito de diminuir o total de *bits* codificados.

3 CODIFICAÇÃO ESCALÁVEL

Escalabilidade, de forma geral, refere-se à capacidade de extração de partes do *bitstream* de um vídeo codificado para adaptá-lo às diferentes necessidades e/ou preferências do usuário, assim também como às diferentes capacidades de receptores e/ou a condições das redes de comunicação (NUNES, 2009).

Se um vídeo é codificado em alta qualidade, altas taxas de *bits* são requeridas. Porém, usuários com redes de baixa velocidade perderão o vídeo pela reduzida velocidade. Da mesma forma, se o vídeo é codificado em baixa qualidade, e baixa taxa de *bits*, usuários com redes e equipamentos rápidos não usufruirão a sua infraestrutura. Ainda, se o vídeo é codificado em qualidade intermediária, tanto os usuários com redes de baixa quanto os usuários com redes de alta velocidade não serão contemplados. A solução está em escolher a escala e trechos do fluxo de bits adequados para cada tipo de usuário ou aplicação.

Codificação de vídeo escalável permite a um decodificador decodificar seletivamente apenas uma parte do fluxo de *bits*, como se pode observar na Figura 1 um decodificador pode selecionar as camadas desde a base até a última camada que confere a maior qualidade ao vídeo (RICHARDSON, 2003).

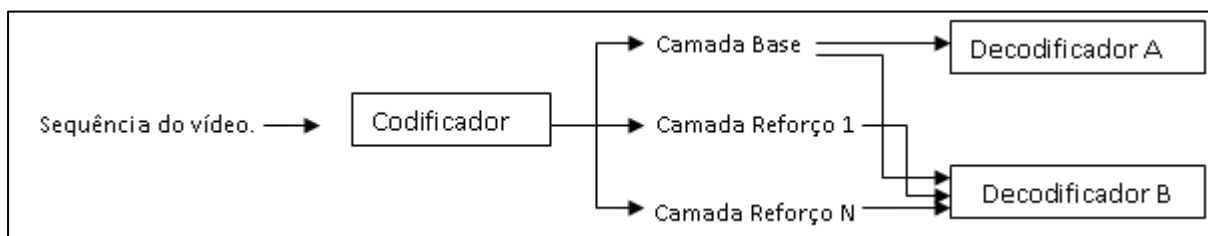


Figura 1. Processo de codificação escalável.
Fonte: RICHARDSON, I. (2003, p.142)

Uma alternativa eficiente para adequar a disponibilidade de banda de transmissão às disponibilidades das redes aos usuários é o uso da escalabilidade, que se baseia no conceito de camadas de enriquecimento, onde a primeira camada, chamada de camada base, possui

menos informação do que as camadas superiores, chamadas de camadas de aprimoramento/enriquecimento (*enhancement layers*).

As camadas de enriquecimento são chamadas assim, pois cada nova camada vai adicionando novas informações para prover mais detalhes, como por exemplo, informações para uma resolução mais elevada. Como existe apenas um sinal sendo transmitido, de forma escalável, o dispositivo do usuário pode selecionar até qual camada de resolução lhe é compatível, processando apenas as informações associadas (NUNES, 2009).

3.1 CLASSIFICAÇÃO DAS TÉCNICAS DE ESCALABILIDADE

Existem três tipos de escalabilidade: escalabilidade espacial, escalabilidade temporal e escalabilidade de qualidade (SNR).

3.1.1 Escalabilidade Espacial

Escalabilidade espacial é dada pela alteração da resolução dos quadros na constituição das camadas do vídeo de acordo com a necessidade. Esta variação de resolução se torna uma característica considerável em transmissões para dispositivos com várias resoluções diferentes (DARONCO, 2009).

A camada base possui uma versão reduzida da resolução completa dos quadros do vídeo codificado. Quando se é decodificado todas as camadas (base + reforços), temos na saída um vídeo em sua resolução completa.

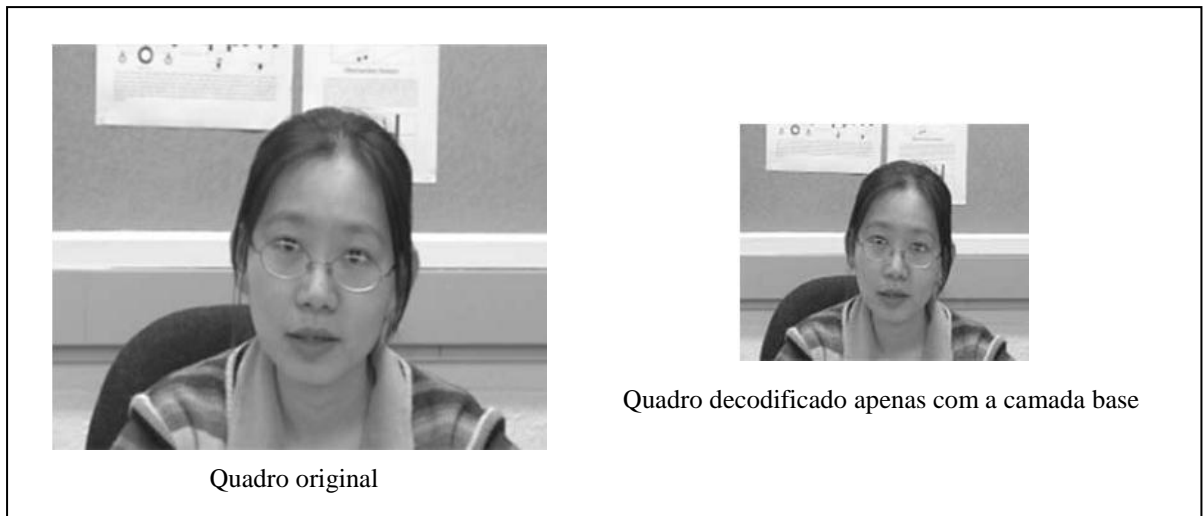


Figura 2. Codificação escalar espacial.
 Fonte: RICHARDSON, I. E. R. (2003 , p.143)

À medida que a decodificação alcança um número maior de camadas, a resolução espacial do vídeo aumenta, como se pode observar na Figura 3.

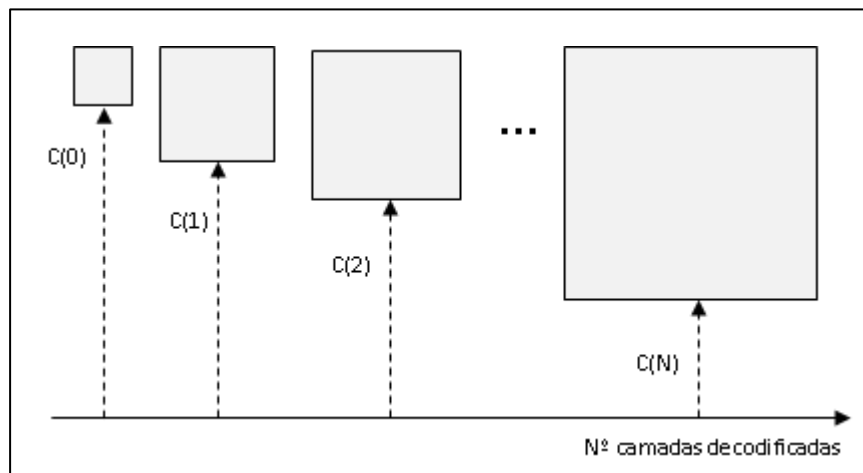


Figura 3. Progressão escalar espacial
 Fonte: ASCENSO, J (2003 , p. 24)

A transmissão de dados em camadas utilizando as técnicas de codificação escalar espacial se adapta tanto a banda disponível, quanto a dispositivos com resoluções diferentes (BRUNO, 2003).

3.1.2 Escalabilidade Temporal

Segundo Fonseca (2008, p.29), a escalabilidade temporal permite com um único *bitstream* extraírem-se múltiplas taxas de quadros, e sua implementação geralmente é pré-determinada pelo padrão de codificação.

Na escalabilidade temporal, existem camadas de mesma resolução uma denominada camada base formada por uma baixa taxa de quadros. As próximas camadas de enriquecimento intercalam com a camada base resultando uma taxa maior de quadros na sequência final (NUNES, 2009).

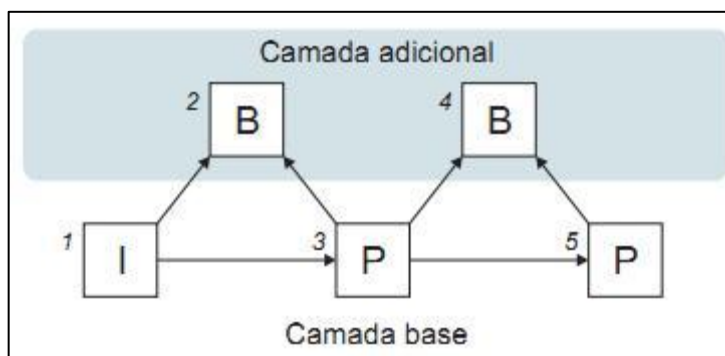


Figura 4. Camada de enriquecimento na escalabilidade temporal.
Fonte: DARONCO, L. C. (2009 ,p. 24)

A camada base demonstrada na Figura 4 é codificada com uma taxa fps (*frames* por segundo) menor e as demais camadas completam a base. A junção de todas as camadas resulta no número de quadros por segundo do vídeo original (DARONCO, 2009).

A seguir é demonstrado, na Figura 5, o comportamento da codificação escalar temporal:

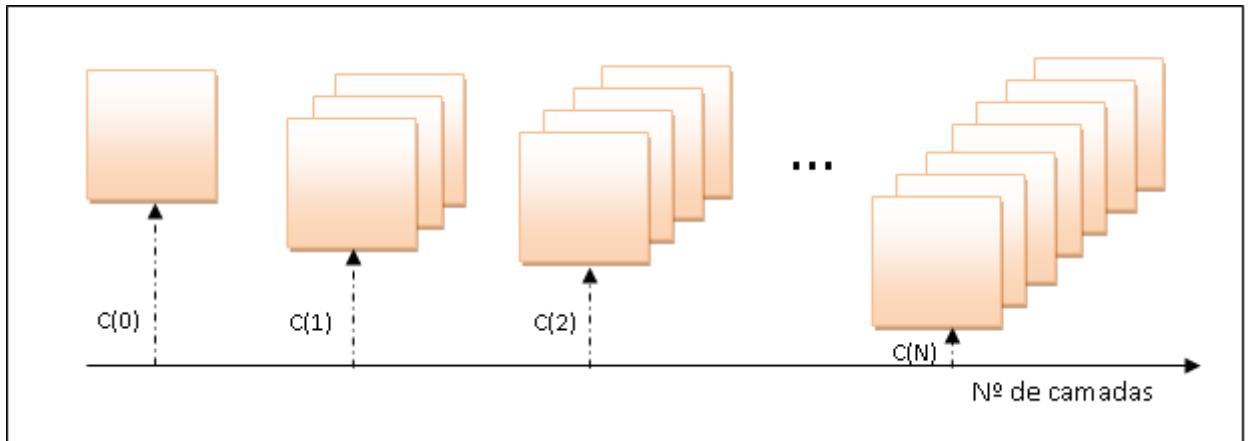


Figura 5. Progressão escalar temporal
 Fonte: ASCENSO, J. (2003, p. 24)

Dois categorias importantes utilizadas no emprego da variação escalar temporal são a codificação em sub-bandas e a baseada nas técnicas de estimativa de movimento.

3.1.3 Escalabilidade Qualitativa (SNR)

Como as demais classificações de escalabilidade, a qualitativa esta condicionada à variação de um fator influência em vídeos digitais, o processo de quantização. Interferindo diretamente no nível de fidelidade dos quadros (DARONCO, 2009).

A escalabilidade qualitativa torna possível a escolha do nível da taxa de *bits*, tendo a camada base como o menor fluxo de *bits* permitido pelo *bitstream* escalável. A taxa de *bits* geralmente é muito próxima ao numero de camadas (FONSECA, 2008).

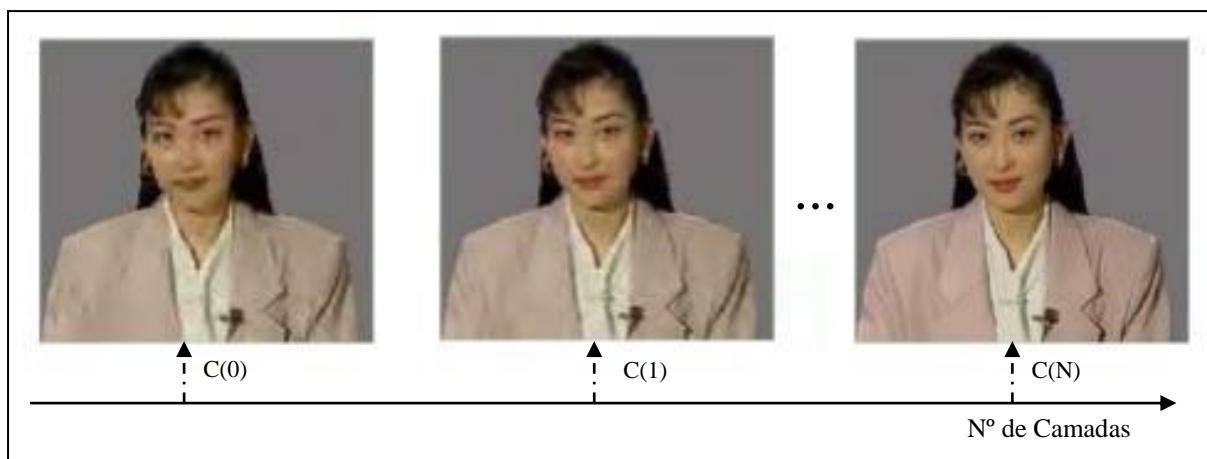


Figura 6. Progressão escalar SNR
Fonte: SOLER, L (2006, p. 29).

3.2 TÉCNICAS DE CODIFICAÇÃO ESCALÁVEL

Existem duas linhas de pesquisa para a compensação da qualidade em um vídeo digital na decodificação:

- a) **Protocolo de transporte:** direcionado a protocolos e arquiteturas inovadoras trabalhando o transporte do vídeo;
- b) **Codificação de fonte:** visa a adaptação do conteúdo as redes e aos terminais.

Neste subitem busca-se explicar sobre as técnicas mais relevantes ao conteúdo desta pesquisa, ou seja, as técnicas mais conhecidas que contenham ligação ao conceito de escalabilidade escalar, temporal e de SNR.

3.2.1 Codificação Piramidal

A denominação piramidal se trata de uma técnica escalar onde a cada camada acrescida à resolução espacial aumenta. A imagem no topo da pirâmide possui a resolução espacial considerada mínima (ASCENSO, 2003).

A Figura 7 demonstra a estrutura escalar pelo método de codificação piramidal, onde é feita uma subamostragem da imagem de entrada e formada a camada base, a subamostragem pode ser quantificada gerando uma resolução espacial com uma taxa inferior de *bits*.

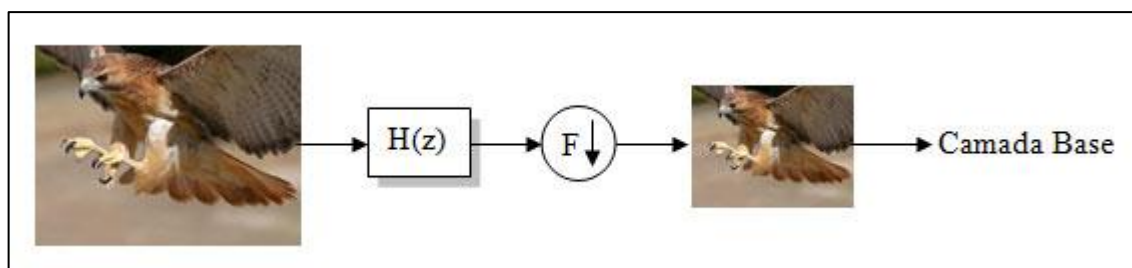


Figura 7. Processo de subamostragem.
Fonte: SOLER, L. (2006, p. 32).

A partir da camada base é realizada uma projeção da camada seguinte, conforme representado na Figura 8, por meio de uma sobreamostragem e um filtro de passo alto. A imagem projetada subtraída da original fornece um erro de previsão, este erro por sua vez possui as frequências mais altas da imagem.

As imagens naturais tem a tendência a se concentrar em níveis de frequência mais baixos, com isso o resultado da estrutura traz dois níveis, um deles é a representação básica da imagem e um erro de projeção quantificado que permite otimizar a qualidade (RICHARDSON, 2003).

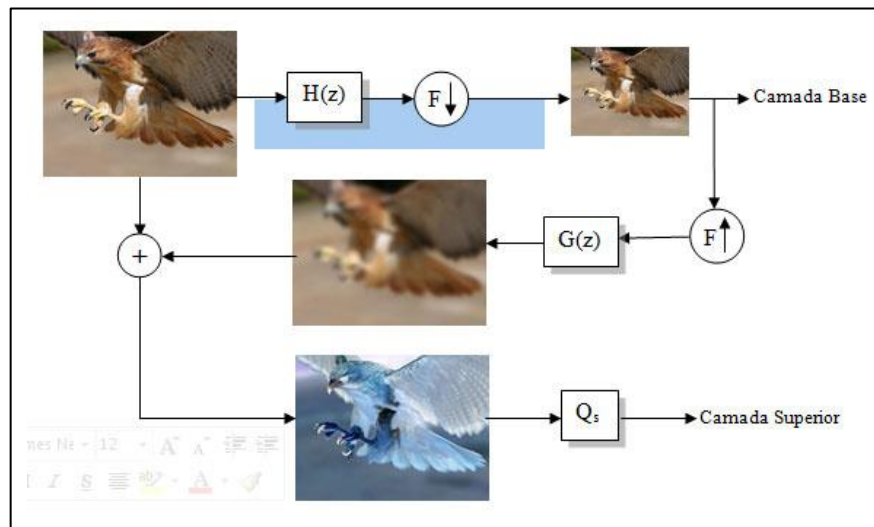


Figura 8. Projeção para camada de enriquecimento.
 Fonte: SOLER, L. (2006, p. 33)

O progresso dos níveis de bitstream, de acordo com a Figura 9, ressalta que a codificação e decodificação aplicada na técnica piramidal têm como base duas pirâmides. A primeira chamada pirâmide de Gauss possui esta denominação, pois seu filtro contém um modelo matemático semelhante ao de Gauss.

A pirâmide de Laplace representa a cadeia de imagens diferença, ou seja, os erros de projeção que possibilitam a reconstrução da imagem original conforme já descrito. Esta pirâmide se chama laplaciana, pois tem comportamento muito semelhante aos operadores de Laplace usados no processamento de imagens também referido com este nome (ASCENSO, 2003).

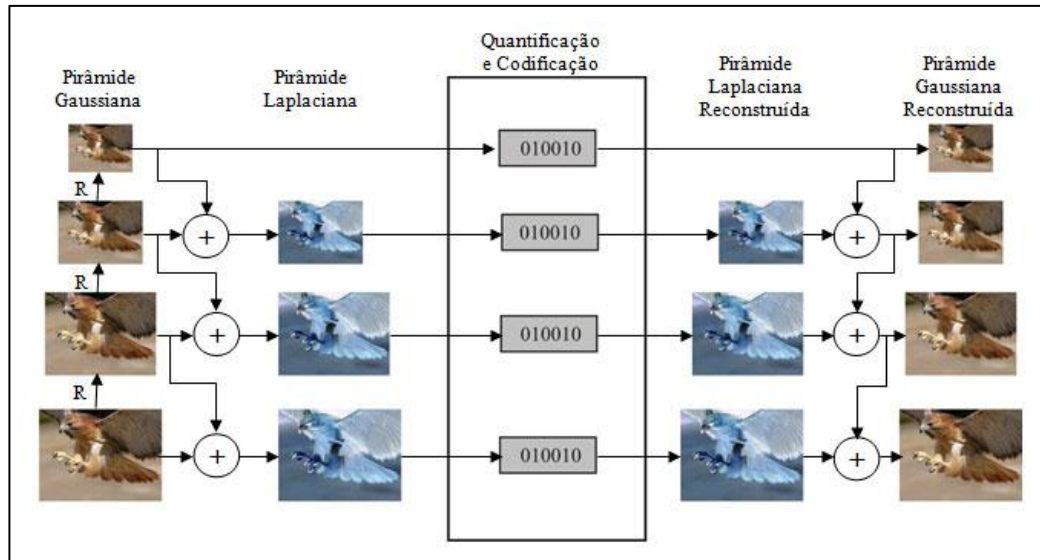


Figura 9. Progressão escalar por meio da codificação piramidal.
 Fonte: SOLER, L. (2006, p. 34)

3.2.2 Codificação em Sub-bandas

Este método consiste na decomposição do sinal em pequenas partes por meio de filtros em pares que permitem a recomposição do sinal com um grau de fidelidade perfeito (ou muito próximo disso) (SOLER, 2006).

Conforme a Figura 10 ocorre à divisão do sinal em duas partes, onde os processos de filtragem e subamostragem (H e G) conferem a cada parte da divisão metade das amostras.

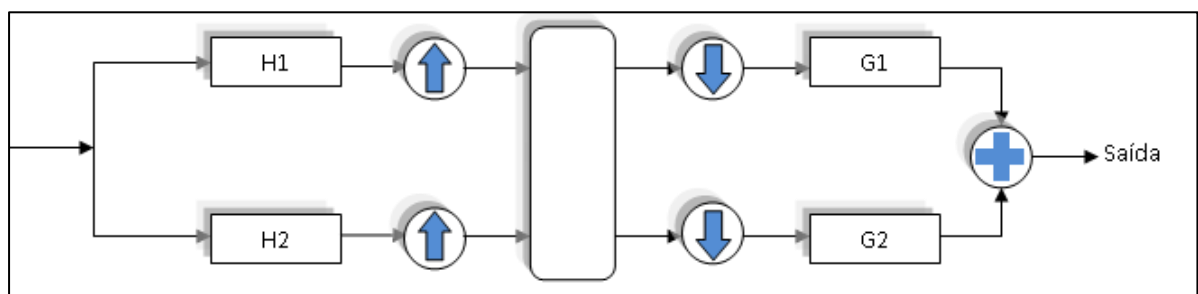


Figura 10. Filtros de sub-bandas.
 Fonte: ASCENSO, J. M. D. (2003, p. 33)

Assim como acontece na codificação piramidal, a primeira parte do sinal decomposto traz uma camada de base (representação de menor qualidade), e a outra parte do sinal se combina com a base formando o sinal original. A partir deste conceito da divisão em sub-bandas e na recepção reconstituir o sinal original, visa reduzir as dificuldades da transmissão em tempo real.

3.2.3 Transformada DCT

A transformada discreta do cosseno (DCT) é um processo aplicado em codificação de vídeos digitais aplicado a blocos de imagem e geralmente associado métodos de compensação de movimento. Esta técnica é utilizada por inúmeros padrões de codificação, tais como: JPEG, H.261, MPEG-1 e MPEG-2.

Diferentemente de outras técnicas a transformada DCT, não pode ser representada em camadas, essa característica criou a necessidade a criação de técnicas para codificação eficiente e progressiva (SOLER, 2006). A seguir serão apresentadas as principais técnicas, referentes ao contexto desta pesquisa, baseadas nas propriedades da transformada DCT:

- a) seleção de *frames*: a seleção de quadros permite criar uma representação escalar em camadas. Este método permite aumentar a frequência de acordo com incremento das camadas, selecionando um numero maior de quadros. Esta técnica manutenção escalar temporal de um vídeo digital;
- b) seleção de coeficientes: a transformada DCT é uma técnica que organiza as frequências de um bloco de imagens em um vetor, em ordem crescente de frequência. Inúmeras técnicas escalares tiram

partido desta característica, utilizando filtros de domínio de frequência;

- c) seleção de componentes: esta metodologia visa dividir em camadas os componentes responsáveis por características dos vídeos digitais, como crominância, luminância e componentes para representação de cores. Esta técnica escalar permite adequar os vídeos digitais a terminais com diferentes resoluções de cores;
- d) re-quantificação: a quantificação em etapas dos coeficientes DCT pode dividir a qualidade do vídeo em camadas. Esta técnica trabalha a escalabilidade SNR, para melhor usabilidade dos vídeos digitais de acordo com os requisitos da aplicação;
- e) DCT piramidal: os coeficientes DCT possuem uma característica muito utilizada neste método escalar a capacidade de subamostragem e sobreamostragem de frequência. Segundo Ascenso (2003, p43), esta propriedade influi diretamente na geração de camadas com resoluções espaciais diferentes, similar à codificação piramidal, mas seu desempenho é superior.

4 PADRÕES DE CODIFICAÇÃO

Uma técnica de codificação deve ser padronizada para que os diversos equipamentos existentes no mercado possam compreendê-la, independentemente do fabricante. Segundo Manoel (2007, p.3), os primeiros esforços para a padronização da codificação de vídeo digital datam do final da década de 80 e início da década de 90, com a criação do padrão H.261 pelo ITU-T.

4.1 EVOLUÇÃO DOS PADRÕES DE CODIFICAÇÃO

Nos últimos anos, tem se verificado uma grande evolução nos padrões internacionais para compressão e codificação de vídeo. Uma consequência desta padronização é o crescente interesse em técnicas de compressão mais complexas e eficientes, que combinem altas taxas de compressão com alta qualidade visual. Diversos esquemas de codificação digital de vídeo têm sido propostos procurando atender às mais diferentes aplicações, classificadas de acordo com a relação entre taxa de *bits* e qualidade requerida.

O acelerado crescimento no desenvolvimento da codificação e compressão de vídeos digitais é dado pelo surgimento de organizações com o objeto de padronizar as técnicas de codificação, criando normas internacionais facilitando a utilização de vídeos em seus diversos meios de aplicação (NUNES, 2009). Dentre estas organizações podemos destacar o *Moving Picture Experts Group* (MPEG) e o *International Telecommunication Union* (ITU).

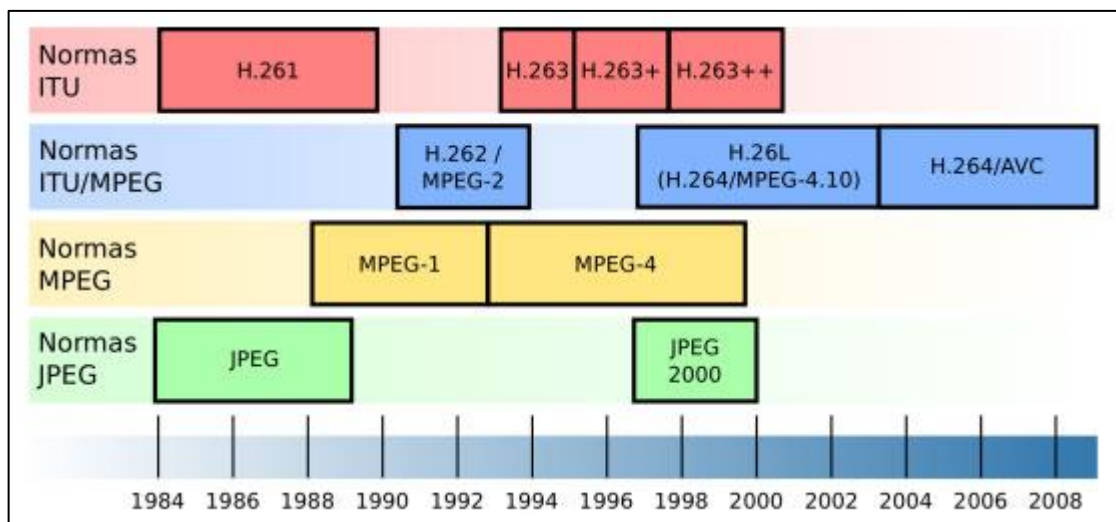


Figura 11. Evolução dos padrões de codificação de vídeo.
 Fonte: NUNES, R. P. (2009 , p. 39)

A Figura 11 traz a sequência evolutiva dos padrões de codificação de vídeo normalizado pela ITU e MPEG, tanto as normas exclusivas de cada órgão quanto às desenvolvidas em parceria.

A seguir descrevem-se brevemente os padrões criados como normas MPEG:

- a) MPEG-1 (1992): foi o primeiro padrão codificação de vídeo MPEG, sendo utilizado para Video CD, VoD, *streaming*, entre outros;
- b) MPEG-2 (1994): como no MPEG1 era um padrão de codificação de vídeo e áudio, além de estabelecer alguns mecanismos de transporte de dados. Com intuito de ser transmitido por televisão via broadcast, foi utilizado em TV Digital ATSC, ISDB e DVB, TV a cabo, também em SVCD's e DVD's ;
- c) MPEG-3: tinha objetivo de ser utilizado em televisão de alta definição (HDTV), mas acabou não sendo utilizado, pois foi comprovado que o MPEG-2 poderia ser utilizado;
- d) MPEG-4 (1998): Obtém um ganho considerável de compressão de vídeo em relação ao MPEG-2, além da codificação de vídeo e áudio

engloba conteúdo 3D, direitos autorais, objetos audiovisuais, além de outros benefícios.

Após essa breve consideração pelos padrões MPEG, a seguir é feita a mesma abordagem, contudo refere-se aos padrões desenvolvidos pela ITU:

- a) H.120 (1984): o primeiro padrão de codificação de vídeo digital. Continha técnicas básicas como quantização escalar, VCL e compensação de movimento;
- b) H.261 (1990): foi primeiramente projetado para trabalhar em linhas de transmissão de 64 *kbit/s*, porém suportava taxa entre 40 *kbit/s* e 2 *Mbit/s*. Acabou sendo o primeiro padrão de codificação utilizado em grande escala em videoconferências;
- c) H.262 (1994): desenvolvido em conjunto com MPEG, é idêntico ao padrão MPEG-2;
- d) H.263 (1995): projetado para aplicações em videoconferências, é uma expansão do padrão H.261, e baseado nos padrões MPEG-1 e 2;
- e) H.263+ (1998): segunda versão H.263 obteve um ganho na eficiência de codificação além da inclusão algumas funcionalidades;
- f) H.263++ (2000): assim como a versão anterior tem algumas melhorias em relação ao processo de codificação;
- g) H.264 (2003): considerado por muitos como o estado da arte em codificação de vídeo é muito semelhante ao padrão MPEG-4. Este padrão alcançou este patamar, com novas aplicação no processo de codificação com um equilíbrio entre custo, complexidade e codificação.

4.2 PADRÃO H.264/AVC

O padrão H.264/AVC - *Advanced Video Coding* é atualmente o estado-da-arte em codificação de vídeo. Ele alcança taxas de compressão superiores aos padrões anteriores, tais como o H.263 (ITU-T, 2005).

O padrão H.264 possibilita que vídeos digitais sejam manipulados por diferentes tipos de dispositivos, tais como dispositivos móveis, aparelhos celulares, com baixas resoluções, e aparelhos de TV, de alta definição. É requerida, portanto a transmissão de vídeo em diferentes resoluções de vídeo (escalabilidade espacial e escalabilidade temporal) compatíveis com estes tipos de dispositivos.

O projeto H.264/AVC veio com intuito de criar um padrão de codificação com boa qualidade de vídeo mesmo trabalhando com uma taxa de bits mais baixa que os padrões até então existentes, MPEG-2 e H.263, por exemplo.

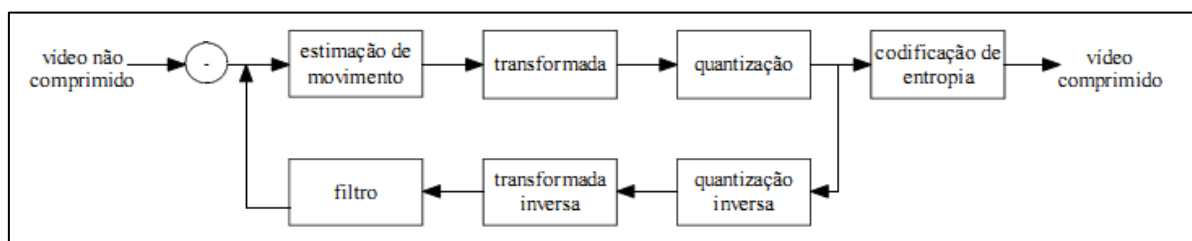


Figura 12. Codificador H.264/AVC
Fonte: CARVALHO, M. G. (2007, p. 19)

No padrão H.264/AVC cada quadro é dividido em blocos de 16x16 *pixels*, denominados macro blocos. No processo de codificação/decodificação deste padrão, a interação é feita a cada macro bloco (CARVALHO, 2007).

Seguindo o processo descrito na Figura 12 o vídeo a ser codificado passa primeiramente pelo módulo de estimação de movimento. Nesta primeira etapa é realizada uma busca dos quadros em relação ao quadro de referência, visando mapear as diferenças

criando um vetor de movimento. A verificação é feita por meio de algoritmos tanto do ponto de vista temporal quanto espacial, esta consolidação dos dados diferenciais é denominada de resíduo.

Os processos seguintes de transformada e quantização são aplicados para redução das redundâncias psico-visuais, ou seja, que não são sensíveis ao olho humano. Após este processamento é feita a codificação de entropia, onde algoritmos específicos são aplicados a forma de codificação dos dados visando otimizar o vídeos em seu processamento futuro.

Dentro do padrão H.264 existem os modelos espaciais e temporais explorando suas respectivas redundâncias, como o objetivo geral desta pesquisa abrange as escalas temporais e espaciais a seguir serão apresentados estes modelos.

4.3 EXTENSÃO ESCALÁVEL DO PADRÃO H.264

O padrão H.264 possui uma extensão desenvolvida que visa o suporte a escalabilidade denominado H.264/SVC – *Scalable Video Coding*. Esta extensão possibilita a adição ou retirada de camadas dentro do fluxo de *bits*, manipularem-se a escala de acordo com a aplicação. O H.264/SVC provê suporte a múltiplas resoluções de exibição com um único *bitstream* (WIEGAND, 2007, tradução nossa).

Adicionalmente, o H.264/SVC também suporta escalabilidade temporal, ou seja, alguns quadros podem ser descartados além da escalabilidade de qualidade, que é o enriquecimento na fidelidade dos quadros mantendo-se a resolução entre as camadas (SEGALL; SULLIVAN, 2007, tradução nossa).

A extensão SVC do padrão de codificação H.264, busca manter a alta qualidade alcançada na extensão H.264/AVC, mas subdividindo em uma ou mais camadas de *bitstreams*

intermediários que possam ser decodificados separadamente (SCHWARZ; MARPE; WIEGAND, 2007, tradução nossa).

A Figura 13 traz uma estrutura básica do padrão H.264/SVC com duas camadas, onde nos mecanismos de predição é gerado um resíduo visual, este resíduo é aplicado à camada de enriquecimento. O resíduo é a representação das modificações ocorridas no vídeo digital, ou seja, o movimento realizado no vídeo é extraído para ser codificado o enriquecimento escalar.

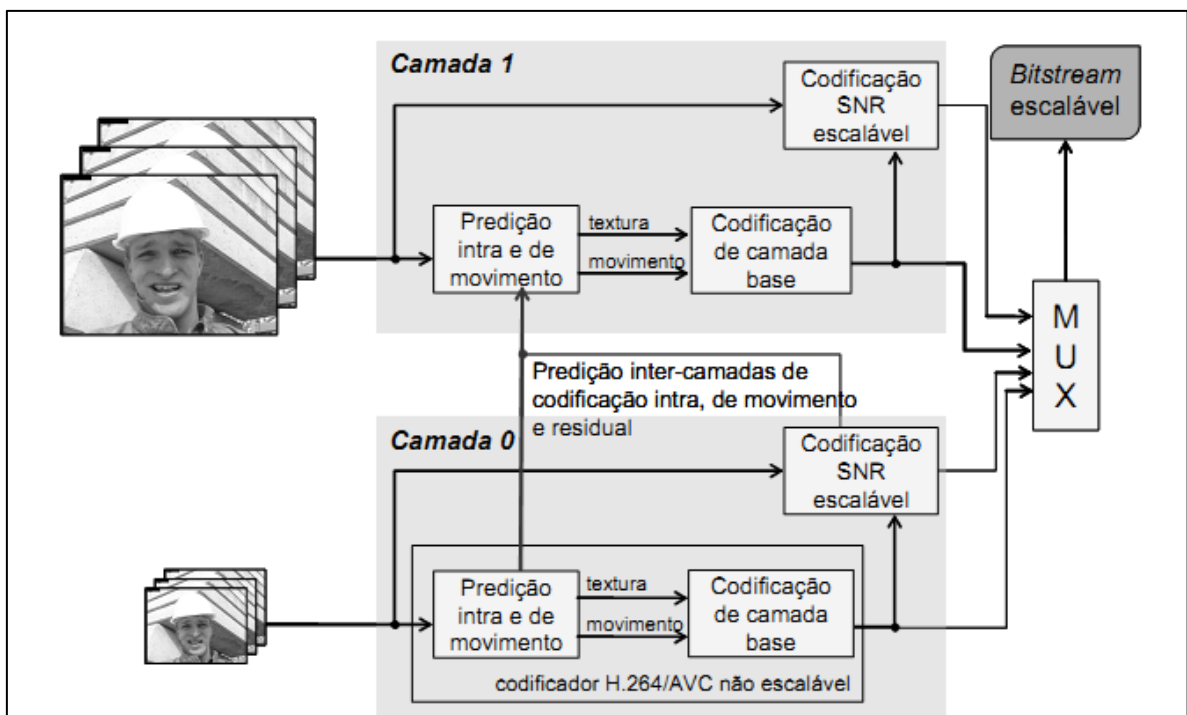


Figura 13. Estrutura do codificador SVC
Fonte: FONSECA, C. (2008, p. 27)

No padrão SVC, de maneira simples como assim foi objetivado o projeto, pode-se trabalhar os três tipos escalares aqui descritos: espacial, temporal e de qualidade. A seguir será aprofundada a maneira como o H.264/SVC trata essas classificações escalares.

4.3.1 Escalabilidade Espacial

A codificação de vídeo escalável espacial segue a abordagem tradicional de codificação em multicamadas para permitir a escalabilidade espacial. Esta técnica era empregada também nos padrões anteriores como, H.262/MPEG-2, H.263 e MPEG-4 (SCHWARZ; MARPE; WIEGAND, 2007, tradução nossa).

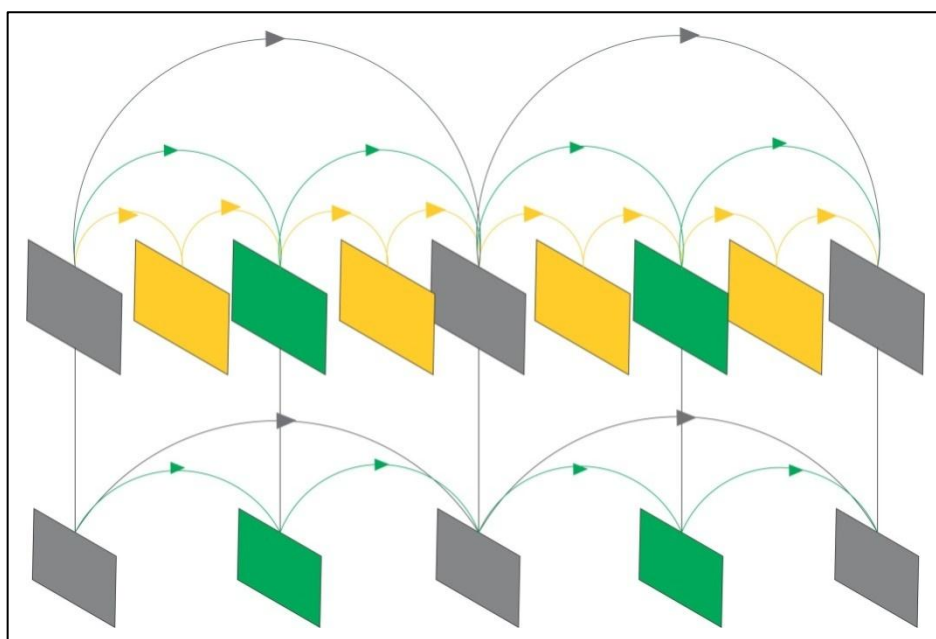


Figura 14. Comportamento da escalabilidade espacial no H.264/SVC.
Fonte: SCHWARZ, H.; WIEGANDM, T. (2007, p.1109)

Nesta classificação escalar espacial do padrão H.264/SVC, utiliza mecanismos de predição entre camadas visando aperfeiçoar a codificação da camada de enriquecimento em relação ao *simulcast* (WIEN; SHWARZ; OELBAUM, 2007, tradução nossa). Este comportamento pode se verificado na Figura 14, onde a camada base é representada pelos quadros em cinza e as restantes são camadas adicionais.

O padrão SVC possibilita a camada adicional ser maior do que a camada inferior ou representar apenas uma região de interesse, aplicando o conceito de escalabilidade por região de interesse (DARONCO, 2009).

O SVC alcançou melhorias na complexidade computacional e nos requisitos de memória, por via de algumas regras. Uma dessas regras é a ordem das figuras nas camadas escaláveis, deve ser a mesma em todos os níveis (NUNES, 2009).

4.3.2 Escalabilidade Temporal

O padrão H.264, por meio de seu controle de memória de quadros de referencia, permite codificar seqüências de quadros com dependências temporais arbitrárias. Com isso pode suportar um número de camadas razoável sem alteração do projeto (SHWARZ; MARPE; WIEGAND, 2007, tradução nossa).

A maior vantagem que o padrão SVC traz com relação a escalabilidade temporal são as múltiplas referencias para estimação do movimento na codificação de um *frame*, uma questão bem restrita em outros padrões de codificação (DARONCO, 2009) .

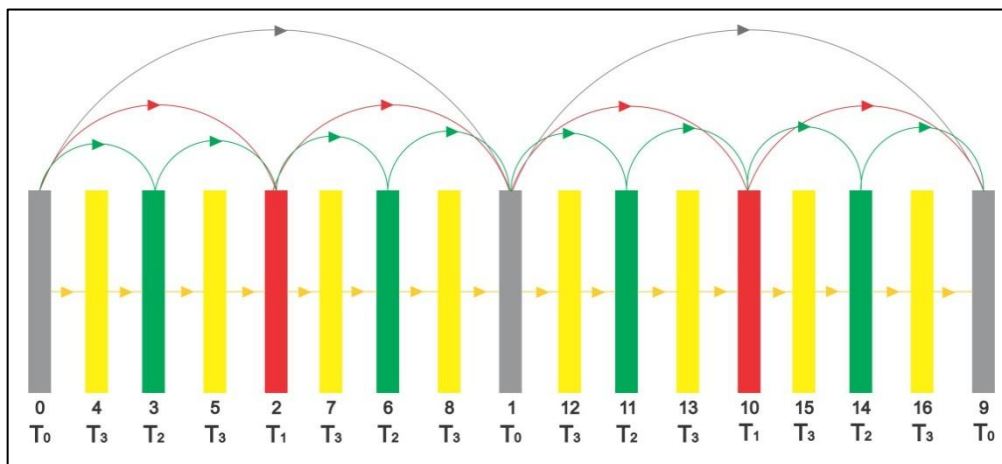


Figura 15. Comportamento da escalabilidade temporal no H.264/SVC.
 Fonte: SCHWARZ; MARPE; WIEGAND, 2007

Na Figura 15 apresenta uma representação da escalabilidade temporal no padrão SVC, a estrutura por ela representada é denominada predição hierárquica, pelos níveis estipulados para as camadas sequenciais.

As cores dispostas na Figura 15 são as camadas com suas imagens distintas. Sendo que toda camada de enriquecimento é codificada com base nas camadas anterior e posterior. A camada base é composta pelas Figuras T_0 onde as camadas T_x (com $x > 0$), são as camadas de enriquecimento.

4.3.3 Escalabilidade por Qualidade (SNR)

No padrão H.264/SVC a codificação escalar de qualidade tem como base o conceito de *Coars-Grain Quality Scalability* (CGS). Segundo Daronco (2009, p.33), este processo consiste em manter a escala temporal e espacial, com o mesmo processo de predição *inter-layer* da codificação escalar espacial, contudo não necessidade de redimensionamento.

As técnicas de predição nas camadas SNR são as mesmas do processo escalar espacial, mas sem a utilização das etapas de *upsampling* e *deblocking* em macroblocos com predição intra na camada referencial, contudo a resolução espacial é mantida entre a camada base e a camada de enriquecimento (FONSECA, 2008).

5 TRABALHOS CORRELATOS

Apesar do padrão de codificação de vídeo H.264/SVC ser muito recente, há diversas linhas de pesquisa para o seu estudo, a seguir foram escolhidos três trabalhos que tiveram pontos em comum com esta pesquisa.

5.1 OVERVIEW OF THE SVC EXTENSION H.264/AVC

Este artigo foi um dos primeiros artigos publicado em 2007, após a padronização da extensão SVC do padrão H.264 de codificação de vídeo publicado pelos autores SCHWARZ, MARPE e WIEGAND. Trata da descrição desta aplicação além de testes de desempenho, além de uma visão geral dos conceitos básicos da prorrogação H.264/AVC para SVC.

As ferramentas básicas para o trabalho das escalas temporais, espaciais, escalabilidade e qualidade são descritas em detalhes e analisadas experimentalmente quanto a sua eficiência e complexidade.

O artigo trata da evidente extensão do H.264 considerando desempenho da taxa de distorção ótima com relação aos padrões anteriores que continham o suporte a escalabilidade, por requerer apenas um único circuito de compensação de movimento no lado do decodificador.

5.2 CODIFICAÇÃO ESCALÁVEL DE VIDEO PARA RECEPÇÃO FIXA DE SBTV

Com o advento da TV digital a partir de 2008, surgiram estudos diversos sobre esta nova tecnologia a tese de mestrado de Rogério Pernas Nunes trata a codificação escalável neste contexto.

A escalabilidade com relação aos formatos de vídeo no SBTVD juntamente com evolução de dispositivos com resoluções melhoradas são pontos retratados neste trabalho. A diversidade de aplicações da TV digital é tratada como fator principal para o estudo escalar do padrão H.264/AVC.

Uma das constatações desta pesquisa é a vantagem significativa da escalabilidade sobre uma transmissão simultânea equivalente, evidenciando uma possível adoção do padrão SVC no SBTVD.

5.3 AVALIAÇÃO SUBJETIVA DE QUALIDADE APLICADA A CODIFICAÇÃO DE VIDEO ESCALÁVEL

Esta tese de mestrado de Leonardo Crauss Daronco é uma análise subjetiva sobre vídeos experimentais codificados no padrão H.264/SVC. Uma pesquisa justificada pelos diversos dispositivos, redes e meios de comunicações existentes.

O objetivo principal desta tese é verificar se o padrão H.264/SVC permite chegar a uma transmissão universal ou o mais próximo disso devido à diversidade de dispositivos e redes existentes.

Ao final do trabalho é constatada que a instabilidade não provoca grandes danos à qualidade subjetiva de um vídeo, mas que o método de escalabilidade temporal entre as classificações de escalabilidade é o mais instável.

6 AMBIENTE DE SIMULAÇÃO

6.1 H.264/SVC JSVM REFERENCE 9.19

O JSVM 9.19 (*Joint Scalable Video Model*) é o software referência do projeto SVC, este software foi desenvolvido em C++ por membros do JVT para o tratamento do padrão H.264/SVC.

Este software contém uma série de executáveis que permitem trabalhar as características do padrão H.264. A seguir será descrito brevemente o objetivo de cada executável:

- a) DownConverStatic: neste deste executável pode se interferir na resolução de um vídeo digital;
- b) H264AVCEncoderLibTestStatic: permite a codificação de um vídeo sem escalabilidade ou com escalabilidade;
- c) H264AVCDecoderLibTestStatic: decodifica vídeos tanto em formato escalável quanto não escalável, contudo quando for necessária a decodificação de uma camada intermediária o aplicativo BitStreamExtractorStatic dever ser utilizado;
- d) AvcRewriterStatic: retira o suporte a escalabilidade de um vídeo digital, mas para isso o vídeo no formato SVC precisa ser codificado com suporte a esta aplicação;

- e) `BitStreamExtractorStatic`: este aplicativo permite extrair sub-bitstreams , ou seja, extrair um vídeo com menor fluxo de bits em relação ao original;
- f) `QualityLevelAssignerStatic`: possibilita a adição de informações na camada de qualidade de um vídeo digital;
- g) `MCTFPreProcessorStatic`: utilizado para pré-filtrar um vídeo antes de codificá-lo, fazia parte do fluxo normal de codificação, mas foi retirado e ficou como opção neste executável;
- h) `PSNRStatic`: este executável tem como umas das principais funções o cálculo do fluxo de bits e do PSNR, métricas citadas no item 2.4;
- i) `FixedQPEncoderStatic`: permite ajustar o fluxo de bits e controlar o codificador H.264;
- j) `SIPAnalyser`: usado para escolha da predição entre as camadas do vídeo digital.

Dentre estes executáveis descritos três foram necessário na realização dos experimentos buscando avaliar o desempenho da extensão SVC do padrão de codificação H.264, os aplicativos utilizados nos experimentos serão descritos a seguir.

6.2 CODIFICADOR “H264AVCENCODERLIBTESTSTATIC”

Como já descrito no item 6.1, o codificador contido no pacote JSVM permite a manipulação de vídeos no padrão H.264/AVC (sem escalabilidade) e o H.264/SVC (com suporte escalar). Este aplicativo tem como saída arquivos reconstruídos no formato YUV e um arquivo compactado no formato H.264.

A chamada do aplicativo deve conter no mínimo os parâmetros estabelecidos na Figura 16.

```
H264AVCEncoderLibTestStatic.exe -pf <mcfg> [command line options]
```

Figura 16. Comando básico do codificador JSVM.
Fonte: Manual JSVM 9.19

O parâmetro “-pf” permite a especificação do arquivo “main” do codificador, responsável pelo parâmetros macro do vídeo a ser codificado, ou seja, todos parâmetros comuns entres as camadas dos vídeos escalares. O arquivo “main” também tem por função especificar os arquivos de parâmetros das camadas escalares espaciais da sequência codificada.

O processo de codificação do codificador de referência possui opções de simulação para uma ou mais camadas, podendo-se modificar a escala espacial, temporal ou de qualidade. A seguir serão descritos os arquivos de configuração necessários para uma simulação escalar com duas camadas de codificação.

6.2.1 Configuração do Arquivo “main”

Na Tabela 5, demonstra os principais parâmetros estudados no arquivo *main* para a realização das simulações. Nele pode-se verificar a existência de parâmetros gerais do vídeo, como a codificação da sequência em formato AVC ou SVC, o arquivo de saída com todas as camadas do vídeo codificado no formato H.264, além da especificação dos arquivos de configuração das *layers*.

Tabela 5. Arquivo de Configuração "Main".

Parâmetro	Valor	Descrição
AVCMode	0	SVC: 0; AVC: 1
OutputFile	exemplo.164	Arquivo de Saída, formato H.264
FrameRate	30.0	Frames por segundo
FramesToBeEncoded	375	Número de frames a serem codificados
GOPSize	1	Número de camadas em escala temporal
NumLayers	2	Numero de camadas em escala espacial
LayerCfg	layer_0.cfg	Arquivo de configuração de cada camada.
LayerCfg	layer_1.cfg	Arquivo de configuração de cada camada.

6.2.2 Configuração das Camadas (*Layers*)

No formato de vídeo multicamada espacial é necessária a configuração de cada *layer*, nas Tabelas 6 e 7 são apresentados os parâmetros de grande importância neste estudo. Nelas pode-se observar que o simulador utilizado necessita de vídeos em no mínimo duas resoluções espaciais, para possibilitar uma simulação espacial.

Tabela 6. Arquivo de configuração da primeira camada “layer0”.

Parâmetro	Valor	Descrição
SourceWidth	176	Largura da sequência de entrada
SourceHeight	144	Altura da sequência de entrada
InputFile	exemplo_qcif.yuv	Sequência de entrada específica da camada.
ReconFile	exemplo_qcif_layer0.yuv	Arquivo YUV reconstruído
QP	32.0	Parâmetro de quantização

Tabela 7. Arquivo de configuração da segunda camada “layer1”.

Parâmetro	Valor	Descrição
SourceWidth	352	Largura da sequência de entrada
SourceHeight	288	Altura da sequência de entrada
InputFile	exemplo_cif.yuv	Sequência de entrada específica da camada.
ReconFile	exemplo_cif_layer1.yuv	Arquivo YUV reconstruído
QP	32.0	Parâmetro de quantização
BaseLayerId	0	Camada base da sequência SVC

6.2.3 Formato de Saída

O codificador H264AVCEncoderLibTestStatic tem como saída um arquivo no formato H.264, definido pelo arquivo de configuração *main* (vide Tabela 5, “exemplo.264”) e um arquivo reconstruído no formato YUV para cada camada espacial codificada. Além dos arquivos de saída este codificador, também fornece dados estatísticos em tempo de codificação conforme demonstrado na Figura 17.

	C1		C2	C3	C4	C5	C6			C7
AU	0:	I	T0	L0	Q0	QP 40	Y 28.1288	U 40.0329	U 41.0214	5600 bit
	0:	I	T0	L1	Q0	QP 51				472 bit
AU	0:	I	T0	L1	Q0	QP 51	Y 25.0604	U 39.6273	U 40.7215	472 bit
	1:	P	T0	L0	Q0	QP 40	Y 28.0311	U 40.0699	U 41.0006	368 bit
AU	1:	P	T0	L1	Q0	QP 51				288 bit
	1:	P	T0	L1	Q0	QP 51	Y 24.9806	U 39.7096	U 40.7333	296 bit
AU	2:	P	T0	L0	Q0	QP 40	Y 27.9201	U 40.0756	U 41.1214	496 bit
	2:	P	T0	L1	Q0	QP 51				280 bit
AU	2:	P	T0	L1	Q0	QP 51	Y 24.8737	U 39.7804	U 40.8413	288 bit
	3:	P	T0	L0	Q0	QP 40	Y 27.7999	U 40.1163	U 41.0817	480 bit
AU	3:	P	T0	L1	Q0	QP 51				312 bit
	3:	P	T0	L1	Q0	QP 51	Y 24.7980	U 39.8427	U 40.8043	320 bit
AU	4:	I	T0	L0	Q0	QP 40	Y 28.1385	U 40.5202	U 40.4745	5784 bit
	4:	I	T0	L1	Q0	QP 51				440 bit
AU	4:	I	T0	L1	Q0	QP 51	Y 25.0112	U 40.1215	U 40.1716	448 bit
	5:	P	T0	L0	Q0	QP 40	Y 28.0586	U 40.5253	U 40.4126	408 bit
AU	5:	P	T0	L1	Q0	QP 51				296 bit
	5:	P	T0	L1	Q0	QP 51	Y 24.9399	U 40.1489	U 40.1673	304 bit
AU	6:	P	T0	L0	Q0	QP 40	Y 27.9066	U 40.4821	U 40.2875	552 bit
	6:	P	T0	L1	Q0	QP 51				280 bit
AU	6:	P	T0	L1	Q0	QP 51	Y 24.7923	U 40.0832	U 40.0013	288 bit
	7:	P	T0	L0	Q0	QP 40	Y 27.8325	U 40.4455	U 40.0280	696 bit
AU	7:	P	T0	L1	Q0	QP 51				312 bit
	7:	P	T0	L1	Q0	QP 51	Y 24.7239	U 40.0681	U 40.0155	320 bit

Figura 17 Saída de dados em tempo de codificação.

Fonte: Software JSVM 9.19 durante codificação.

A Tabela 8 traz os dados de saída em tempo de execução do H264AVCEncoderLibTestStatic de acordo com as colunas apresentadas na Figura 17. Ao fim da codificação o software fornece dados gerais da codificação como: taxa de *bits*, tempo de codificação, velocidade de codificação, PSNR e o número de frames codificados. Todas estas informações são separadas de acordo com as camadas estipuladas nos arquivos de configuração, sejam elas em escala temporal, espacial ou de qualidade, conforme demonstrado na Figura 18.

Tabela 8. Tabela saída de dados JSVM.

Nome	Finalidade
C1	Número do quadro
C2	Camada em escala temporal
C3	Camada em escala espacial
C4	Camada em escala de qualidade
C5	Parâmetro de quantização estipulado para a camada
C6	Valores de PSNR
C7	Bit usados na codificação.

SUMMARY :						
	bitrate	Min-bitr	Y-PSNR	U-PSNR	U-PSNR	
176x144 @ 30.0000	79.7912	79.7912	27.3628	38.7202	39.9351	
352x288 @ 30.0000	108.3504	108.3504	24.0902	38.4304	39.5772	
Encoding speed: 631.013 ms/frame, Time:189304.000 ms, Frames: 300						

Figura 18 Resultados apresentados ao final da codificação.
 Fonte: Software JSVM 9.19 ao fim da codificação.

6.3 APLICATIVO DE AMOSTRAGEM “BITSTREAMEXTRACTORSTATIC”

O executável BitStreamExtractorStatic permite a extração das camadas estipuladas na codificação do vídeo H.264/SVC, este processo utiliza como arquivo de entrada o vídeo gerado com extensão *.264 na codificação.

A seguir tem-se a descrição da chamada principal do aplicativo BitStreamExtractorStatic:

```
BitStreamExtractorStatic <in> <out> [-l] [-t]
```

Figura 19. Comando básico do extrator JSVM.
 Fonte: Manual JSVM 9.19]

Os parâmetros apresentados na Figura 19 fornecem os nomes dos vídeos de entrada e saída dispostos nos espaços nomeados como *in* e *out*, além dos parâmetros “-l” especificando a camada espacial desejada e o parâmetro “-t” que defini a camada temporal desejada.

6.4 DECODIFICADOR “H264AVCDECODERLIBTESTSTATIC”

O aplicativo H264AVCDecoderLibTestStatic é responsável pela decodificação dos vídeos codificados pelo software de referência JSVM, tanto para o padrão H.264/AVC quanto o SVC. A chamada deste executável é descrita na figura a seguir:

```
H264AVCDecoderLibTestStatic <str> <rec>
```

Figura 20. Comando básico do decodificador JSVM.
Fonte: Manual JSVM 9.19

O comando apresentado na Figura 20 recebe apenas os dois parâmetros demonstrados, onde “*str*” é o arquivo de entrada a ser decodificado no formato H.264, e o espaço reservado “*rec*” corresponde ao arquivo de saída no formato YUV.

6.5 PROCEDIMENTO EXPERIMENTAL

Neste Capítulo serão demonstradas e comentadas as sequências de vídeos utilizados nos testes escalares, as métricas para análise dos vídeos codificados, além dos dados utilizados como indicativos nas simulações.

6.5.1 Sequências de Testes

Os testes foram realizados com três sequências que permitiam as simulações das escalas espaciais e temporais. Para isso é necessária a disponibilidade dos vídeos em no mínimo duas resoluções espaciais.

Nesta pesquisa foram realizados testes com sequências nas resoluções CIF e QCIF, conforme descrito na Figura 19. Estas resoluções são utilizadas com frequência em estudos por via de softwares de simulação, além disso, são muito utilizadas em dispositivos móveis. As sequências utilizadas possuem bastantes movimentos, tanto de fundo quanto no primeiro foco do vídeo. A Figura 21 traz os vídeos utilizados e suas respectivas resoluções.




	<p>Coastguard</p> <p>Resolução Qcif 176x144 YUV 4:2:0 300 Frames</p>	<p>Coastguard</p> <p>Resolução Cif 352x288 YUV 4:2:0 300 Frames</p>
	<p>Foreman</p> <p>Resolução Qcif 176x144 YUV 4:2:0 300 Frames</p>	<p>Foreman</p> <p>Resolução Cif 352x288 YUV 4:2:0 300 Frames</p>
	<p>Walk</p> <p>Resolução Qcif 176x144 YUV 4:2:0 376 Frames</p>	<p>Walk</p> <p>Resolução Cif 352x288 YUV 4:2:0 376 Frames</p>

Figura 21. Informações das sequências utilizadas.

6.5.2 Métricas de Comparação

O nível de qualidade de um vídeo digital é influenciado por diversos fatores, medidas como PSNR e MSE são muito utilizadas para estimar a qualidade de um vídeo digital com base na diferença de luminância dos *pixels* comparado ao vídeo original. A definição destes critérios é muito importante para o decorrer deste trabalho.

O primeiro parâmetro é a comparação da eficiência da codificação, por meio da análise da redução do volume de dados após o processo de codificação de um determinado algoritmo de compressão (SANBORN, 2005).

Além disso, podemos destacar dois parâmetros de comparação vídeos digitais comumente utilizados na comparação de vídeos digitais. O Erro Médio Quadrático (*Mean*

Squared Error – MSE) é obtido conforme o cálculo representado na equação 3, onde i e j são duas imagens monocromáticas de dimensão m e n , uma imagem é a original a outra é a que está em avaliação.

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

A Razão Pico Sinal-Ruído (PSNR) é um dos parâmetros mais utilizados para a avaliação da qualidade objetiva de um vídeo digital (SANBORN, 2005). Muitas bibliografias também na análise dos seus padrões de codificação o PSNR, esta medição serve muito bem como parâmetro comparativo em testes em mesmo padrão de codificação. Na equação 4 a variável “MAX” corresponde ao valor máximo de luminância que o vídeo pode atingir.

$$PSNR = 20 \cdot \log \left(\frac{MAX}{\sqrt{MSE}} \right) \quad (4)$$

Outra característica levada em consideração nos critérios de avaliação da codificação de vídeos digitais é a complexidade computacional. Alguns métodos de codificação buscam muita precisão em seu desempenho, mas acabam inviabilizando sua utilização pelo tempo elevado de processamento, dificultando sua utilização em tempo real (DARONCO, 2009).

6.6 ANÁLISE DAS MÉTRICAS EM FUNÇÃO DA QUANTIZAÇÃO

Os testes realizados consistem na busca por meio da variação de quantização, verificar as características do mesmo. A quantização de um vídeo varia de 0 a 51, com isso foram estipulados intervalos para que fosse possível mapear o comportamento do padrão H.264/SVC. Vale ressaltar que o parâmetro de quantização com valor 51 é tido como mais grosseiro, progredindo até valor 0 onde se tem vídeos com melhor qualidade.

Os valores de quantização estipulados para as simulações foram 0, 16, 32 e 51. Sendo que nos testes com a variação de escala espacial pode se especificar o valor de quantização para cada camada, por meio de seu arquivo de configuração específico.

6.6.1 Dados Utilizados na Análise

Para analisar os resultados dos testes descritos, foram colhidas as informações de PSNR de Y, taxa de *bits* e tempo de codificação. Todos os valores estão condicionados a variação da quantização de codificação.

Cada sequência estipulada para teste gerou 147 condições codificação, onde cada teste geravam-se valores para cada camada do vídeo.

7 RESULTADOS E ANÁLISE

Neste Capítulo serão descritos os resultados e posterior análise obtidos a partir de simulações realizadas com o codificador de referência do padrão H.264/SVC. Será considerada codificação escalável espacialmente. Os resultados estão dispostos em gráficos visando facilitar a análise e compreensão, onde as grandezas apresentadas são de PSNR, tempo de codificação e taxa de *bits*, em função da variação da quantização dos vídeos. Nos gráficos contendo o desempenho de PSNR, foi adotado como padrão o valor de luminância Y, pois os demais valores UV dependem deste componente de brilho.

Os testes foram aplicados as três sequências descritas no Item 7.1 contemplando todos os quadros que constituem as mesmas. Todas as simulações utilizaram um vídeo na resolução QCIF juntamente com um na resolução CIF, onde o a primeira camada QCIF corresponde a “QP1” na análise dos resultados e a camada de incremento CIF a “QP2”. Nos gráficos demonstrados, existe a indicação *a*, *b* e *c* que correspondem as sequências *coastguard*, *foreman* e *walk* respectivamente.

7.1 TAXA DE BITS EM FUNÇÃO DA QUANTIZAÇÃO

Neste subitem serão apresentados os resultados de taxa de bits em função da quantização. Na Figura 22, pode se verificar nos testes com camada base codificada com a quantização entre 16 e 32 a taxa de bits se mantém sempre reduzida. Já nos valores extremos de QP, 0 e 51, a taxa de *bits* em determinados pontos se apresenta maior do que demais condições. Como isso pode se perceber a existência de uma zona de equilíbrio no parâmetro

de quantização da camada base, para que a camada de incremento tenha uma taxa de *bits* reduzida.

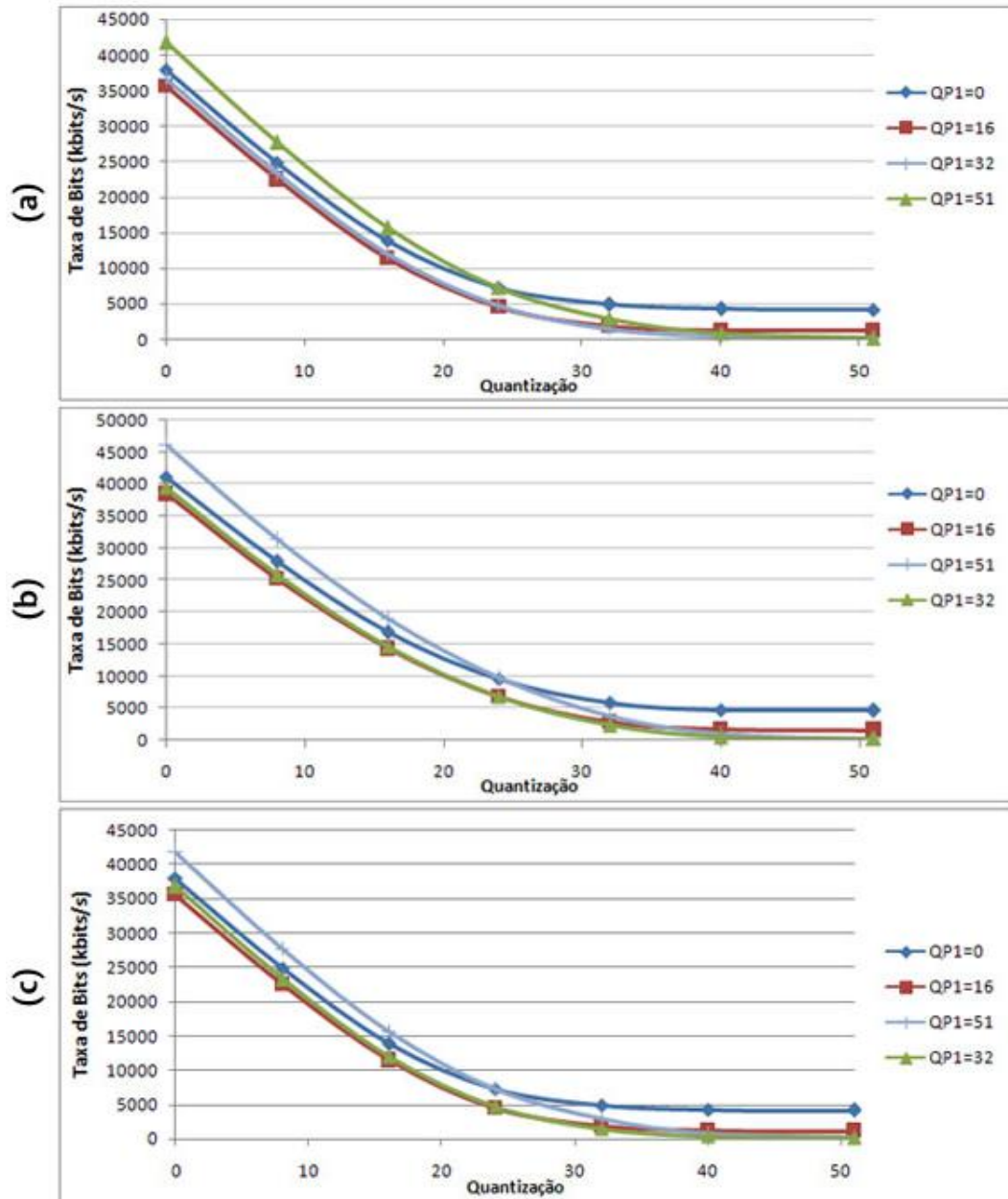


Figura 22. Taxa de bits em função da quantização.

Na Figura 23, se tem uma análise do ambiente demonstrado na Figura 22, focada na região onde as curvas começam tornar sua taxa de bits mais constantes. Pode se observar a evolução diferenciada da linha “QP=51”, cortando as demais em função do aumento da quantização.

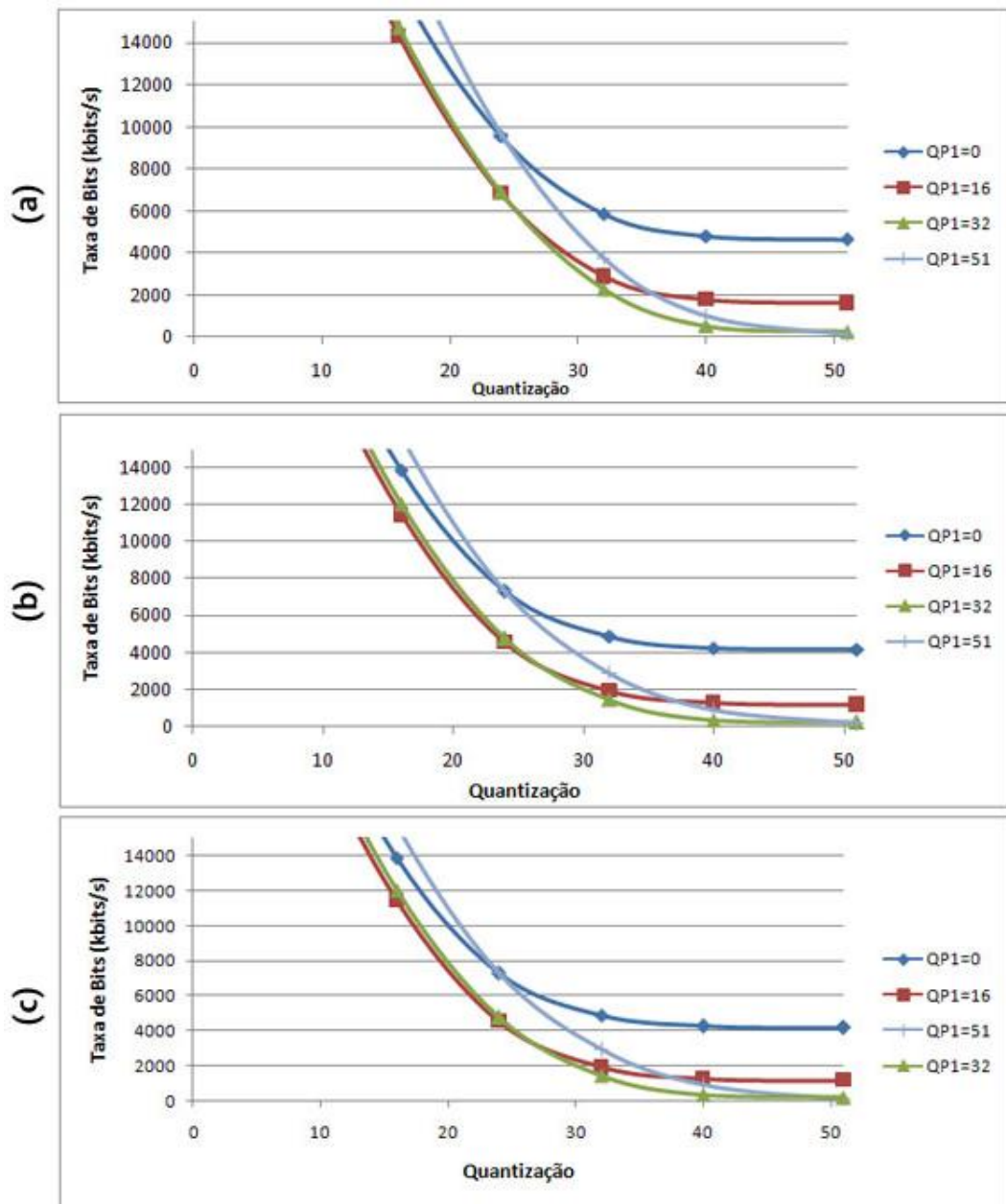


Figura 23. Interseção das curvas de taxa de bits.

7.2 PSNR EM FUNÇÃO DA QUANTIZAÇÃO

Para análise da PSNR foram coletados apenas os valores de Y, na Figura 24 são apresentados estes valores em função da variação de quantização. Pode se verificar a proximidade dos valores de PSNR nas linhas demonstradas até os pontos próximos de 32. Os

valores de quantização acima de 32 passam a obter uma qualidade considerada rudimentar. Com isso quando temos a quantização entre 0 e 32, uma oportunidade de ganho na codificação é demonstrada, pois se a qualidade se mantém próxima independentemente da camada base neste intervalo, outros indicadores como taxa de *bits* e tempo de codificação podem ser explorados.

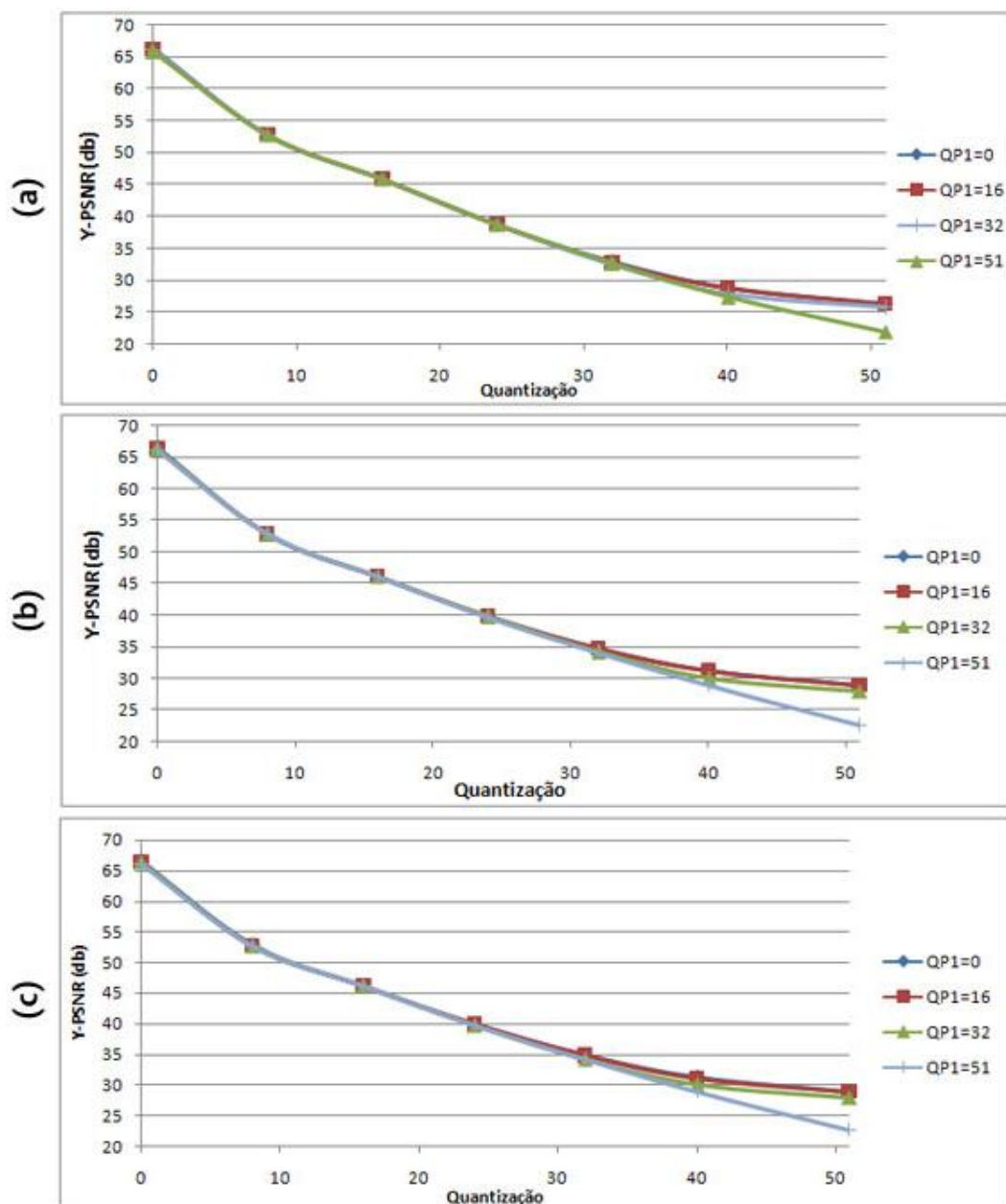


Figura 24. PSNR em função da quantização.

7.3 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA QUANTIZAÇÃO

A Figura 25 retrata se uma das oportunidades de ganho na codificação após a análise da PSNR, o tempo de codificação. No gráfico a seguir verifica se a dependência da camada base codificada para melhor desempenho de tempo de codificação. Quando codificada a camada base com quantização 0 (“QP1=0”), o tempo de codificação tende a ser maior. Porém, quando se codifica a camada base com quantização igual a 51, o tempo de codificação tende a ser menor.

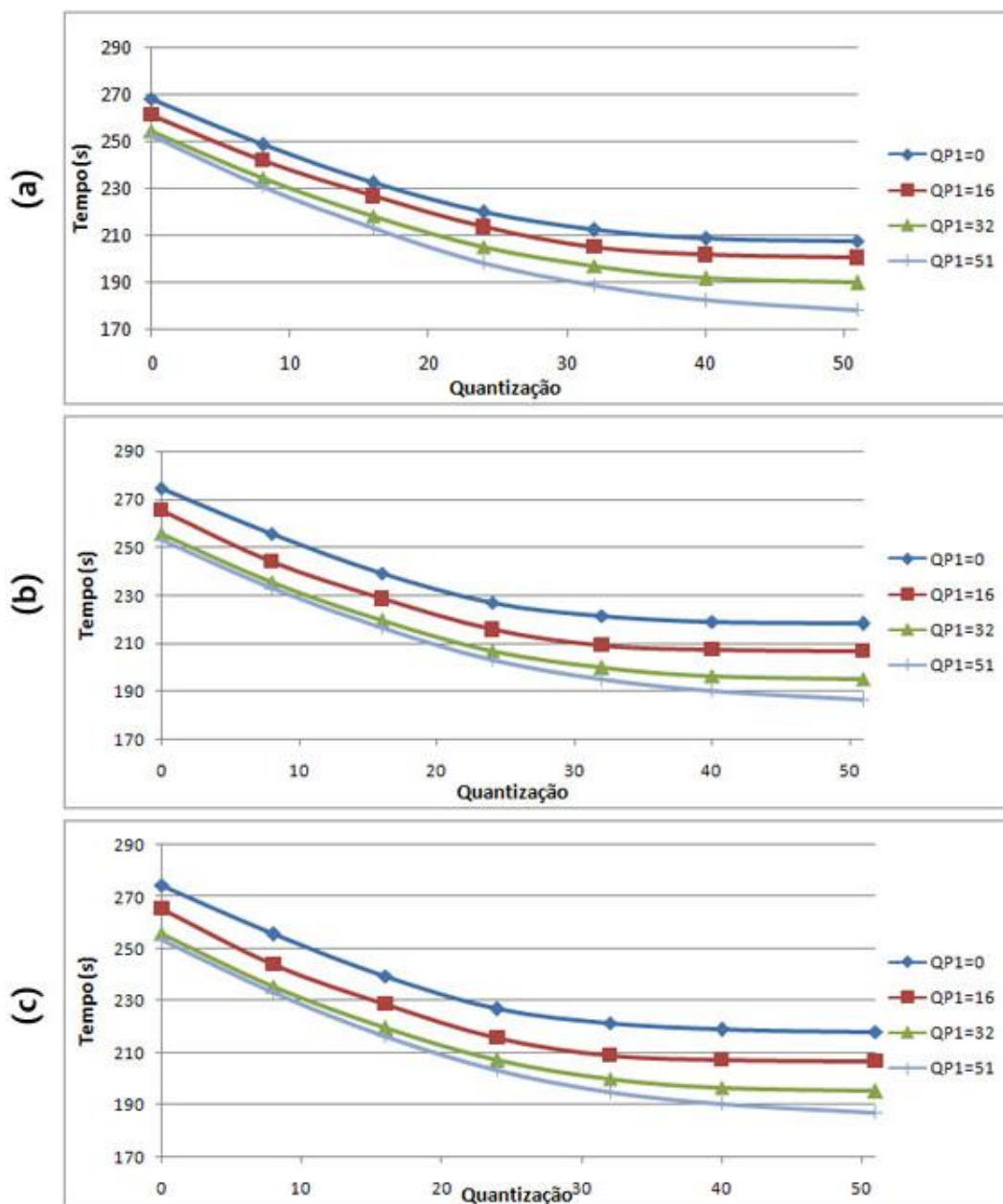


Figura 25. Tempo em função da quantização.

7.4 PSNR EM FUNÇÃO DA TAXA DE BITS

Os dados a seguir apresentados tratam da relação PSNR versus a taxa de *bits*. As linhas com quantização da camada base iguais a 16 e 32, apresentaram uma evolução semelhante com a taxa de *bits* sempre menor que as demais condições. Ao contrário dos vídeos codificados com camada base igual a 0 e 51, que resultaram em taxas de bits maiores em determinados pontos.

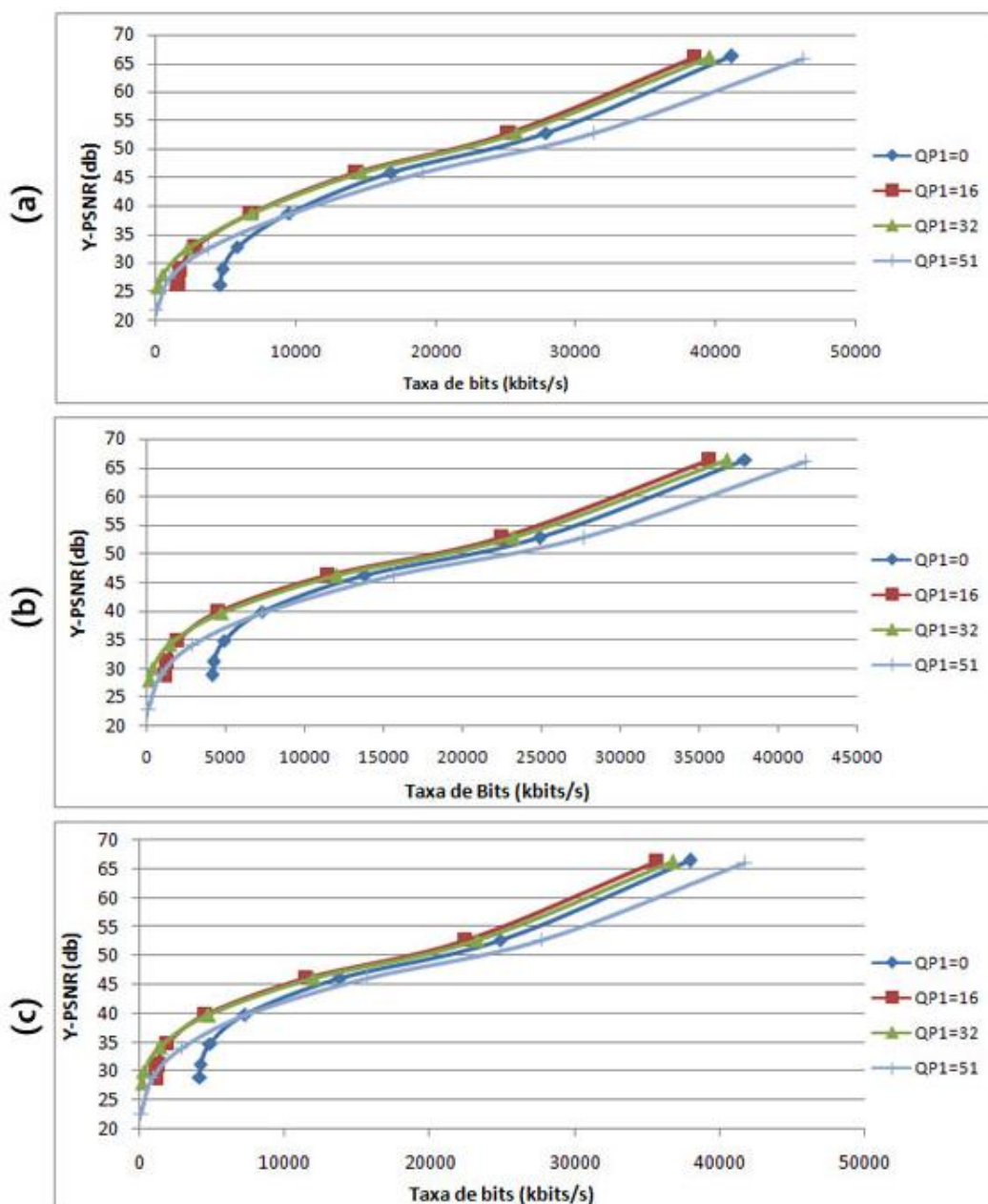


Figura 26. PSNR em função da taxa de bits.

A Figura 27 destaca a inversão de valores das sequências codificadas com camada base igual 51 de quantização. Pois quando seu valor de PSNR é menor exige menos espaço de armazenamento em relação às demais. Contudo no decorrer do gráfico passa a exigir maior espaço de armazenamento, isso mostra que quando camada base é codificada com maior qualidade, menor é a taxa de bits resultado do vídeo.

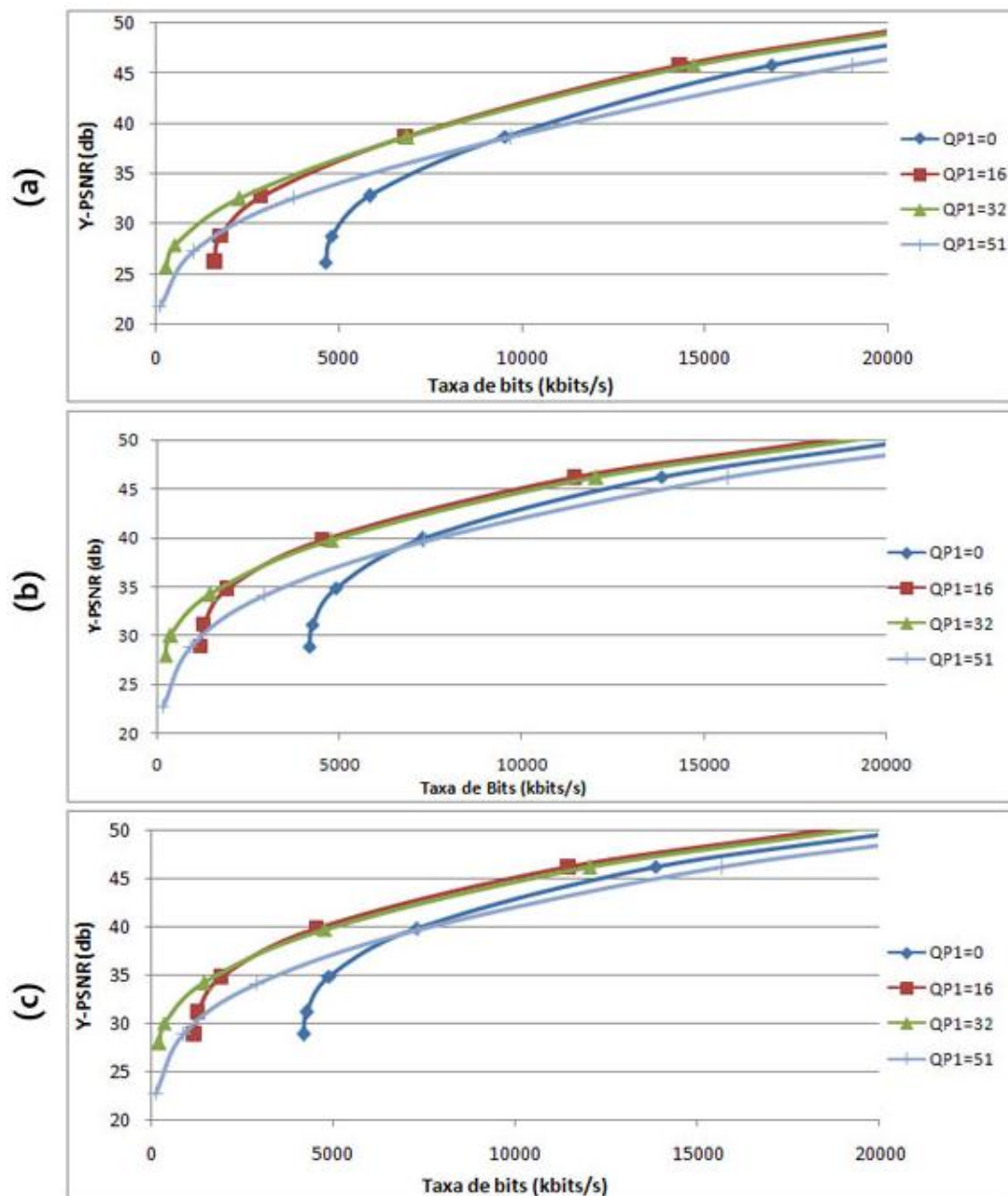


Figura 27. Interseção das curvas de PSNR.

7.5 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DE PSNR

Na Figura 28 pode se verificar a dependência da camada base com relação ao tempo total de codificação do vídeo, onde “QP1=0” necessita sempre de maior tempo para codificação das camadas. Deve se ressaltar que o tempo de codificação tratado na Figura 28 é tempo total de todas as camadas codificadas.

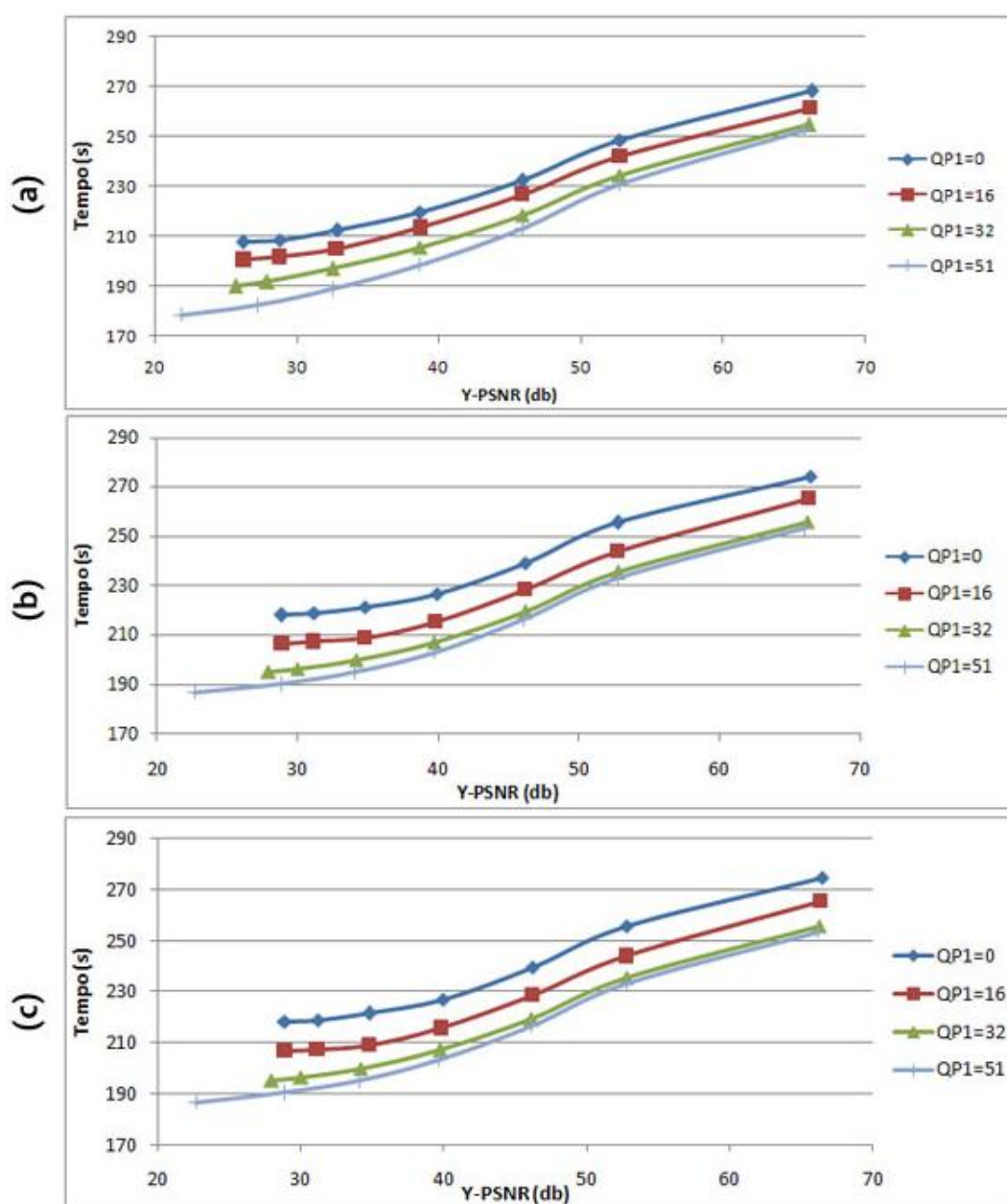


Figura 28. Tempo de codificação em função da PSNR.

7.6 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA TAXA DE BITS

Os resultados de tempo de codificação versus a taxa de bits apresentados na Figura 29 demonstram a relação direta com a camada base codificada. A série “QP1=32” obteve a melhor relação destes dois indicadores, tendo o tempo de codificação próximo à série “QP1=51” e a taxa de *bits* próxima a “QP1=16”.

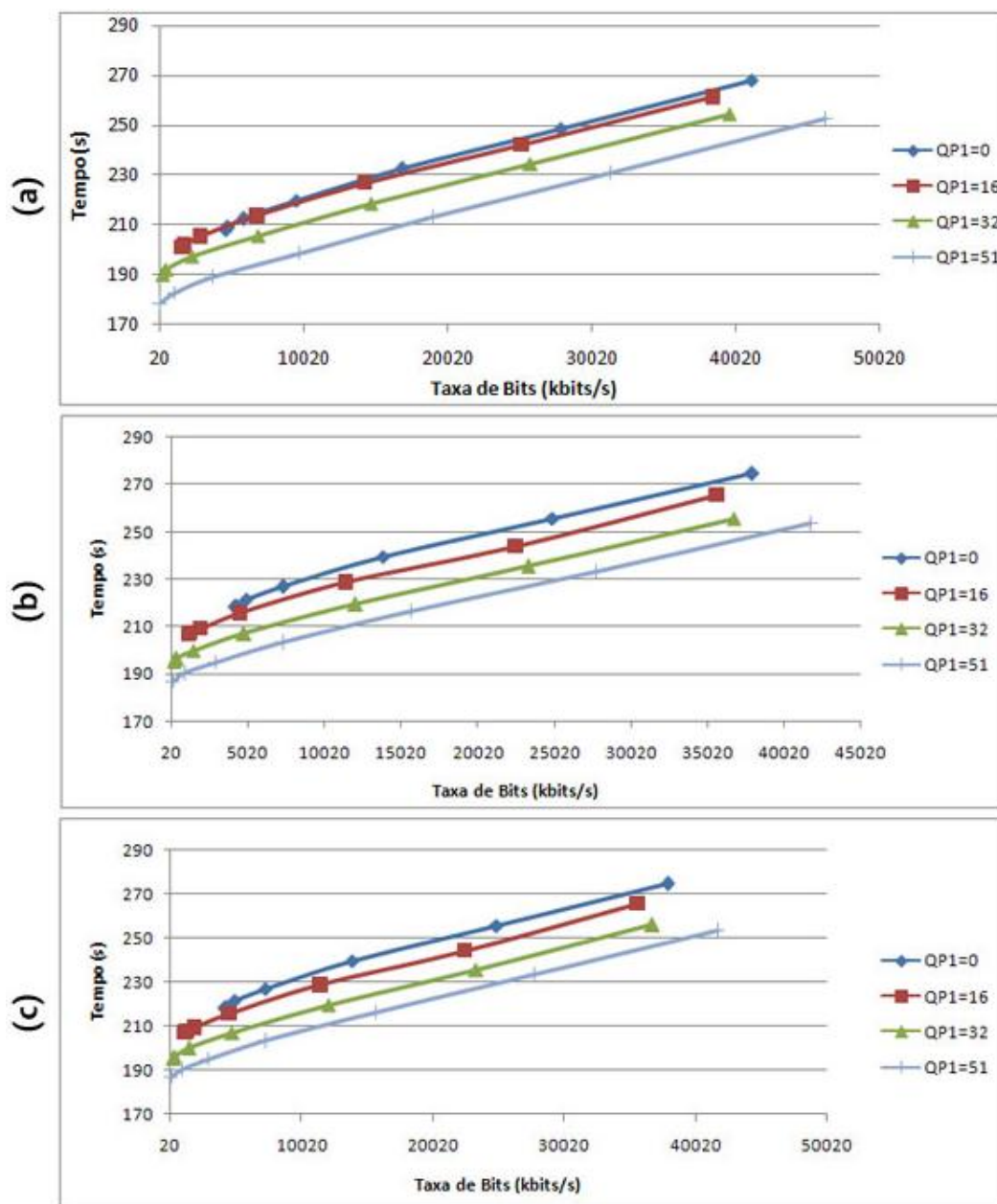


Figura 29. Tempo de codificação em função da taxa de bits.

7.7 ANÁLISE DAS SEQUÊNCIAS RESULTANTES

Após a demonstração e análise dos indicativos métricos das simulações, neste capítulo é tratada a percepção das mudanças nos vídeos testados comparando-se o frame original com o simulado. Com isso pode ser ter a exata noção do limite aceitável da condição do vídeo, levando se em consideração a aplicação do mesmo no dia-a-dia.

7.7.1 Camada Base dos Vídeos Simulados

A Figura 30 traz um comparativo de quadros da camada base dos vídeos simulados. Levando em consideração que a camada base foi codificada numa resolução QCIF de 172x144, pode se dizer que apenas a condição do vídeo simulado com quantização igual 51 é tida como ruim.



Figura 30. Comparativo de frames da camada base.

7.7.2 Camada de Incremento dos Vídeos Simulados

Nas Figuras 31, 32 e 33 estão dispostos os quadros originais além das simulações consideradas mais interessantes com relação à PNSR, taxa de *bits* e tempo de codificação.

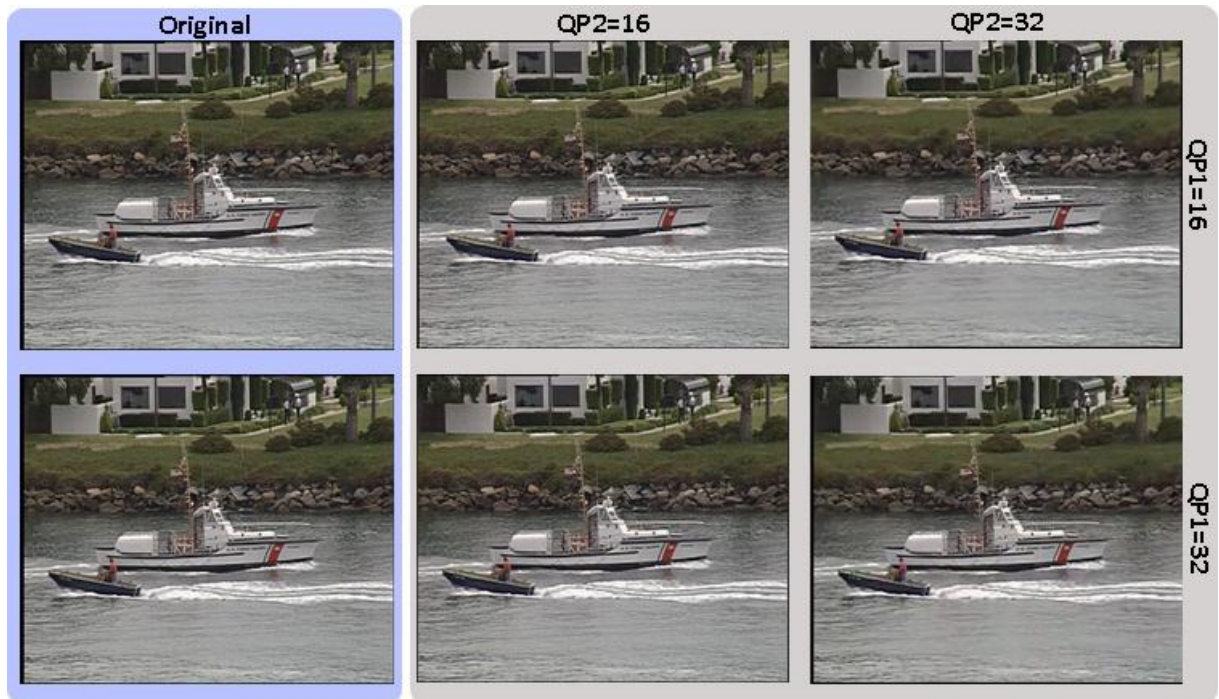


Figura 31. Camada de Incremento - Coastguard.

Pode se verificar a semelhanças dos quadros comparados aos vídeos originais, comprovando a condição de melhor sintonia dos vídeos, tanto de qualidade de imagem quanto em relação a taxa de *bits* e tempo de codificação. Porém em uma análise mais rigorosa, a condição simulada com “QP1=32” e “QP2=16” apresentou maior qualidade, com isso subentende-se que a medida que se aumenta resolução do vídeo é necessário elevar a quantização.



Figura 32. Camada de Incremento - Foreman.

A partir dos quadros demonstrados pode-se afirmar que a codificação da camada base entre 16 e 32 traz ótimas relações entre complexidade computacional e qualidade de vídeos levando-se em consideração os resultados apresentados anteriormente.



Figura 33. Camada de Incremento - Walk.

CONCLUSÃO

Neste trabalho, foi realizada uma análise qualitativa e também de quantitativa por meio de métricas utilizada para avaliação de vídeos digitais no padrão H.264/SVC. A pesquisa traz conceitos básicos sobre vídeos digitais, além de uma introdução a padrão H.264/AVC e posteriormente um estudo de sua extensão SVC.

Para realização das simulações, foi utilizado o software de referência JSVM, que também teve seus e aplicativos e características descritas nesta pesquisa. Este software permitiu a realização das simulações escalares, feitas com três sequências nas resoluções QCIF (176x144 *pixels*) e CIF (352x288 *pixels*).

Os testes realizados visavam buscar a melhor condição de quantização para as layers de uma sequência com escalabilidade espacial. Foram realizadas 147 simulações e por meio delas foram feitas as análises sobre o efeito da camada base sobre a camada de enriquecimento.

A variação da quantização demonstrou que a camada base pouco influência na qualidade da camada de incremento de uma sequência com escalabilidade espacial. Contudo outras características como, taxa de *bits* e tempo de codificação são diretamente afetadas pela primeira camada. Conseqüentemente, as sequências codificadas no H.264/SVC com escalabilidade espacial, se a sintonia de quantização for adequada pode se explorar da melhor maneira esta característica.

Analisando os resultados nas resoluções testadas, é possível afirmar que a melhor condição de codificação está quando codificada a camada base com quantização entre 16 e 32, ou seja, os extremos dos valores de quantização 0 e 51 acarretam piores condições de taxa de *bits* e tempo de codificação. Pode se dizer também que se as simulações fossem realizadas com sequências de resolução espacial maior, os ganhos seriam ainda mais acentuados por se

trabalhar com um volume de dados maiores do que os testados nesta pesquisa. As condições de quantização da camada base entre 16 e 32 foram consideradas de boa qualidade, avaliando se por meio da métrica subjetiva.

Como trabalho futuro, sugere-se a realização de simulações com a escalabilidade de temporal e de qualidade, utilizando este mesmo método de análise. Com isso se poderia sugerir a melhor condição utilizando as três variações de escala possíveis. Além disso, a avaliação do impacto das variações escalares em maiores resoluções além das resoluções testadas nesta pesquisa seria de grande valia no estudo do padrão H.264/SVC.

REFERÊNCIAS

AGOSTINI, L. V. **Desenvolvimento de arquiteturas de alto desempenho dedicadas à compressão de vídeo segundo o padrão H.264/AVC.** 2007. 172 p. Porto Alegre: Tese (Doutorado) - UFRGS, 2007.

ASCENSO, J. M. D. **Codificação escalável de Vídeo com elevada granularidade.** 2003. 297 p. Lisboa: Dissertação (Mestrado) - Universidade Técnica de Lisboa - UTL, 2003.

BRUNO, G. J. E. **Vebit: um algoritmo para codificação de vídeo com escalabilidade.** 2003. 94p. Porto Alegre: Dissertação (Mestrado) - UFRGS, 2003.

CARVALHO, M. G. **Implementação hardware/software da estimação de movimento segundo o padrão H.264.** 2007. 102p. Natal: Dissertação (Mestrado) - UFRN, 2007.

DARONCO, L. C. **Avaliação subjetiva de qualidade aplicada à codificação de vídeo escalável.** 2009. 145 p. Porto Alegre: Dissertação (Mestrado) – UFRGS, 2009.

FERREIRA, R. U. **Codificador H.264/AVC com compensação de movimento baseada em partições alternativas de macrobloco.** 2009. 83 p. Brasília: Dissertação (Mestrado) - UnB, 2009.

FONSECA, C. M. **Investigação em software sobre o padrão H.264 escalável de compressão de vídeos.** 2008. 73 p. Pelotas: UFPel, 2008.

GHANBARI, M. **Standard Codecs: Image Compression to Advanced Video Coding.** United Kingdom: The Institution of Electrical Engineers, 2003.

GONZALEZ, R.; WOODS, R. **Processamento de imagens Digitais.** 2003. São Paulo: Edgard Blücher, 2003.

HUNG, E. M. **Compensação de movimento utilizando blocos multi-escala e forma variável em um codec de vídeo híbrido.** 2007. 73 p. Brasília: Dissertação (Mestrado) – UnB, 2007.

ITU-T. Recommendation H.264. **Advanced video coding for generic audiovisual services. Int'l telecomm.** Union, Geneva, Switzerland, 2005.

MANOEL, E. T. M. **Codificação de vídeo h.264: estudo de codificação mista de macroblocos.** 2007. 117 p. Florianópolis: Dissertação (Mestrado) - Departamento de Engenharia Elétrica - UFSC, 2007.

NICOLEIT, E. R. ; SEARA, R. ; BERMUDEZ, J. C. M. . **Compressão de vídeo com alta qualidade e baixa taxa de bits através de codificação multi-estágio In: XV Simpósio Brasileiro de Telecomunicações,** 1997, Recife. XV Simpósio Brasileiro de Telecomunicações. Recife: , 1997. p.590 - 593.

NUNES, R. P. **Codificação escalável de vídeo para recepção fixa no sistema brasileiro de televisão digital.** 2009. 170 p. São Paulo: Dissertação (Mestrado) – USP, 2009.

PURI, A.; CHEN, X.; LUTHRA, A. **Video coding using the H.264/MPEG-4 AVC compression standard.** [S.l.]: Elsevier B. V., 2004. 57p.

RICHARDSON, I. E. G. **H.264 and MPEG-4 Video Compression: Video coding for next-generation multimedia,** John Wiley & Sons Publishers, USA, 2003.

SANBORN, S. **Quantifying information content in data compression using the autocorrelation function.** IEEE Signal Processing Letters. Mar, 2005, p. 230-233.

SEGALL, C.A.; SULLIVAN, G.J. **Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension.** IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, pp. 1121-1135. Sept. 2007.

SHAFER, R; SCHWARZ, H; DETLEV, M.; SCHIERL, T; WIEGAND, T. **MCTF and Scalability Extension of H.264/AVC and its Application to Video Transmission, Storage, and Surveillance.**

SCHWARZ, H.; MARPE, D.; WIEGAND, T. **Overview of the Scalable Video Coding Extension of the H.264/AVC Standard.** IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, p. 1103-1120, Sept. 2007, invited paper.

SILVA, André M. C.. **Desenvolvimento de um compensador de movimento para o padrão H.264 de compressão de vídeo.** 2006. 65 p. Pelotas: UFPel, 2006.

SOLARI, S.J. **Digital video and audio compression**. McGraw-Hill. 1997.

SOLER, Luciano. **Codificação de Vídeo Escalável**. 2006. 186 p. Porto Alegre: Dissertação (Mestrado) – UFRGS, 2006.

SONG, Daewon; CHEN, Chang W. **Scalable H.264/AVC Video Transmission Over MIMO Wireless Systems With Adaptive Channel Selection Based on Partial Channel Information**. IEEE Transactions on Circuits and Systems for Video Technology, p.1218-1226, set. 2007, invited paper.

SULLIVAN, G.; WIEGAND, T. **Video Compression – From Concepts to the H.264/AVC Standard**. Proceedings of the IEEE, v. 93, n. 1, p. 18-31, Jan. 2005.

TERAOKA, K. **Implementação de um Codificador de Video H.264 com Transformada Fliect**. 2003. Campinas: Dissertação (Mestrado) – UNICAMP, 2003.

Wikipedia. **Televisão digital no Brasil**.

Disponível em: <http://pt.wikipedia.org/wiki/Televis%C3%A3o_digital_no_Brasil>. Acesso em: 25 de Maio 2011.

WIEGAND, T.; SULLIVAN, G.; REICHEL, J.; SCHWARZ, H.; WIEN, M. . **Joint draft 11 of SVC amendment, joint video team**, Doc. JVT-X201, Jul. 2007.

WIEN, M.; SCHWARZ, H.; OELBAUM, T. **Performance Analysis of SVC**. IEEE Transactions on Circuits and Systems for Video Technology. Vol. 17, No. 9, pp.1194-1203, Sept. 2007, invited paper.

APÊNDICE A – ARTIGO CIENTÍFICO

Avaliação de Desempenho da Escalabilidade Espacial na Recomendação H.264/SVC para Codificação de Vídeo.

Márlon Henrique de Souza ¹, Evânio Ramos Niceoleit²

¹ Acadêmico do Curso Ciência da Computação - Universidade do Extremo Sul Catarinense (UNESC) - Criciúma – Brasil.

² M.Eng.. Professor do Curso Ciência da Computação - Universidade do Extremo Sul Catarinense (UNESC) - Criciúma – Brasil.

marlon.h@hotmail.com, evaniorn@gmail.com

RESUMO: O padrão H.264 é o estado da arte atualmente com relação à codificação de vídeo, conseguindo aumentar em duas vezes a taxa de compressão em relação MPEG-2. A diversidade de dispositivos que permitem à visualização de vídeos digitais, aliada a evolução dos padrões de codificação resulta em um aumento na dificuldade de execução em dispositivos de baixa resolução, sem contar no desperdício de banda para transmissão a estes dispositivos. O padrão H.264/SVC incorpora escalabilidade ao padrão AVC, permitindo a manipulação da escala do vídeo seja ela espacial, temporal ou de qualidade. Consequentemente pode se extrair sequências com escalares inferiores ao vídeo original, para facilitar utilização do padrão H.264 em dispositivos com baixa resolução. Este projeto discute o desempenho do H.264/SVC, em especial a sintonia de quantização para melhor explorar as condições de escala espacial. Os resultados comprovaram que as condições de quantização da camada base de um vídeo com escalabilidade espacial, influenciam diretamente na taxa de bits e no tempo de codificação do vídeo resultante.

ABSTRACT: The H.264/AVC standard is the current state of the art with respect to video encoding, thus increasing at twice the rate of compression ratio MPEG-2. The diversity of devices that allow viewing of digital video, coupled with changing patterns of coding results in an increase in the difficulty of running on devices with low resolution, not to mention the waste of bandwidth for transmission to these devices. The H.264/SVC standard incorporates scalability to the H.264/AVC standard, allowing the manipulation of the scale of the video whether spatial, temporal or quality. Result can be extracted sequences with scalar below the original video, to facilitate use of the H.264 standard on devices with low resolution. This project discusses the performance of H.264/SVC, especially tuning quantization to better explore the conditions of spatial scale. The results showed that the conditions of quantization of a base layer of video spatial scalability, directly influence the bit rate and encoding time of the resulting video.

1 INTRODUÇÃO

A chave para este recente sistema de TV são as técnicas de compressão. Nos últimos anos, essas técnicas passaram por uma grande evolução. Frequentemente surgem novas técnicas e padrões que melhoram a qualidade do vídeo digital para uma dada taxa de bits, ou reduzem a taxa de *bits* mantendo a qualidade.

O padrão H.264 possibilita que vídeos digitais sejam manipulados por diferentes tipos de dispositivos, tais como dispositivos móveis, aparelhos celulares, com baixas resoluções, e aparelhos de TV de alta definição. É requerida, portanto a transmissão em diferentes resoluções de vídeo (escalabilidade espacial e escalabilidade temporal) compatíveis com estes tipos de dispositivos.

Escalabilidade, de forma geral, refere-se à capacidade de extração de partes do *bitstream* de um vídeo codificado para adaptá-lo às diferentes necessidades e/ou preferências do usuário, assim também como às diferentes capacidades de receptores e/ou a condições das redes de comunicação (NUNES, 2009).

O padrão H.264/AVC possui uma extensão desenvolvida que visa o suporte a escalabilidade denominado H.264/SVC – Scalable Video Coding (WIEGAND ET AL., 2007, tradução nossa). Esta extensão possibilita a adição ou retirada de camadas dentro do fluxo de bits, manipularem-se a escala de acordo com a aplicação. O H.264/SVC provê suporte a múltiplas resoluções de exibição com um único *bitstream* (SEGALL; SULLIVAN, 2007, tradução nossa).

Esta pesquisa visa com uso da escalabilidade, avaliar a qualidade do vídeo na decodificação de vídeos no padrão H.264/SVC, em diferentes resoluções comparativamente a métricas objetivas tais como PSNR e MSE e Complexidade Computacional.

2.2 CODIFICAÇÃO E DIGITALIZAÇÃO DE VÍDEOS DIGITAIS

A busca pela compactação de um vídeo digital consiste na eliminação de dados redundantes, visando otimizar o espaço de armazenamento e a transmissão do vídeo nas diferentes redes existentes. A codificação de vídeos digitais visa manter a maior fidelidade possível com relação ao vídeo original para uma taxa de bits determinada, ou a menor taxa de *bits* se mantendo a qualidade do vídeo original (SULLIVAN; WIEGAND, 2005, tradução nossa).

A codificação de sistemas de multimídia é tratada pelos codificadores e decodificadores denominados *codec's*. Os *codec's* são algoritmos desenvolvidos para compressão de áudio e vídeo digital, alcançando reduções taxas de redução sem perdas entre 2:1 e 5:1 (NUNES, 2009).

2.4 TÉCNICAS DE COMPRESSÃO DE VÍDEO

Existem dois tipos de técnicas de compressão de vídeo, com perdas e sem perdas. As técnicas de compressão sem perdas visam reproduzir o vídeo exatamente igual o vídeo original, muito utilizado onde a riqueza de detalhes é extremamente necessária, como por exemplo, em arquivos médicos onde todo detalhe influencia no diagnóstico. Nas técnicas de codificação com perdas, o vídeo codificado perde algumas características em relação ao vídeo original, um exemplo disso é o caso de vídeos utilizados na web onde não há a necessidade de um vídeo totalmente fiel ao original para que a informação seja repassada (SILVA, 2006).

Dentro desses dois tipos de compressão existem técnicas utilizadas para atingir o objetivo de cada um, a seguir algumas técnicas serão apresentadas que contenham correlação a resolução espacial de vídeos digitais, foco desta pesquisa.

2.4.1 Predição Espacial

A predição espacial também conhecida como predição intra usa como referencia amostras do quadro atual que estão decodificadas para predizer as amostras que estão sendo codificadas (FERREIRA R., 2009).

A predição intra-quadro é realizada através da interpolação de valores dos pixels das bordas de blocos espacialmente vizinhos que foram anteriormente codificados. Esta interpolação é realizada direcionalmente, em múltiplos modos, cada um deles empregado em uma direção espacial diferente (PURI; CHEN; LUTHRA, 2004).

2.4.3 Quantização

A quantização tem por objetivo reduzir a amplitude do sinal dividindo-o e posteriormente o arredondando. Com isso pode se transmitir a informação com menos *bits*, e quando recebidos aplica-se o processo inverso multiplicando os coeficientes para reconstruir o sinal original (FERREIRA R., 2009).

Por meio do processo de quantização acontecem as maiores perdas de qualidade e também as maiores taxas de compressão. A redução causada é decorrente do processo de aproximação onde um ou mais valores são substituídos por índices de uma escala finita (BRUNO, 2003).

3 CODIFICAÇÃO ESCALÁVEL

Escalabilidade, de forma geral, refere-se à capacidade de extração de partes do *bitstream* de um vídeo codificado para adaptá-lo às diferentes necessidades e/ou preferências do usuário, assim também como às diferentes capacidades de receptores e/ou a condições das redes de comunicação (NUNES, 2009).

Codificação de vídeo escalável permite a um decodificador decodificar seletivamente apenas uma parte do fluxo de *bits*, como se pode observar na Figura 1 um decodificador pode selecionar as camadas desde a base até a ultima camada que confere a maior qualidade ao vídeo (RICHARDSON, 2003).

As camadas de enriquecimento são chamadas assim, pois cada nova camada vai adicionando novas informações para prover mais detalhes, como por exemplo, informações para uma resolução mais elevada. Como existe apenas um sinal sendo transmitido, de forma escalável, o dispositivo do usuário pode selecionar até qual camada de resolução lhe é compatível, processando apenas as informações associadas (NUNES, 2009).

4.3 PADRÃO H.264/SVC

O padrão H.264 possui uma extensão desenvolvida que visa o suporte a escalabilidade denominado H.264/SVC – *Scalable Video Coding*. Esta extensão possibilita a adição ou retirada de camadas dentro do fluxo de *bits*, manipularem-se a escala de acordo com a aplicação. O H.264/SVC provê suporte a múltiplas resoluções de exibição com um único *bitstream* (WIEGAND, 2007, tradução nossa).

A extensão SVC do padrão de codificação H.264, busca manter a alta qualidade alcançada na extensão H.264/AVC, mas subdividindo em uma ou mais camadas de *bitstreams*

intermediários que possam ser decodificados separadamente (SCHWARZ; MARPE; WIEGAND, 2007, tradução nossa).

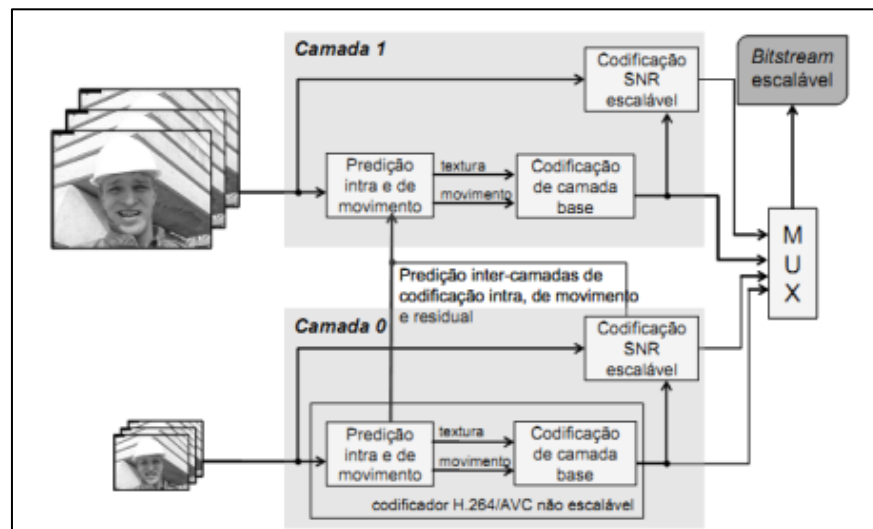


Figura 34. Estrutura do codificador SVC
Fonte: FONSECA, C. (2008, p. 27)

Conforme pode se verificar na figura 1 o padrão SVC, de maneira simples como assim foi objetivado o projeto, pode-se trabalhar os três tipos escalares: espaço, tempo e de qualidade. Tudo isso por meio de sua estruturação por camadas, a seguir será aprofundada a maneira como o H.264/SVC trata a escalabilidade espacial.

4.3.1 Escalabilidade Espacial

A escalabilidade explora a redundância espacial de um vídeo em função da resolução necessitada. A redundância espacial é também chamada de redundância intraquadro ou redundância *interpixel*, a comparação e análise de *pixels* vizinhos em um quadro que tendem a possuir valores semelhantes. Este tipo de redundância pode ser reduzida em uma operação chamada de codificação intraquadro, presente em alguns padrões de codificação de vídeo atuais (RICHARDSON, 2003).

A codificação de vídeo escalável espacial segue a abordagem tradicional de codificação em multicamadas para permitir a escalabilidade espacial. Esta técnica era empregada também nos padrões anteriores como, H.262/MPEG-2, H.263 e MPEG-4 (SCHWARZ; MARPE; WIEGAND, 2007, tradução nossa).

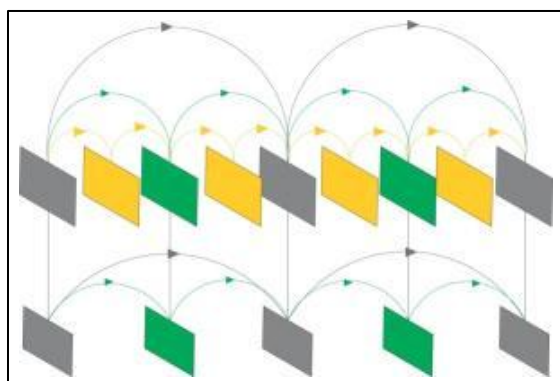


Figura 2. Comportamento da escalabilidade espacial no H.264/SVC.
Fonte: SCHWARZ, H.; WIEGAND, T. (2007, p.1109)

Nesta classificação escalar espacial do padrão H.264/SVC, utiliza mecanismos de predição entre camadas visando aperfeiçoar a codificação da camada de enriquecimento em relação ao *simulcast* (WIEN; SHWARZ; OELBAUM, 2007, tradução nossa). Este comportamento pode se verificado na Figura 2, onde a camada base é representada pelos quadros em cinza e as restantes são camadas adicionais.

6.1 H.264/SVC JSVM REFERENCE 9.19

O JSVM 9.19 (*Joint Scalable Video Model*) é o software referência do projeto SVC, este software foi desenvolvido em C++ por membros do JVT para o tratamento do padrão H.264/SVC. Este software contém uma série de executáveis que permitem trabalhar as características do padrão H.264. Dentre estes executáveis descritos três foram necessário na realização dos experimentos buscando avaliar o desempenho da extensão SVC do padrão de codificação H.264, os aplicativos utilizados nos experimentos serão descritos a seguir.

6.2 CODIFICADOR “H264AVCENCODERLIBTESTSTATIC”

Como já descrito no item 6.1, o codificador contido no pacote JSVM permite a manipulação de vídeos no padrão H.264/AVC (sem escalabilidade) e o H.264/SVC (com suporte escalar). Este aplicativo tem como saída arquivos reconstruídos no formato YUV e um arquivo compactado no formato H.264.

O parâmetro “-pf” permite a especificação do arquivo “main” do codificador, responsável pelo parâmetros macro do vídeo a ser codificado, ou seja, todos parâmetros comuns entres as camadas dos vídeos escalares. O arquivo “main” também tem por função especificar os arquivos de parâmetros das camadas escalares espaciais da sequência codificada.

O processo de codificação do codificador de referência possui opções de simulação para uma ou mais camadas, podendo-se modificar a escala espacial, temporal ou de qualidade. A seguir serão descritos os arquivos de configuração necessários para uma simulação escalar com duas camadas de codificação.

6.2.1 Configuração do Arquivo “main”

Neste arquivo estão os principais parâmetros estudados no arquivo *main* para a realização das simulações. Nele pode-se verificar a existência de parâmetros gerais do vídeo, como a codificação da sequência em formato AVC ou SVC, o arquivo de saída com todas as camadas do vídeo codificado no formato H.264, além da especificação dos arquivos de configuração das *layers*.

6.2.2 Configuração das Camadas (*Layers*)

No formato de vídeo multicamada espacial é necessária a configuração de cada *layer*, na Tabelas 1 são apresentados os parâmetros determinados para cada camada. Pode-se observar que o simulador utilizado necessita de vídeos de entrada em cada camada estipulada, no caso de duas resoluções espaciais, são necessários dois arquivos de parâmetros.

Tabela 1. Arquivo de configuração da primeira camada “layer0”.

Parâmetro	Valor	Descrição
SourceWidth	176	Largura da sequência de entrada
SourceHeight	144	Altura da sequência de entrada
InputFile	exemplo_qcif.yuv	Sequência de entrada específica da camada.
ReconFile	exemplo_qcif_layer0.yuv	Arquivo YUV reconstruído
QP	32.0	Parâmetro de quantização

6.2.3 Formato de Saída

O codificador H264AVCEncoderLibTestStatic tem como saída um arquivo no formato H.264, definido pelo arquivo de configuração *main* e um arquivo reconstruído no formato YUV para cada camada espacial codificada.

Ao fim da codificação o software fornece dados gerais da codificação como: taxa de *bits*, tempo de codificação, velocidade de codificação, PSNR e o número de frames codificados. Estes mesmos dados são fornecidos em tempo de codificação.

Todas estas informações são separadas de acordo com as camadas estipuladas nos arquivos de configuração, sejam elas em escala temporal, espacial ou de qualidade, conforme demonstrado na Figura 3.

SUMMARY :						
	bitrate	Min-bitr	Y-PSNR	U-PSNR	U-PSNR	
176x144 @ 30.0000	79.7912	79.7912	27.3628	38.7202	39.9351	
352x288 @ 30.0000	108.3504	108.3504	24.0902	38.4304	39.5772	
Encoding speed: 631.013 ms/frame, Time:189304.000 ms, Frames: 300						

Figura 3. Resultados apresentados ao final da codificação.
Fonte: Software JSVM 9.19 ao fim da codificação.

6.5 PROCEDIMENTO EXPERIMENTAL

Nesta pesquisa foram realizados testes com sequências nas resoluções CIF e QCIF, estas resoluções são utilizadas com frequência em estudos por via de softwares de simulação, além disso, são resoluções típicas de dispositivos móveis.

Os testes realizados consistem na busca por meio da variação de quantização, verificar as características do mesmo. A quantização de um vídeo varia de 0 a 51, com isso foram estipulados intervalos para que fosse possível mapear o comportamento do padrão H.264/SVC. Vale ressaltar que o parâmetro de quantização com valor 51 é tido como mais grosseiro, progredindo até valor 0 onde se tem vídeos com melhor qualidade.

Os valores de quantização estipulados para as simulações foram 0, 16, 32 e 51. Sendo que nos testes com a variação de escala espacial pode se especificar o valor de quantização para cada camada, por meio de seu arquivo de configuração específico.

Para analisar os resultados dos testes descritos, foram colhidas as informações de PSNR de Y, taxa de *bits* e tempo de codificação. Todos os valores estão condicionados a variação da quantização de codificação.

7 RESULTADOS E ANÁLISE

Neste Capítulo serão descritos os resultados e posterior análise obtidos a partir de simulações realizadas com o codificador de referência do padrão H.264/SVC. Será considerada codificação escalável espacialmente. Os resultados estão dispostos em gráficos visando facilitar a análise e compreensão, onde as grandezas apresentadas são de PSNR, tempo de codificação e taxa de *bits*, em função da variação da quantização dos vídeos. Nos gráficos contendo o desempenho de PSNR, foi adotado como padrão o valor de luminância Y, pois os demais valores UV dependem deste componente de brilho.

Todas as simulações utilizaram um vídeo na resolução QCIF juntamente com um na resolução CIF, onde o a primeira camada QCIF corresponde a “QP1” na análise dos resultados e a camada de incremento CIF a “QP2”.

7.1 TAXA DE BITS EM FUNÇÃO DA QUANTIZAÇÃO

Neste subitem serão apresentados os resultados de taxa de bits em função da quantização. Na Figura 4, pode se verificar nos testes com camada base codificada com a quantização entre 16 e 32 a taxa de bits se mantém sempre reduzida. Já nos valores extremos de QP, 0 e 51, a taxa de *bits* em determinados pontos se apresenta maior do que demais condições. Como isso pode se perceber a existência de uma zona de equilíbrio no parâmetro de quantização da camada base, para que a camada de incremento tenha uma taxa de *bits* reduzida.

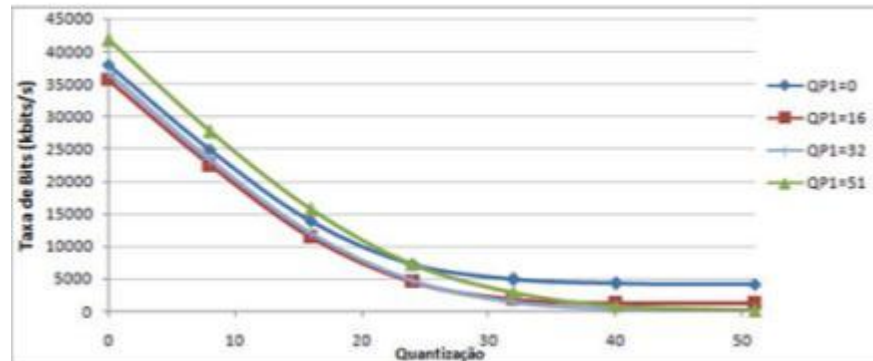


Figura 4. Taxa de bits em função da quantização.

Pode se verificar uma região onde as curvas começam tornar sua taxa de bits mais constantes. Pode se observar a evolução diferenciada da linha “QP=51”, cortando as demais em função do aumento da quantização.

7.2 PSNR EM FUNÇÃO DA QUANTIZAÇÃO

Para análise da PSNR foram coletados apenas os valores de Y, na Figura 5 são apresentados estes valores em função da variação de quantização. Pode se verificar a proximidade dos valores de PSNR nas linhas demonstradas até os pontos próximos de 32. Os valores de quantização acima de 32 passam a obter uma qualidade considerada rudimentar. Com isso quando temos a quantização entre 0 e 32, uma oportunidade de ganho na codificação é demonstrada, pois se a qualidade se mantém próxima independentemente da

camada base neste intervalor, outros indicadores como taxa de *bits* e tempo de codificação podem ser explorados.

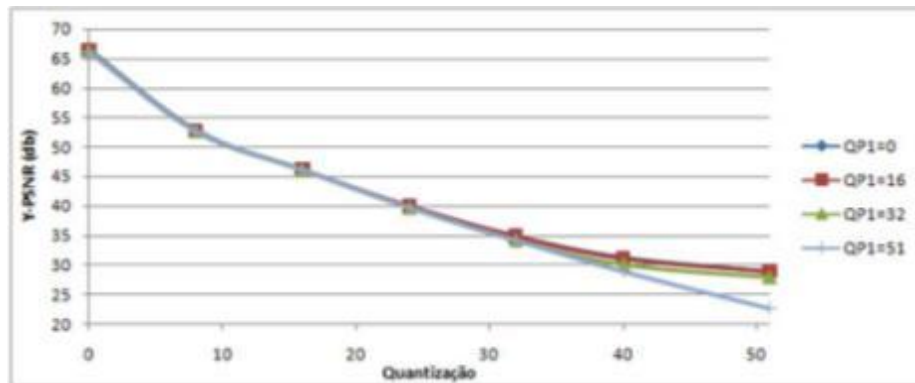


Figura 5. PSNR em função da quantização.

7.3 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA QUANTIZAÇÃO

A Figura 6 retrata se uma das oportunidades de ganho na codificação após a análise da PSNR, o tempo de codificação. No gráfico a seguir verifica se a dependência da camada base codificada para melhor desempenho de tempo de codificação. Quando codificada a camada base com quantização 0 (“QP1=0”), o tempo de codificação tende a ser maior. Porém, quando se codifica a camada base com quantização igual a 51, o tempo de codificação tende a ser menor.

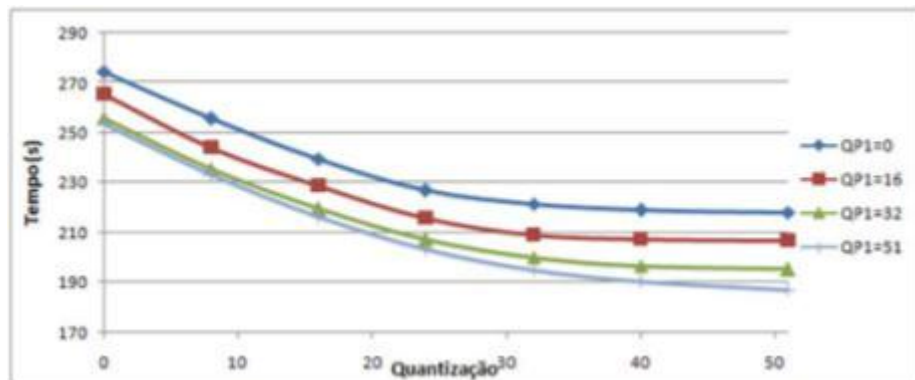


Figura 6. Tempo em função da quantização.

7.4 PSNR EM FUNÇÃO DA TAXA DE BITS

Os dados apresentados na figura 7 tratam da relação PSNR versus a taxa de *bits*. As linhas com quantização da camada base iguais a 16 e 32, apresentaram uma evolução semelhante com a taxa de *bits* sempre menor que as demais condições. Ao contrário dos vídeos codificados com camada base igual a 0 e 51, que resultaram em taxas de bits maiores em determinados pontos.

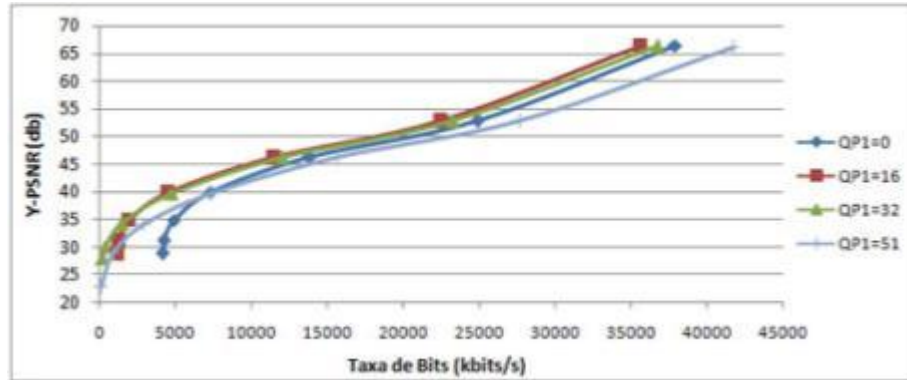


Figura7. PSNR em função da taxa de bits.

Destaca-se a inversão de valores das sequências codificadas com camada base igual 51 de quantização. Pois quando seu valor de PSNR é menor exige menos espaço de armazenamento em relação às demais. Contudo no decorrer do gráfico passa a exigir maior espaço de armazenamento, isso mostra que quando camada base é codificada com maior qualidade, menor é a taxa de bits resultado do vídeo.

7.5 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DE PSNR

Na Figura 8 pode se verificar a dependência da camada base com relação ao tempo total de codificação do vídeo, onde “QP1=0” necessita sempre de maior tempo para codificação das camadas. Deve se ressaltar que o tempo de codificação tratado na Figura 28 é tempo total de todas as camadas codificadas.

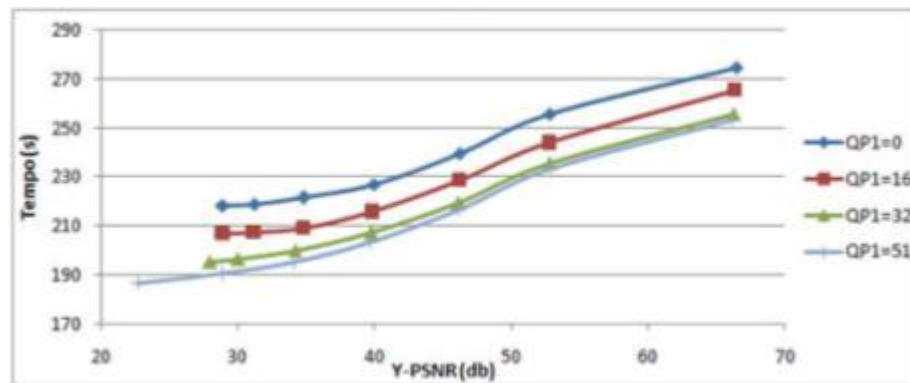


Figura 35. Tempo de codificação em função da PSNR.

7.6 TEMPO DE CODIFICAÇÃO EM FUNÇÃO DA TAXA DE BITS

Os resultados de tempo de codificação versus a taxa de bits apresentados na Figura 9 demonstram a relação direta com a camada base codificada. A série “QP1=32” obteve a melhor relação destes dois indicadores, tendo o tempo de codificação próximo à série “QP1=51” e a taxa de *bits* próxima a “QP1=16”.

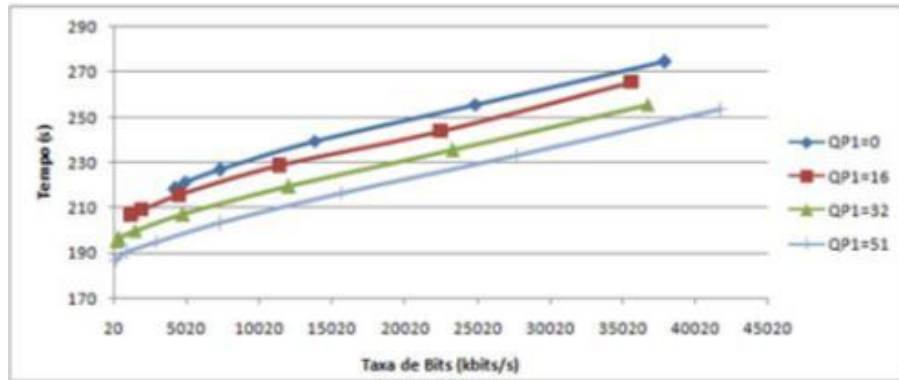


Figura 36. Tempo de codificação em função da taxa de bits.

7.7 ANÁLISE DAS SEQUÊNCIAS RESULTANTES

Após a demonstração e análise dos indicativos métricos das simulações, neste capítulo é tratada a percepção das mudanças nos vídeos testados comparando-se o frame original com o simulado. Com isso pode ser ter a exata noção do limite aceitável da condição do vídeo, levando se em consideração a aplicação do mesmo no dia-a-dia.

CONCLUSÃO

Neste trabalho, foi realizada uma análise qualitativa e também quantitativa por meio de métricas utilizada para avaliação de vídeos digitais no padrão H.264/SVC. Para realização das simulações, foi utilizado o software de referência JSVM, que também teve seus aplicativos e características descritas nesta pesquisa. Este software permitiu a realização das simulações escalares, feitas com três sequências nas resoluções QCIF (176x144 *pixels*) e CIF (352x288 *pixels*).

A variação da quantização demonstrou que a camada base pouco influência na qualidade da camada de incremento de uma sequência com escalabilidade espacial. Contudo outras características como, taxa de *bits* e tempo de codificação são diretamente afetadas pela primeira camada. Consequentemente, as sequências codificadas no H.264/SVC com escalabilidade espacial, se a sintonia de quantização for adequada pode se explorar da melhor maneira esta característica.

Analisando os resultados nas resoluções testadas, é possível afirmar que a melhor condição de codificação está quando codificada a camada base com quantização entre 16 e 32, ou seja, os extremos dos valores de quantização 0 e 51 acarretam piores condições de taxa de *bits* e tempo de codificação. Pode se dizer também que se as simulações fossem realizadas com sequências de resolução espacial maior, os ganhos seriam ainda mais acentuados por se trabalhar com um volume de dados maiores do que os testados nesta pesquisa. As condições de quantização da camada base entre 16 e 32 foram consideradas de boa qualidade, avaliando se por meio da métrica subjetiva.

REFERÊNCIAS

- BRUNO, G. J. E. **Vebit: um algoritmo para codificação de vídeo com escalabilidade**. 2003. 94p. Porto Alegre: Dissertação (Mestrado) - UFRGS, 2003.
- CARVALHO, M. G. **Implementação hardware/software da estimação de movimento segundo o padrão H.264**. 2007. 102p. Natal: Dissertação (Mestrado) - UFRN, 2007.
- FERREIRA, R. U. **Codificador H.264/AVC com compensação de movimento baseada em partições alternativas de macrobloco**. 2009. 83 p. Brasília: Dissertação (Mestrado) - UnB, 2009.
- FONSECA, C. M. **Investigação em software sobre o padrão H.264 escalável de compressão de vídeos**. 2008. 73 p. Pelotas: UFPel, 2008.
- ITU-T. Recommendation H.264. **Advanced video coding for generic audiovisual services. Int'l telecomm**. Union, Geneva, Switzerland, 2005.
- NUNES, R. P. **Codificação escalável de vídeo para recepção fixa no sistema brasileiro de televisão digital**. 2009. 170 p. São Paulo: Dissertação (Mestrado) – USP, 2009.
- PURI, A.; CHEN, X.; LUTHRA, A. **Video coding using the H.264/MPEG-4 AVC compression standard**. [S.l.]: Elsevier B. V., 2004. 57p.
- RICHARDSON, I. E. G. **H.264 and MPEG-4 Video Compression: Video coding for next-generation multimedia**, John Wiley & Sons Publishers, USA, 2003.
- SEGALL, C.A.; SULLIVAN, G.J. **Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension**. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, pp. 1121-1135. Sept. 2007.
- SCHWARZ, H.; MARPE, D.; WIEGAND, T. **Overview of the Scalable Video Coding Extension of the H.264/AVC Standard**. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, p. 1103-1120, Sept. 2007, invited paper.
- SILVA, André M. C.. **Desenvolvimento de um compensador de movimento para o padrão H.264 de compressão de vídeo**. 2006. 65 p. Pelotas: UFPel, 2006.
- SULLIVAN, G.; WIEGAND, T. **Video Compression – From Concepts to the H.264/AVC Standard**. Proceedings of the IEEE, v. 93, n. 1, p. 18-31, Jan. 2005.
- WIEGAND, T.; SULLIVAN, G.; REICHEL, J.; SCHWARZ, H.; WIEN, M. . **Joint draft 11 of SVC amendment, joint video team**, Doc. JVT-X201, Jul. 2007.
- WIEN, M.; SCHWARZ, H.; OELBAUM, T. **Performance Analysis of SVC**. IEEE Transactions on Circuits and Systems for Video Technology. Vol. 17, No. 9, pp.1194-1203, Sept. 2007, invited paper.