

# UNIVERSIDADE FEDERAL DE PELOTAS

Instituto de Física e Matemática  
Departamento de Informática



Trabalho Acadêmico

**Investigação em Software sobre o Padrão H.264  
Escalável de Compressão de Vídeos**

**Carolina Marques Fonseca**

Pelotas, 2008

**CAROLINA MARQUES FONSECA**

**Investigação em Software sobre o Padrão H.264  
Escalável de Compressão de Vídeos**

Trabalho de conclusão de curso apresentado ao curso de Bacharelado em Ciência da Computação da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Luciano Volcan Agostini

**Pelotas, 2008**

Dados de catalogação na fonte:

Ubirajara Buddin Cruz – CRB-10/901

Biblioteca de Ciência & Tecnologia - UFPel

F676i Fonseca, Carolina Marques

Investigação em software sobre padrão H.264 escalável de compressão de vídeos / Carolina Marques Fonseca ; orientador Luciano Volcan Agostini. – Pelotas, 2008. – 73f. : il. - Monografia (Conclusão de curso). Curso de Bacharelado em Ciência da Computação. Departamento de Informática. Instituto de Física e Matemática. Universidade Federal de Pelotas. Pelotas, 2008.

1.Informática. 2.Compressão de vídeo. 3.H.264/AVC. 4.H.264/SVC. 5.Escalabilidade. 6.Análise de qualidade. 7.Análise de complexidade. I.Agostini, Luciano Volcan.

II.Título.

CDD: 006.72

Dedico este trabalho a minha mãe, Anair e aos meus irmãos, Michel e Rafael, por todo amor, apoio e incentivo em todos os dias da minha vida.

## **AGRADECIMENTOS**

Em primeiro lugar quero agradecer a minha mãe Anair, pelo amor, dedicação e paciência. Sem ela nada na minha vida seria possível.

Em segundo lugar, quero agradecer meus irmãos Michel e Rafael, por sempre me incentivar e apoiar em tudo.

Agradeço ao meu namorado Peterson, por estar sempre ao meu lado me apoiando e ajudando nos momentos difíceis. No início era apenas um colega de aula, mas aos poucos foi me conquistando e hoje é uma das pessoas mais importantes da minha vida.

Agradeço ao professor Luciano Agostini, meu orientador, pela sua dedicação, compreensão e conhecimento transmitido.

Agradeço também, aos meus colegas: Bianca, César, Christian, Cristian Douglas, Eduardo, Fabiane, Fabio, Gabriel, João, Marcelo, Paulo, Peterson, Piero, Roberto e Vitor, que sempre fizeram as aulas parecerem menos cansativas e que de alguma forma me incentivaram, me ajudaram e tornaram os dias mais alegres nesses cinco anos do curso.

Agradeço aos colegas do grupo de pesquisa GACI: Fabiane, Guilherme, Felipe, Daniel e Robson, em especial aos colegas de TCC, Fabiane e Guilherme que sempre estiveram dispostos a me ajudar. Gostaria de agradecer também a Thaísa, mestrande da UFRGS, por ter me ajudado, compartilhando seu conhecimento.

Por fim, agradeço a todos os professores do curso de Bacharelado em Ciência da Computação da UFPel, que contribuíram muito para a minha formação.

A todos, muito obrigado!

## Resumo

FONSECA, Carolina Marques Fonseca. **Investigação em Software sobre o Padrão H.264 Escalável de Compressão de Vídeos**. 2008. 73f. Monografia – Curso de Bacharelado em Ciência da Computação. Universidade Federal de Pelotas.

Com o crescimento do número de dispositivos que manipulam vídeos digitais, o uso da compressão de vídeos se torna fundamental. O H.264/AVC é o padrão de compressão de vídeos mais atual, conseguindo alcançar uma taxa de compressão duas vezes maior que o MPEG-2. Como a grande maioria dos dispositivos, de celulares a TV digital de alta definição (HDTV), apresentam configurações diferentes, o vídeo deve ser codificado com diferentes configurações. Este processo é computacionalmente caro, além de gerar um grande desperdício do canal de transmissão, já que ocorre o envio de detalhes de alta definição que são eliminados nos dispositivos de baixa resolução. Com isto, o JVT (*Joint Video Team*) desenvolveu o H.264/SVC, uma extensão escalável do padrão H.264/AVC. O conceito de escalabilidade baseia-se na codificação de camadas e o SVC incorpora três tipos de escalabilidade: espacial, temporal e de qualidade. Este trabalho apresenta uma investigação em software sobre o padrão de compressão de vídeo escalável H.264/SVC, com o objetivo de analisar a qualidade, a complexidade e a eficiência de codificação deste padrão em relação ao padrão não escalável. Esta investigação foi realizada utilizando cinco vídeos com duas resoluções, uma 176x144 *pixels* e outra 352x288 *pixels*. No total foram realizadas 350 execuções neste trabalho. Todos os resultados obtidos com este trabalho são apresentados neste texto.

Palavras-chave: H.264/AVC. H.264/SVC. Escalabilidade. Análise de Qualidade. Análise de Complexidade.

## Abstract

FONSECA, Carolina Marques Fonseca. **Software Investigation about the H.264 Scalable Video Compression Standard**. 2008. 73f. Monografia – Curso de Bacharelado em Ciência da Computação. Universidade Federal de Pelotas.

With the growth number of devices that handle digital videos, the use of video compression is essential. The H.264/AVC standard is the current standard for video compression and it can achieve a compression rate twice higher than MPEG-2. As the majority of devices, from mobile phones to digital high-definition TV (HDTV), have different configurations, the video should be encoded with different settings. This process is computationally expensive and it generates a huge waste of the transmission channel, since high-definition details are sent and they are discarded in the low resolution devices. Considering this situation, the JVT (Joint Video Team) developed the H.264/SVC, a scalable extension of the H.264/AVC video coding standard. The concept of scalability is based on encoding layers. The SVC includes three types of scalability: spatial, temporal and quality. This work presents a software investigation about the scalable video coding standard, which has as objective the evaluation of the quality, the complexity and coding efficiency of this standard in comparison with the non scalable standard. This investigation has done using five videos in two resolutions: 176x144 pixels and 352x288 pixels. In the total 350 executions were performed in this work. A summary of these results is presented in this text.

Keywords: H.264/AVC. H.264/SVC. Scalability. Quality Analysis. Complexity Analysis.

## Lista de Figuras

Figura 1 - Diagrama de blocos de um codificador H.264/AVC .....	21
Figura 2 - Diagrama de blocos de um decodificador H.264/AVC .....	22
Figura 3 - Estrutura do codificador SVC com duas camadas espaciais.....	26
Figura 4 - Codificação com quadros tipo B Hierárquica .....	29
Figura 5 - Primeiro quadro das cinco amostras de vídeos utilizados nos testes. ....	35



## Lista de Tabelas

Tabela 1 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade. ....	47
Tabela 2 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade.....	47
Tabela 3 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade. ....	48
Tabela 4 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com as três escalabilidades em conjunto e sem escalabilidade.....	49
Tabela 5 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade. ....	50
Tabela 6 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade. ....	50
Tabela 7 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade.....	51
Tabela 8 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com as três escalabilidades e sem escalabilidade.....	51
Tabela 9 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade.....	52
Tabela 10 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade.....	53
Tabela 11 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade. ....	53
Tabela 12 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com as três escalabilidades juntas e sem escalabilidade.....	54

## Lista de Abreviaturas e Siglas

AVC	<i>Advanced Video Coding</i>
Cb	<i>Chrominance blue</i>
CIF	<i>Common Intermediate Format</i>
Cr	<i>Chrominance red</i>
DVD	<i>Digital Versatile Disk</i>
GACI	Grupo de Arquiteturas e Circuitos Integrados
GOP	<i>Group of Pictures</i>
HDTV	<i>High Definition Digital Television</i>
HSI	<i>Hue, Saturation, Intensity</i>
Hz	<i>Hertz</i>
IEC	<i>International Electrotechnical Commission</i>
ISO	<i>International Organization for Standardization</i>
ITU-T	<i>International Telecommunication Union - Telecommunication</i>
JSVM	<i>Joint Scalable Video Model</i>
JVT	<i>Join Video Team</i>
KB	Kilo Byte
MC	<i>Motion Compensation</i>

MCTF	<i>Motion-Compensated Temporal Filtering</i>
ME	<i>Motion Estimation</i>
MPEG	<i>Moving Picture Experts Group</i>
PSNR	<i>Peak Sinal-to-Noise Ratio</i>
Q	Quantização
Q <sup>-1</sup>	Quantização Inversa
QCIF	<i>Quarter Common Intermediate Format</i>
QP	<i>Quantization Parameter</i>
SBTVD	Sistema Brasileiro de Televisão Digital
SDTV	<i>Standard Definition Television</i>
SVC	<i>Scalable Video Coding</i>
T	Transformada
T <sup>-1</sup>	Transformada Inversa
TV	Televisão
UFPeI	Universidade Federal de Pelotas
VCEG	<i>Video Coding Experts Group</i>
VLC	<i>Variable Length Coding</i>
Y	<i>Luminance</i>
YCbCr	<i>Luminance, Chrominance Blue, Chrominance Red</i>

## Sumário

1	INTRODUÇÃO .....	13
1.1	Motivação.....	14
1.2	Objetivos .....	15
1.3	Organização do Trabalho.....	15
2	COMPRESSÃO DE VÍDEO .....	16
2.1	Espaço de Cores e Subamostragem de Cores.....	16
2.2	Redundâncias de Dados na Representação de Vídeos.....	18
2.3	Histórico dos Padrões anteriores .....	19
2.4	Padrão H.264/AVC.....	20
3	ESCALABILIDADE.....	24
3.1	Introdução à escalabilidade.....	24
3.2	Tipos de Escalabilidade .....	25
3.3	Uso da escalabilidade nos padrões anteriores.....	25
3.4	Extensão escalável do Padrão H.264/AVC .....	26
3.4.1	Escalabilidade Espacial .....	26
3.4.1.1	Predição de Movimento .....	27
3.4.1.2	Predição Residual entre camadas .....	27
3.4.1.3	Predição Intra entre camadas .....	28
3.4.2	Escalabilidade Temporal.....	28
3.4.2.1	Estruturas de Predição Hierárquica .....	29
3.4.3	Escalabilidade de Qualidade.....	30
4	SOFTWARE DE REFERÊNCIA E DESCRIÇÃO DOS EXPERIMENTOS REALIZADOS .....	31
4.1	Software de Referência.....	31
4.1.1	Executáveis Utilizados .....	32
4.2	Descrição dos Experimentos Realizados.....	35
4.2.1	Avaliações com o uso da escalabilidade espacial.....	36
4.2.2	Avaliações com o uso da escalabilidade temporal.....	37
4.2.3	Avaliações com o uso da escalabilidade de qualidade .....	38

4.2.4	Avaliações com o uso dos três tipos de escalabilidade .....	39
4.2.5	Avaliações sem o uso de escalabilidade (AVC) .....	41
4.2.6	Experimentos para a Análise da Qualidade .....	43
4.2.7	Experimentos para a Análise da Eficiência de Codificação .....	44
4.2.8	Experimentos para a Análise da Complexidade.....	44
5	RESULTADOS.....	46
5.1	Resultados de Tamanho do vídeo .....	46
5.2	Resultados de PSNR .....	49
5.3	Resultados de Tempo de Execução .....	52
6	CONCLUSÕES.....	55
	Referências.....	58
	Apêndices.....	61

# 1 INTRODUÇÃO

A compressão de vídeos está se tornando cada vez mais importante para aplicações que manipulam vídeos digitais. Em função da enorme quantidade de dados a serem armazenados ou transmitidos neste tipo de aplicação, a compressão é essencial, pois é capaz de reduzir drasticamente a quantidade de dados necessários para a representação da informação. O H.264/AVC é o mais novo padrão de compressão de vídeo e conseguiu atingir as mais elevadas taxas de compressão dentre todos os padrões existentes. Para atingir essas altas taxas de compressão, foi necessário um grande aumento na complexidade computacional dos codificadores e decodificadores, dificultando, assim, a sua utilização em software para aplicações que demandam processamento em tempo real. Por essa razão, torna-se necessário o desenvolvimento de arquiteturas de hardware para resolver este problema.

Com o avanço das tecnologias de compressão de vídeos, cada vez mais dispositivos manipulam vídeos digitais, como aparelhos celulares, DVD *players*, câmeras, filmadoras digitais e televisão digital de alta resolução. Como a maioria destes dispositivos suporta diferentes configurações e capacidades, o vídeo deveria ser codificado com diferentes configurações, mas isso se torna muito caro, além de gerar um aumento no consumo de energia e de potência dissipada pelos dispositivos e também causa um desperdício do canal de envio de dados, já que ocorre o envio de detalhes de alta resolução que são desnecessários para receptores de baixa resolução. Em razão disto, o *Joint Video Team* (JVT), formado por especialistas da ITU-T *Video Coding Group* (VCEG) e da ISO/IEC *Moving Picture Experts Group* (MPEG) desenvolveu o H.264/SVC - *Scalable Video Coding*, uma extensão escalável do padrão de compressão de vídeo H.264/AVC.

O H.264/SVC provê suporte à escalabilidade, que é o suporte a múltiplos níveis de resoluções de exibição ou qualidade a partir de um único *bitstream*. A

escalabilidade se baseia em codificação de vídeo em camadas. A camada base codifica o vídeo com uma baixa representação na taxa de envio de quadros, na resolução e/ou na qualidade. As camadas acima, chamadas de camadas de enriquecimento, codificam informações adicionais, utilizando a camada base para reconstruir o vídeo com maior taxa de envio de quadros, maior resolução e/ou maior qualidade no processo de decodificação. Desta forma, os dispositivos podem selecionar até qual camada eles têm capacidade de codificar/decodificar. Dispositivos de baixa capacidade decodificam apenas os dados presentes nas camadas mais baixas, enquanto que dispositivos de alta capacidade decodificam os dados das camadas superiores. Neste trabalho, serão usadas duas camadas, uma camada base e uma camada de enriquecimento.

O SVC incorpora três tipos de escalabilidade. A Escalabilidade Espacial, que apresenta uma variação na resolução, a Escalabilidade Temporal, que está relacionada com a taxa de quadros e a Escalabilidade de Qualidade, que está relacionada com a qualidade do vídeo.

## **1.1 Motivação**

Com o crescimento de aplicações de vídeo digital, a compressão de vídeo tem se tornado cada vez mais essencial, pois os padrões de compressão de vídeo reduzem drasticamente a quantidade de dados contida nos vídeos, viabilizando o seu armazenamento e transmissão.

O padrão H.264/AVC é o mais novo padrão de compressão de vídeo e foi desenvolvido com o objetivo de dobrar a taxa de compressão em relação aos demais padrões existentes. O fato do H.264/AVC ser o padrão estado da arte em compressão de vídeos é um motivador importante para este trabalho.

Outra motivação para este trabalho é o estudo da escalabilidade, pois seria economicamente inviável a codificação de um vídeo independente para cada resolução diferente, por exemplo, já que os atuais dispositivos suportam diferentes resoluções. Assim, com o uso da escalabilidade, apenas uma codificação é realizada onde as camadas superiores complementam os dados gerados nas camadas inferiores.

Uma outra motivação relevante para este trabalho é a contribuição com o Grupo de Arquiteturas e Circuitos Integrados da UFPel (GACI) e com o Sistema Brasileiro de Televisão Digital (SBTVD).

## **1.2 Objetivos**

O objetivo principal deste trabalho é analisar a qualidade, a complexidade e a eficiência de codificação do padrão de compressão de vídeo escalável em relação ao padrão de compressão de vídeo não escalável, através de experimentos com vídeos reais.

Outros objetivos deste trabalho são: contribuir para as novas gerações do Sistema Brasileiro de televisão Digital (SBDTV) e contribuir com o desenvolvimento do Grupo de Arquiteturas e Circuitos Integrados (GACI) da UFPel.

## **1.3 Organização do Trabalho**

O trabalho de conclusão de curso se divide em 6 capítulos e está organizado como segue:

O capítulo 2 apresenta alguns conceitos básicos de compressão de vídeo, como espaço de cores, subamostragem de cores, redundância de dados, histórico dos padrões e por fim, o padrão H.264/AVC de compressão de vídeos.

No capítulo 3, estão apresentados os conceitos e os tipos de escalabilidade, o uso de escalabilidade nos padrões anteriores e a extensão escalável do padrão H.264/AVC.

O capítulo 4 descreve o software de referência JSVM, os executáveis usados para a realização das execuções e a descrição dos experimentos realizados.

No capítulo 5 são apresentados os resultados obtidos com o desenvolvimento deste trabalho. Finalmente, no capítulo 6, são apresentadas as conclusões deste trabalho de conclusão de curso.



## **2 COMPRESSÃO DE VÍDEO**

Este capítulo tem por objetivo introduzir alguns dos conceitos básicos da área de compressão de vídeos, tais como o espaço de cores, a subamostragem de cores, a redundância de dados, o histórico dos padrões de compressão e, por fim, o padrão H.264/AVC.

### **2.1 Espaço de Cores e Subamostragem de Cores**

A representação digital de um vídeo colorido está associada à interpretação das cores pelo sistema visual humano. O sistema humano de visão possui elementos sensíveis à luz chamados bastonetes e cones. Os bastonetes são sensíveis à intensidade luminosa, enquanto os cones são sensíveis às cores primárias (GONZALEZ; WOODS, 2003). O sistema visual humano é capaz de distinguir milhares de cores distintas a partir de combinações de intensidades distintas das cores primárias. Por outro lado, o sistema visual humano não consegue distinguir mais do que duas dúzias de tons de cinza, que, na verdade, indicam a intensidade luminosa da imagem (GONZALEZ; WOODS, 2003).

Existem algumas formas de se representar as cores de forma digital. Um sistema para representar cores é chamado de espaço de cores e a definição do espaço de cor a ser utilizado para representar um vídeo é essencial para a eficiência da codificação deste vídeo.

São vários os espaços de cores usados para representar imagens digitais, tais como: RGB, HSI e YCbCr (SHI; SUN, 1999). O espaço de cores RGB é um dos mais comuns e conhecidos, ele representa, em três matrizes distintas, as três cores primárias captadas pelo sistema visual humano: vermelho, verde e azul (AGOSTINI, 2007). No espaço de cores YCbCr, as três componentes utilizadas são luminância

(Y), que define a intensidade luminosa ou o brilho; cromaância azul (Cb) e cromaância vermelha (Cr) (BHASKARAN, 1997).

Os componentes R, G e B possuem um elevado grau de correlação, o que não é desejável do ponto de vista da compressão de vídeos. Por isso, a compressão é aplicada para espaços de cores do tipo luminância e cromaância, como o YCbCr (RICHARDSON, 2002).

Outra vantagem do espaço de cor YCbCr sobre o espaço RGB é que, no espaço YCbCr, a informação de cor está completamente separada da informação de brilho. Assim, estas informações podem ser tratadas de forma diferenciada pelos codificadores de imagens estáticas e de vídeos (AGOSTINI, 2007).

O sistema visual humano é mais sensível a informações de luminância do que de cromaância. Então, os padrões de compressão de imagens estáticas e vídeos podem explorar esta característica humana para aumentar a eficiência de codificação através da redução da taxa de amostragem dos componentes de cromaância em relação aos componentes de luminância (RICHARDSON, 2002). Esta operação é chamada de subamostragem de cores e é realizada sob o espaço de cores YCbCr nos padrões de compressão de vídeo atuais.

Existem diversas formas de relacionar os componentes de cromaância com o componente de luminância para realizar a subamostragem. Os formatos mais comuns são o 4:4:4, que para cada quatro amostras de luminância (Y), existem quatro amostras de cromaância azul (Cb) e quatro amostras de cromaância vermelha (Cr), o 4:2:2 onde para cada quatro amostras de Y na direção horizontal, existem apenas duas amostras de Cb e duas amostras de Cr. Já no formato 4:2:0, para cada quatro amostras de Y, existe apenas uma amostra de Cb e uma amostra de Cr. Neste caso, as amostras de cromaância possuem metade da resolução horizontal e metade da resolução vertical do que as amostras de luminância. A nomenclatura 4:2:0 é usada por motivos históricos, pois os números não representam a relação lógica entre os componentes de cor, que seria 4:1:1 (RICHARDSON, 2003).

A subamostragem de cor aumenta significativamente a eficiência da codificação, uma vez que parte da informação da imagem é simplesmente descartada, sem causar impacto visual perceptível. Considerando o formato 4:2:0 como exemplo, uma vez que cada componente de cromaância possui exatamente um quarto das amostras presentes no componente de luminância, então um vídeo

YCbCr no formato 4:2:0 irá utilizar exatamente a metade das amostras necessárias para um vídeo RGB ou YCbCr no formato 4:4:4. Isso implica em uma taxa de compressão de 50%, considerando apenas a subamostragem. Neste trabalho é usado o espaço de cores YCbCr com subamostragem de cor no formato 4:2:0.

## 2.2 Redundâncias de Dados na Representação de Vídeos

O uso da compressão de vídeos visa diminuir a quantidade de dados considerados redundantes em uma imagem ou vídeo. Considera-se redundante todo dado que se repete inúmeras vezes em uma sequência de quadros, como o fundo da cena ou objetos que não mudam de posição. Existem três tipos de redundâncias exploradas na compressão de vídeos: redundância espacial, redundância temporal e redundância entrópica. Esta classificação apresentada é baseada em (GONZALEZ; WOODS, 2003). A seguir, estão explicadas resumidamente cada uma dessas redundâncias.

- **Redundância Espacial:** A redundância espacial também chamada de “redundância intra-quadro” (GHANBARI, 2003), “redundância interpixel” ou “redundância geométrica” (GONZALEZ; WOODS, 2003) é a redundância presente na correlação existente entre os pixels espacialmente distribuídos em um quadro (GHANBARI, 2003).
- **Redundância Temporal:** A redundância temporal, também chamada de “redundância inter-quadros” (GHANBARI, 2003), é a redundância causada pela correlação entre quadros próximos em um vídeo. Muitos blocos de pixels simplesmente não mudam de valor de um quadro para outro em um vídeo, como por exemplo, em um fundo que não foi alterado de um quadro para outro. A exploração eficiente da redundância temporal conduz a elevadas taxas de compressão, o que é fundamental para o sucesso dos codificadores (AGOSTINI, 2007).
- **Redundância Entrópica:** A redundância entrópica está relacionada com a forma de representação computacional dos símbolos codificados e não se relaciona diretamente ao conteúdo da imagem. A

entropia é uma medida da quantidade média de informação transmitida por símbolo do vídeo (SHI; SUN, 1999). A quantidade de informação nova transmitida por um símbolo diminui na medida em que a probabilidade de ocorrência deste símbolo aumenta.

### 2.3 Histórico dos Padrões anteriores

Antes do padrão H.264/AVC surgir, alguns padrões já haviam sido criados e já estavam consolidados, servindo de base para o desenvolvimento do novo padrão. O primeiro padrão relevante para a construção do H.264/AVC foi o H.261 da ITU-T (ITU-T, 1990). Este padrão lançou as bases do que é usado até hoje na maioria dos padrões de compressão de vídeo (AGOSTINI, 2007).

Após o padrão H.261, surgiu o padrão MPEG-1 da ISO/IEC (ISO/IEC, 1993), seguido do padrão MPEG-2 da ISO/IEC, que também foi padronizado pela ITU-T como H.262 (ITU-T, 1994). O MPEG-2 ou H.262 se tornou um padrão popular e é muito utilizado até a atualidade em diversas aplicações. Apesar do grande sucesso do padrão MPEG-2, a evolução dos padrões de compressão de vídeo não parou.

O padrão H.263 (ITU-T, 2000) foi lançado e incorporou alguns avanços obtidos pelos padrões MPEG-1 e MPEG-2, bem como técnicas novas que vinham sendo pesquisadas intensamente tanto pela indústria quanto pela academia.

Em 2001, o grupo MPEG (*Motion Photographic Experts Group*) da ISO/IEC (ISO/IEC, 2007) finalizou o desenvolvimento do padrão MPEG-4 Parte 2 (ISO/IEC, 1999). Ainda neste ano, o MPEG construiu uma nova chamada de propostas, similar à do H.26L da ITU-T, para melhorar ainda mais a eficiência de codificação atingida pelo MPEG-4. Então, o VCEG (*Video Coding Experts Group*), da ITU-T resolveu submeter seu *draft* H.26L em resposta à chamada de propostas do MPEG e propôs a união de esforços para completar o trabalho.

Para permitir um avanço mais acelerado na construção do novo padrão e para evitar duplicação de esforços, o ITU-T e o ISO/IEC concordaram em unir esforços para desenvolverem, em conjunto, a próxima geração de padrão para codificação de vídeo e concordaram em usar o H.26L como ponto de partida. Então,

foi criado, em dezembro de 2001, o JVT (*Joint Video Team*) (ITU-T, 2007a), formado por especialistas do VCEG e do MPEG. O JVT tinha o objetivo de completar o desenvolvimento técnico do padrão até o ano de 2003. A ITU-T decidiu adotar o padrão com o nome de H.264 e a ISO/IEC decidiu adotar o padrão com o nome de MPEG-4 parte 10 – AVC (*Advanced Video Coding* - Codificação de Vídeo Avançada). O padrão H.264 teve seu *draft* final (ITU-T, 2003) aprovado em outubro de 2003 (SULLIVAN, 2005).

Em um artigo de 2004, descrevendo o padrão (SULLIVAN, 2004), o coordenador do JVT, Gary Sullivan, decidiu utilizar o nome H.264/AVC para se referir ao padrão de forma equilibrada entre os nomes dados ao padrão pela ITU-T e pela ISO. O H.264 foi retirado do nome dado ao padrão pela ITU-T, enquanto que o AVC foi extraído do nome do padrão na ISO. Este nome, com um misto dos nomes oficiais do padrão, acabou sendo usado por muitos autores desde então e, deste ponto em diante neste texto, o padrão sempre será referenciado desta forma.

Em julho de 2004, o JVT adicionou algumas novas funcionalidades ao padrão H.264/AVC através de uma extensão do padrão chamada de *Fidelity Range Extensions* (FRExt) (ITU-T, 2004; ITU-T, 2005). Depois disso, o JVT prosseguiu em duas direções: o suporte à escalabilidade (ITU-T, 2007) e a múltiplas visões (*multi-view*). A proposta do H.264/AVC escalável foi concluída e incluída no padrão H.264/AVC no final de 2007 (HUANG, 2007). A extensão é chamada H.264/SVC – *Scalable Video Coding* (SCHWARZ, 2007).

## 2.4 Padrão H.264/AVC

Como já mencionado, o padrão H.264/AVC é o mais novo padrão de codificação de vídeo e foi desenvolvido pelo JVT (*Joint Video Team*) (ITU-T, 2007a). O JVT é formado por especialistas do VCEG da ITU-T (ITU-T, 2007) e do MPEG da ISO (ISO/IEC, 2007).

Este padrão foi desenvolvido com o objetivo principal de dobrar a taxa de compressão de seu predecessor, o padrão MPEG-2. Através do uso de diversas técnicas avançadas e inovadoras de compressão, este objetivo foi conquistado, o

que ocasionou um aumento significativo na sua complexidade computacional, que é cerca de quatro vezes maior que a do MPEG-2 (RICHARDSON, 2003).

A Fig. 1 apresenta o diagrama de blocos do codificador H.264/AVC. Os principais módulos do codificador são:

- **Predição Inter-Quadros:** é formada pela estimação de movimento (ME) e pela compensação de movimento (MC) e explora a redundância temporal;
- **Predição Intra-Quadro:** explora a redundância espacial;
- **Transformadas Diretas (T) e Inveras ( $T^{-1}$ ):** preparam os dados para exploração da redundância espacial na quantização;
- **Quantização Direta (Q) e Inversa ( $Q^{-1}$ ):** exploram a redundância espacial;
- **Filtro Redutor de Efeito de Bloco;**
- **Codificação de Entropia:** que explora a redundância de codificação.

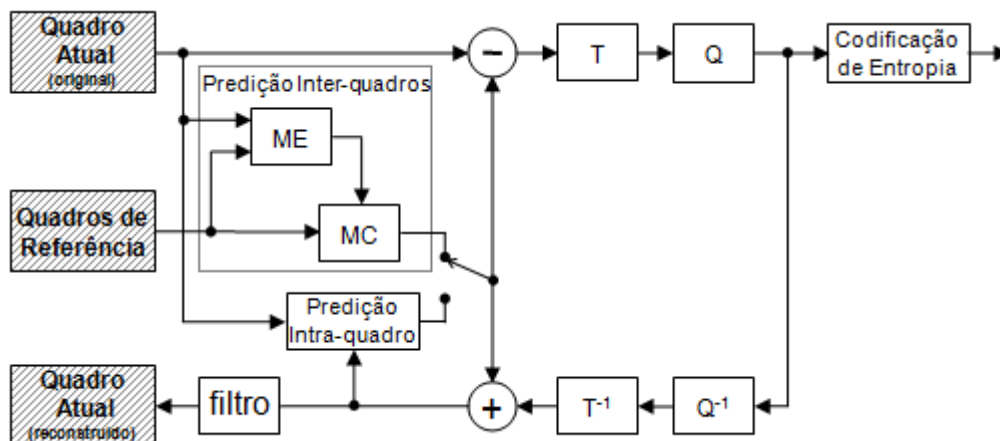


Figura 1 - Diagrama de blocos de um codificador H.264/AVC  
Fonte: AGOSTINI, 2008.

O decodificador H.264/AVC é formado por um grupo destes módulos: compensação de movimento (MC), predição intra-quadro, transformadas inveras ( $T^{-1}$ ), quantização inversa ( $Q^{-1}$ ), filtro e codificação de entropia. A Fig. 2 apresenta o diagrama de blocos de um decodificador H.264/AVC.

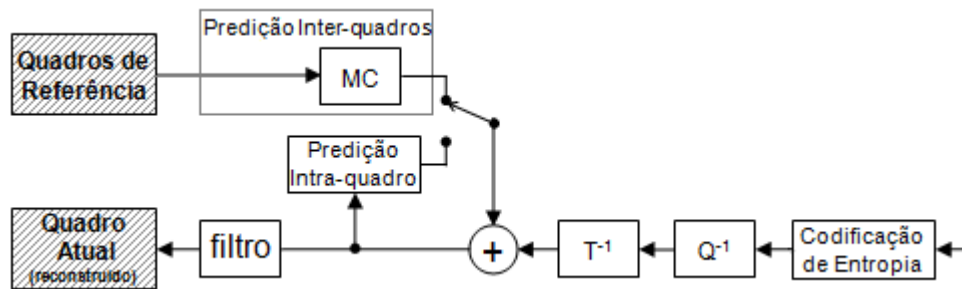


Figura 2 - Diagrama de blocos de um decodificador H.264/AVC  
Fonte: AGOSTINI, 2008.

O módulo da estimação de movimento (ME) está presente apenas nos codificadores. Este bloco é o que apresenta a maior complexidade computacional dentre todos os módulos de um codificador H.264/AVC (PURI, 2004). Este aumento da complexidade computacional é função das inovações inseridas neste módulo do padrão, que tiveram o objetivo de atingir elevadas taxas de compressão. Residem nos módulos da ME e MC (compensação de movimento) as principais fontes de ganhos do H.264/AVC em relação aos demais padrões de compressão de vídeo (WIEGAND, 2003; RICHARDSON, 2003).

A compensação de movimento funciona de forma complementar à estimação de movimento. A ME localiza o melhor casamento dentre os quadros de referência e gera um vetor de movimento. É função da compensação de movimento, a partir do vetor de movimento gerado pela ME, copiar os blocos de melhor casamento na memória de quadros anteriormente codificados e montar o quadro predito. Este quadro será subtraído do quadro atual para gerar o quadro de resíduos que passará pela transformada (AGOSTINI, 2007). A compensação de movimento é realizada tanto no codificador quanto no decodificador.

A predição intra-quadro é outra inovação introduzida pelo padrão H.264/AVC, pois é realizada no domínio espacial. Ela é responsável por reduzir a redundância presente dentro do próprio quadro (chamada de redundância espacial), em função da similaridade entre *pixels* vizinhos (GHANBARI, 2003). Vários algoritmos podem ser utilizados para realizar a codificação intra-quadro. O módulo de predição intra-quadro está presente nos codificadores e nos decodificadores H.264/AVC.

Nas transformadas direta (T) e inversa ( $T^{-1}$ ), o H.264/AVC introduziu algumas inovações. A primeira delas é que as dimensões da transformada foram inicialmente definidas em 4x4, ao invés do tradicional 8x8. Assim as transformadas passam a ter dimensões variáveis de 4x4 ou 8x8 amostras. Outra inovação é relativa ao uso de uma aproximação inteira da transformada DCT direta e inversa, de modo a facilitar a sua implementação em hardware de ponto fixo.

A quantização no padrão H.264/AVC, é uma quantização escalar (RICHARDSON, 2002). O fator de quantização é função de um parâmetro QP de entrada, que é usado no codificador para controlar a qualidade da compressão e o *bit-rate* de saída. A transformada direta e a quantização direta estão presentes apenas no codificador H.264/AVC.

O módulo Filtro nas figuras 1 e 2 representa o filtro redutor de efeito de bloco normatizado pelo padrão H.264/AVC e que está presente nos codificadores e nos decodificadores que seguem este padrão. O objetivo deste filtro é suavizar o efeito de blocos do quadro reconstruído antes do mesmo ser usado para fazer a predição de um novo macrobloco do tipo inter-quadros. Este filtro é adaptativo, distinguindo uma aresta real da imagem, que não deve ser filtrada, de um artefato gerado por um elevado passo de quantização, que deve ser filtrado (AGOSTINI, 2007).

A codificação de entropia está presente nos codificadores e decodificadores. A principal inovação introduzida na codificação de entropia do padrão H.264/AVC, tanto na codificação aritmética quanto na codificação VLC (Variable Length Coding) é o uso de codificação adaptativa baseada em contextos. Nela, a maneira com que são codificados os diversos elementos sintáticos depende do elemento a ser codificado, da fase em que se encontra o algoritmo de codificação e dos elementos sintáticos que já foram codificados.



## 3 ESCALABILIDADE

Este capítulo irá apresentar uma introdução aos conceitos envolvidos com a escalabilidade na compressão de vídeo, os tipos de escalabilidade existentes, o uso da escalabilidade nos padrões anteriores e a extensão escalável do padrão H.264/AVC, conhecida como H.264/SVC.

### 3.1 Introdução à escalabilidade

Na atualidade, cada vez mais dispositivos manipulam vídeos digitais como DVD *players*, Câmeras Digitais, Celulares e TVs de alta definição por exemplo. Mas como a maioria destes dispositivos suporta configurações diferentes (resolução, taxa de quadro e qualidade), uma maneira de lidar com esta questão seria fazer com que os dispositivos tivessem a mesma capacidade de receber e decodificar o sinal que está sendo enviado. No entanto, esta solução não é eficiente, pois causa um consumo de energia desnecessário nos dispositivos que suportam configurações mais baixas, já que a manipulação de informações de elevada resolução ou qualidade possuem um elevado custo computacional.

Outra maneira de lidar com esta questão seria transmitir diversos sinais independentes, o que é chamado de *simulcast* (SEGALL; SULLIVAN, 2007). Cada um dos sinais atenderia às necessidades específicas de um tipo de dispositivo, em função das suas configurações. No entanto, esta solução também não é eficiente, já que, neste caso, ocorreria um grande desperdício de largura banda de transmissão.

Uma alternativa eficiente para solucionar os grandes desperdícios de banda de transmissão e o elevado consumo de energia, 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 a camada ou as

camadas acima, chamadas de camadas de enriquecimento. As camadas de enriquecimento são chamadas assim, pois cada nova camada vai adicionando novas informações para alcançar, por exemplo, uma resolução mais elevada. Como existe apenas um sinal sendo transmitido, o dispositivo escolhe até qual camada lhe interessa, processando apenas as informações necessárias. Neste trabalho estão sendo consideradas duas camadas, uma camada base e uma camada de enriquecimento.

### **3.2 Tipos de Escalabilidade**

Existem três tipos de escalabilidade. A escalabilidade espacial, a escalabilidade temporal e a escalabilidade de qualidade.

Na escalabilidade espacial as camadas de enriquecimento apresentam uma variação na resolução do vídeo. Na escalabilidade temporal, as camadas de enriquecimento apresentam um aumento na taxa de quadros por segundo do vídeo, Já na escalabilidade de qualidade, se mantém a mesma resolução espacial e temporal, mas apresenta alguns incrementos na qualidade do vídeo. Cada um dos tipos de escalabilidade serão melhor apresentados nas próximas seções deste texto.

### **3.3 Uso da escalabilidade nos padrões anteriores**

Através de uma série de ferramentas, alguns modos de escalabilidade eram suportados pelos padrões anteriores (SHWARZ; MARPE; WIEGAND, 2007). Porém, o uso da escalabilidade espacial e de qualidade faziam com que a eficiência de codificação caísse muito e a complexidade do decodificador aumentasse bastante, fazendo com que o uso da escalabilidade fosse raramente utilizada.

Em dezembro de 2001, na 58ª reunião do MPEG através de um grupo *ad hoc*, começaram as atividades de padronização do SVC. As atividades deste grupo resultaram em uma chamada para apresentação de propostas para uma nova atividade de padronização em outubro de 2003, a qual resultou no *Scalable Video Model* 1.0 que resumiu os conceitos promissores submetidos em março de 2004 (WIEN; SCHWARZ; OELBAUM, 2007).

Em janeiro de 2005, esta atividade de padronização tornou-se um item de trabalho do JVT – *Joint Video Team*. (WIEN; SCHWARZ; OELBAUM, 2007). A versão final da norma escalável foi publicada em dezembro de 2007.

### 3.4 Extensão escalável do Padrão H.264/AVC

O objetivo da padronização do H.264/SVC é permitir a codificação de um *bitstream* de um vídeo de alta qualidade que contenha um ou mais subconjuntos de *bitstreams* que possam ser decodificados de forma independente com uma complexidade e qualidade similar à alcançada usando o codificador H264/AVC sem escalabilidade e com uma quantidade de dados equivalente àquela utilizada no subconjunto do *bitstream* escalável (SCHWARZ; MARPE; WIEGAND, 2007).

A estrutura típica de um codificador H.264/SVC com duas camadas está apresentada na Fig. 3. O *bitstream* da camada base é gerado de forma a ser compatível com o padrão não escalável H.264/AVC.

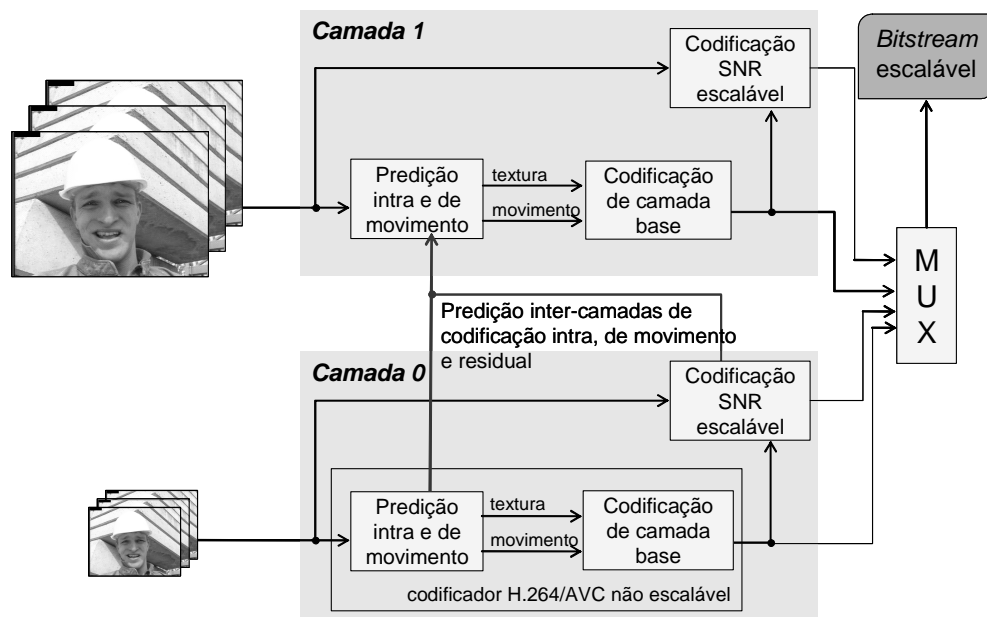


Figura 3 - Estrutura do codificador SVC com duas camadas espaciais.

#### 3.4.1 Escalabilidade Espacial

A codificação de vídeo escalável (SVC), segue a abordagem tradicional de codificação em multicamadas para permitir a escalabilidade espacial. Esta abordagem era usada também nos padrões anteriores como, H.262/MPEG-2, H.263

e MPEG-4 (SCHWARZ; MARPE; WIEGAND, 2007). Cada camada corresponde a uma resolução espacial suportada.

Para melhorar a eficiência de codificação da camada de enriquecimento em relação ao *simulcast*, foram introduzidos mecanismos de predição entre camadas. Um codificador pode escolher livremente que informação da camada de referência será explorada (WIEN; SCHWARZ; OELBAUM, 2007).

Na codificação de vídeo escalável, existem três mecanismos que podem ser utilizados na predição entre camadas. Nas seções a seguir, serão apresentados estes três mecanismos.

#### **3.4.1.1 Predição de Movimento**

Nas camadas de enriquecimento foi criado um novo tipo de macrobloco. Este macrobloco é sinalizado pelo elemento “*base mode flag*” e para ele apenas um sinal residual e nenhuma outra informação (modos de predição intra, parâmetros de movimento, etc) é transmitido.

Quando um macrobloco na camada de enriquecimento é codificado com *base mode flag* igual a 1 e o submacrobloco 8x8 correspondente na camada de referência tiver sido codificado com predição inter quadros, o macrobloco na camada de enriquecimento também é codificado com predição inter quadros. Com isto, os dados de particionamento, os índices de referência e os vetores de movimento serão derivados do bloco 8x8 co-localizado da camada de referência.

#### **3.4.1.2 Predição Residual entre camadas**

A Predição Residual entre Camadas pode ser usada para qualquer macrobloco codificado com predição Inter (se foi usado o novo tipo de macrobloco citado na seção acima ou se forem usados os tipos de macroblocos convencionais).

Devido à Predição de Movimento entre Camadas, camadas espaciais subsequentes provavelmente terão informações de movimento similares. Então, os resíduos de camadas consecutivas provavelmente terão uma forte correlação.

Para o uso deste tipo de predição, um *flag* é adicionado à sintaxe do macrobloco, o *residual prediction flag*. Quando este *flag* é igual a 1, é realizado um *upsampling* (SEGALL; SULLIVAN, 2007) no sinal residual do bloco 8x8

correspondente na camada de referência. Este *upsampling* usa um filtro que não ultrapassa as bordas do bloco, o que poderia gerar alguns componentes de perturbação visual. Para permitir a aplicação do filtro, é realizada uma extensão de borda no bloco.

O resíduo gerado na camada base é usado como predição para o sinal residual da camada de enriquecimento e, portanto, somente a diferença correspondente é codificada.

### 3.4.1.3 Predição Intra entre camadas

Quando um macrobloco na camada de enriquecimento é codificado com o *base mode flag* igual a 1 e o submacrobloco 8x8 correspondente na camada de referência foi codificado com predição intra quadro, o sinal de predição na camada de enriquecimento será obtido através do *upsampling* do sinal intra correspondente da camada de referência.

### 3.4.2 Escalabilidade Temporal

A escalabilidade temporal é a técnica que permite que um único *bitstream* suporte múltiplas taxas de quadros. Ela é, geralmente, implementada por uma estrutura de predição temporal pré-determinada pelo padrão. No SVC, podem ser utilizados vários níveis através da estrutura de quadros tipo B hierárquica (SCHWARZ; MARPE; WIEGAND, 2005, 2006) e também é utilizada a ferramenta de pré-processamento chamada MCTF – *Motion-Compensated Temporal Filtering* (compensação de movimento com filtragem temporal), para aumentar a eficiência de codificação (HUANG, et al, 2007).

Para codificadores de vídeo híbridos (chamado assim por explorar diversos tipos de redundâncias), a escalabilidade temporal pode, geralmente, ser ativada restringindo-se a estimação de movimento a quadros de referência de camadas temporais com identificadores menores ou iguais ao identificador da camada do quadro atual. Os padrões de compressão de vídeo anteriores, como MPEG-1, H.262/MPEG-2, H.263 e MPEG-4, suportavam escalabilidade temporal em algum grau. O H.264/AVC aumentou significativamente a flexibilidade para escalabilidade temporal, por causa do seu controle de memória de quadros de referência. Ele

permite a codificação de sequências de quadros com dependências temporais arbitrárias e, por isso, para suportar escalabilidade temporal com um número razoável de camadas temporais não são necessárias alterações no projeto do H.264/AVC (SCHWARZ; MARPE; WIEGAND, 2007).

### 3.4.2.1 Estruturas de Predição Hierárquica

A escalabilidade temporal com camadas de enriquecimento diádicas, as quais são caracterizadas por uma duplicação da taxa de quadros de uma camada para outra, podem ser providas de forma eficiente com o conceito de hierarquia de quadros tipo B, como ilustrado na Fig. 4. Os quadros das camadas de enriquecimento são tipicamente codificados com quadros tipo B, onde a lista 0 de quadros de referência é restrita a quadros temporalmente passados e a lista 1 de quadros de referência é restrita a quadros temporalmente futuros. Ambos os quadros, neste caso, possuem identificador de camada temporal menor do que o quadro que está sendo predito.

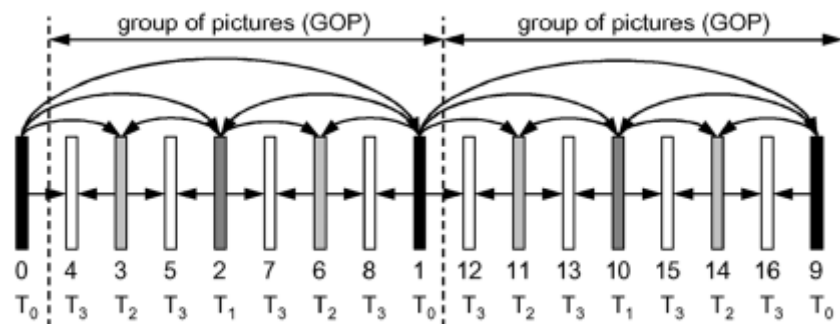


Figura 4 - Codificação com quadros tipo B Hierárquica  
Fonte: SCHWARZ; MARPE; WIEGAND, 2007

O conjunto formado por quadros entre dois quadros sucessivos na camada base junto com o quadro seguinte, forma o GOP – *group of pictures*. É possível ter uma eficiência de codificação maior através do aumento do tamanho do GOP e assim, o atraso de codificação/decodificação (SCHWARZ; MARPE; WIEGAND, 2007). Neste trabalho, o tamanho do GOP utilizado é de dois.

### 3.4.3 Escalabilidade de Qualidade

A escalabilidade de qualidade é chamada de escalabilidade de fidelidade ou escalabilidade SNR (*Signal-to-Noise Ratio*) (SCHWARZ; MARPE; WIEGAND, 2007). Este tipo de escalabilidade pode ser considerada como um caso especial da escalabilidade espacial, onde a resolução se mantém entre as camadas base e de enriquecimento. Este caso é suportado pelo conceito geral de codificação escalável espacial e também é referido como CGS (*Coarse-grain quality scalable coding*). Os mesmos mecanismos de predição entre camada usados na codificação escalável espacial são empregados, mas sem usar as operações correspondentes ao *upsampling* e *deblocking* para macroblocos codificados com predição intra na camada de referência.

Quando se utiliza a predição entre camadas para a escalabilidade CGS no SVC, um refinamento na informação de textura é normalmente realizado na camada de enriquecimento, com um parâmetro de quantização (QP) menor, comparado com o usado para a camada predita.

O conceito de multicamadas na codificação escalável de qualidade, permite apenas uma pequena taxa de *bit* selecionada ser suportada no *bitstream* escalável. Em geral, o número de taxa de *bit* é idêntico ao número de camadas.

Para aumentar a flexibilidade de adaptação do *bitstream* está incluído no SVC uma variação do CGS, o MGS (*Medium-grain quality scalability*). A diferença para o conceito CGS é um sinalizador de alto nível modificado, que permite uma troca entre diferentes camadas MGS em qualquer *bitstream*. Com o conceito de MGS, qualquer camada de enriquecimento pode ser descartada a partir do *bitstream* com escalabilidade de qualidade.

Uma outra técnica é o FGS (*Fine-grain quality scalability*) que foi padronizada pela primeira vez no MPEG-4. O FGS correlaciona diretamente sucessivos refinamentos em imagens progressivas. Esta técnica foi avaliada, mas ainda não foi incluída no padrão SVC.

## 4 SOFTWARE DE REFERÊNCIA E DESCRIÇÃO DOS EXPERIMENTOS REALIZADOS

Este trabalho de conclusão tem como principal objetivo, analisar a qualidade, a complexidade e a eficiência do padrão escalável em relação ao padrão não escalável de compressão de vídeos.

Para atingir este objetivo, foi utilizado o software de referência JSVM – *Joint Scalable Video Model* (ITU-T; ISO/IEC, 2008), o qual é descrito na seção 4.1, que também irá mostrar quais dos executáveis do software foram utilizados para a realização dos experimentos. Além disso, na seção 4.2, será descrito de forma detalhada todos os experimentos realizados para analisar a qualidade e a complexidade do padrão escalável.

### 4.1 Software de Referência

O JSVM ) (ITU-T; ISO/IEC, 2008) é um software de referência projetado pelo *Joint Video Team* (JVT), escrito em C++. Neste software estão contidas todas as funcionalidades citadas no capítulo 3. Abaixo estão listados os executáveis contidos no pacote do software JSVM. No total são 10 executáveis.

- **DownConvertStatic:** executável usado para aumentar ou diminuir a resolução de um vídeo.
- **H264AVCEncoderLibTestStatic:** executável usado para codificar o vídeo no formato escalável ou no formato sem escalabilidade.
- **H264AVCDecoderLibTestStatic:** executável usado para decodificar o vídeo.



- **AvcRewriter:** executável usado para reescrever um SVC para um *bitstream* AVC quando o *flag* que indica isto está presente na camada de enriquecimento.
- **BitStreamExtractorStatic:** executável usado para extrair *sub-bitstreams* com escalabilidade temporal, espacial e de qualidade.
- **QualityLevelAssignerStatic:** executável usado para gerar um *bitstream* com informações adicionais na camada de qualidade.
- **MCTFPreProcessor:** executável usado para uma pré filtragem de sequências de imagens.
- **PSNRStatic:** executável usado para comparar a qualidade entre dois vídeos.
- **FixedQPEncoderStatic:** executável usado para controlar o codificador e ajustar a taxa de bits do *bitstream* gerado.
- **SIPAnalyser:** executável usado para escolher a predição entre camadas.

Para a realização dos experimentos deste trabalho, apenas os executáveis descrito na próxima seção foram utilizados.

#### 4.1.1 Executáveis Utilizados

Para a realização dos experimentos utilizando o software de referência, foram usados quatro dos executáveis citados na Seção 4.1:

- H264AVCEncoderLibTestStatic
- BitStreamExtractorStatic
- H264AVCDecoderLibTestStatic
- PSNRStatic

O executável *H264AVCEncoderLibTestStatic* tem a função de codificar o vídeo, tanto no formato escalável (SVC) quanto no formato sem escalabilidade (AVC). Abaixo segue o modo da chamada deste executável pela linha de comando.

```
C:\H264AVCEncoderLibTestStatic -pf <arquivo_de_configuração>
```

Existem vários arquivos de configuração utilizados no processo. O primeiro deles é genérico e é usado para todas as camadas e cada camada possui um arquivo de configuração adicional. O arquivo de configuração genérico (ver exemplo no Apêndice A) é passado pela linha de comando e deve ser modificado para cada vídeo, devendo ter, no mínimo, as seguintes especificações:

- Largura do vídeo
- Altura do vídeo
- Formato da codificação (0:Escalável ou 1:Normal)
- Vídeo de entrada
- Nome do arquivo para o vídeo codificado
- Taxa de Frame
- Números de quadros utilizados
- Tamanho do GOP – *Group of Pictures*
- Número de camadas
- Chamada para os arquivos de configurações de cada camada (caso o formato da codificação seja escalável)

Já para os arquivos de configurações de cada camada (ver exemplo no Apêndice A), as principais especificações são as seguintes:

- Largura do vídeo
- Altura do vídeo
- Taxa de Frame de entrada
- Taxa de Frame de saída
- Vídeo de entrada
- Nome do vídeo de saída
- QP – *Quantization Parameter*

O *BitStreamExtractorStatic* é o executável responsável por extrair do arquivo gerado na etapa de codificação (formato SVC), sub-*bitstreams* com as especificações passadas no arquivo de configuração de cada camada. A chamada deste executável pela linha de comando esta definida abaixo:

```
C:\BitStreamExtractorStatic <in> <out> [-l] [-t]
```

Onde:

- **in**: arquivo gerado na etapa de codificação;
- **out**: nome do arquivo que será gerado nesta etapa de extração.
- **-l**: indica qual camada espacial está sendo extraída
- **-t**: indica qual camada temporal está sendo extraída

Já o *H264AVCDecoderLibTestStatic* é o executável responsável por decodificar o vídeo gerado após a etapa de codificação (caso o formato da codificação seja AVC) ou por decodificar os arquivos gerados após a etapa de extração de camadas (caso o formato da codificação seja SVC). A chamada do executável de decodificação pela linha de comando é mostrado a seguir:

```
C:\H264AVCDecoderLibTestStatic <arq> <name>
```

Onde:

- **arq**: arquivo gerado na etapa de extração, ou o arquivo gerado na etapa de codificação, caso o arquivo codificado seja do tipo não escalável.
- **name**: nome do arquivo que irá ser gerado após esta etapa.

Por fim, o executável *PSNRStatic* informa a qualidade entre o video original e o vídeo decodificado com o uso do software de referência. A chamada deste executável pela linha de comando é:

```
C:\PSNRStatic <l_video> <a_video> <video_orig> <video_dec>
```

Onde:

- **I\_video**: largura do vídeo
- **a\_video**: altura do vídeo
- **video\_orig**: vídeo original
- **video\_dec**: vídeo decodificado usando o software de referência

## 4.2 Descrição dos Experimentos Realizados

Os experimentos realizados foram executados no Software de Referência JSVM Versão 9.12.2, utilizando os computadores do Grupo de Arquiteturas e Circuitos Integrados (GACI) da UFPel. A configuração da maioria desses computadores é Pentium IV, 3,2GHz e 1GB de RAM ou Intel Core 2 6300, 1,86GHz e 2GB de RAM.

Para a realização dos experimentos, foram utilizados os 105 primeiros quadros de cinco vídeos. Os vídeos são bem variados, pois alguns possuem pequenas movimentações de objetos ou câmera e outros têm grande movimentação. No total foram realizadas 350 execuções utilizando os executáveis citados na seção anterior, estes executáveis são utilizados em cinco grupos que serão citados nas próximas seções. A Fig. 5 apresenta o primeiro quadro das cinco amostras de vídeos utilizadas.



Figura 5 - Primeiro quadro das cinco amostras de vídeos utilizados nos testes.

Para cada um dos cinco vídeos utilizados neste trabalho, foram realizados cinco tipos de avaliações diferentes usando o software de referência. Estas avaliações estão descritas nas seções a seguir:

#### 4.2.1 Avaliações com o uso da escalabilidade espacial

Nestas avaliações foram consideradas duas camadas de escalabilidade espacial, uma QCIF (176x174 pixels) e outra CIF (352x288 pixels). Em ambos os casos o QP foi definido 28 para o processo de quantização e a resolução temporal foi definida em 15 quadros por segundo. Para cada um dos cinco vídeos utilizados nesta avaliação, foram desenvolvidos um arquivo de configuração (ver exemplo no Apêndice A), chamado **main.cfg** e dois arquivos de configuração de camada chamados de **layer0.cfg** e **layer1.cfg**, já que neste trabalho foi definido o uso de duas camadas. Cada um desses arquivos de configuração de camadas foram criados especificamente para o uso da escalabilidade espacial (ver exemplo no Apêndice A).

Após o passo da criação dos arquivos de configuração, o executável *H264AVCEncoderLibtestStatic* é chamado na linha de comando com o parâmetro **main.cfg**. Quando esta etapa acaba, um arquivo com o vídeo comprimido é gerado (video\_comprimido.264).

Logo após, é executado o *BitStreamExtractorStatic* duas vezes, pois são geradas duas camadas espaciais: a camada base com uma resolução QCIF e a camada de enriquecimento com uma resolução CIF. Para ambas as execuções, o parâmetro **-t** foi mantido com valor **0**, indicando que foi usada a mesma resolução temporal para ambos experimentos, ou seja, com uma taxa de 15 quadros por segundo para os dois casos. Na primeira chamada, os parâmetros adicionais são o video\_comprimido.264, o nome para o primeiro *bitstream* que será extraído (video\_comprimido\_I0.264) e **-l** igual a **0** (indicando que será realizada a extração da camada base). Já na segunda chamada, passa-se o video\_comprimido.264, o nome para o segundo *bitstream* que será extraído (video\_comprimido\_I1.264) e **-l** igual a **1** (indicando que é a camada de enriquecimento que será extraída).

Com a etapa da extração dos sub-*bitstreams* encerrada, o executável *H264AVCDecoderLibTestStatic* é chamado duas vezes. A primeira vez com os parâmetros video\_comprimido\_I0.264 (camada base) e o nome do arquivo do vídeo

decodificado após esta etapa (*video\_decodificado\_l0.yuv*). A segunda vez com os parâmetros *video\_comprimido\_l1.264* (camada de enriquecimento) e o nome do vídeo decodificado nesta etapa (*video\_decodificado\_l1.yuv*).

Concluídas todas as etapas descritas acima, foi possível obter a informação de qualidade do vídeo original em relação ao vídeo reconstruído, utilizando todos os quatro executáveis citados acima.

Para obter esta informação de qualidade foi utilizado o executável *PSNRStatic* duas vezes. A primeira vez para a camada base, com os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original com resolução QCIF (*video\_original\_176x144.yuv*) e o arquivo do vídeo reconstruído (*video\_decodificado\_l0.yuv*). Na segunda chamada (camada de enriquecimento), foram passados os parâmetros largura do vídeo (352), altura do vídeo (288), o arquivo do vídeo original na resolução CIF (*video\_original\_358x288.yuv*) e o vídeo reconstruído (*video\_decodificado\_l1.yuv*).

#### 4.2.2 Avaliações com o uso da escalabilidade temporal

Nestas avaliações foi considerada uma camada de escalabilidade temporal, com GOP igual a 2, gerando assim duas camadas temporais. Uma camada temporal QCIF (176x174 pixels) com 15 quadros por segundo e uma camada temporal QCIF com 30 quadros por segundo. O QP foi definido 28 para o processo de quantização. Nesta avaliação, para cada um dos cinco vídeos utilizados, foram desenvolvidos os arquivos de configuração **main.cfg** e o arquivo de configuração de camada **layer0.cfg**. O arquivo de configuração de camada foi desenvolvido especificamente para o uso da escalabilidade temporal (ver exemplo no Apêndice A).

Quando os arquivos de configuração estão prontos, o executável *H264AVCEncoderLibtestStatic* é chamado na linha de comando com o parâmetro **main.cfg**. Após terminar a codificação do vídeo, um arquivo com o vídeo comprimido é gerado (*video\_comprimido.264*).

Então, é executado o *BitStreamExtractorStatic* duas vezes. Nas duas execuções, o parâmetro **-l** foi mantido com valor **0**, indicando que foi usada a

mesma resolução espacial para ambos experimentos, ou seja, a resolução QCIF. Na primeira vez ele tem como parâmetros `video_comprimido.264`, o nome para o primeiro bitstream que será extraído (`video_comprimido_I0t0.264`), e `-t` igual a **0** (indicando que a camada base será extraída, com a taxa de quadros igual a 15 quadros por segundo). Já na segunda chamada, passa-se o `video_comprimido.264`, o nome para o segundo *bitstream* que será extraído (`video_comprimido_I0t1.264`), e `-t` igual a **1** (indicando que a camada de enriquecimento será extraída, com a taxa de quadros igual a 30 quadros por segundo).

O próximo passo é a execução do executável *H264AVCDecoderLibTestStatic*, que será chamado duas vezes. A primeira vez com os parâmetros `video_comprimido_I0t0.264` (camada base) e o nome do arquivo do vídeo decodificado após esta etapa (`video_decodificado_I0t0.yuv`). A segunda vez com os parâmetros `video_comprimido_I0t1.264` (camada de enriquecimento) e o nome do vídeo decodificado nesta etapa (`video_decodificado_I0t1.yuv`).

Finalmente o executável *PSNRStatic* foi usado para gerar a informação de qualidade do vídeo original em relação ao vídeo reconstruído. O executável *PSNRStatic* foi chamado duas vezes. A primeira vez com os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original com resolução QCIF (`video_original_176x144.yuv`) e o arquivo do vídeo reconstruído com taxa de quadros igual a 15 quadros por segundo (`video_decodificado_I0t0.yuv`). Na segunda chamada, foi passado os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original na resolução QCIF (`video_original_176x144.yuv`) e o vídeo reconstruído com taxa de quadros igual a 30 quadros por segundo (`video_decodificado_I0t1.yuv`).

### 4.2.3 Avaliações com o uso da escalabilidade de qualidade

Nestas avaliações, foram consideradas duas camadas de escalabilidade de qualidade, uma com QP igual a 28 e outra com QP igual a 22. Em ambos os casos a resolução espacial foi QCIF (176x174 pixels) e a resolução temporal foi de 15 quadros por segundo. Para cada um dos cinco vídeos utilizados foram desenvolvidos os arquivos **main.cfg** e os dois arquivos **layer0.cfg** e **layer1.cfg** para

cada camada. Os arquivos de configuração de camadas são específicos para o uso da escalabilidade de qualidade (ver exemplo no Apêndice A).

Após o passo da criação dos arquivos de configuração, o executável *H264AVCEncoderLibtestStatic* foi chamado com o parâmetro **main.cfg**. Quando esta etapa acaba, um arquivo com o vídeo comprimido é gerado (video\_comprimido.264).

Logo após, é executado o *BitStreamExtractorStatic* duas vezes. Para ambas as execuções, o parâmetro **-t** foi mantido com valor **0**, indicando que foi usada a mesma resolução temporal para ambos experimentos, ou seja, com uma taxa de 15 quadros por segundo para os dois casos. Na primeira vez ele tem como parâmetros o video\_comprimido.264, o nome para o primeiro *bitstream* que será extraído (video\_comprimido\_I0.264) e **-I** igual a **0** (indicando que é a camada base, com QP 28). Já na segunda chamada, passa-se o video\_comprimido.264, o nome para o segundo *bitstream* que será extraído (video\_comprimido\_I1.264) e **-I** igual a **1** (indicando que é a camada de enriquecimento, com QP 22).

Com a etapa da extração dos sub-*bitstreams* encerrada, o executável *H264AVCDecoderLibTestStatic* é chamado duas vezes. A primeira vez com os parâmetros video\_comprimido\_I0.264 (camada base) e o nome do arquivo do vídeo decodificado após esta etapa (video\_decodificado\_I0.yuv). A segunda vez com os parâmetros video\_comprimido\_I1.264 (camada de enriquecimento) e o nome do vídeo decodificado nesta etapa (video\_decodificado\_I1.yuv).

Então a informação de qualidade foi gerada através do executável *PSNRStatic*, que foi chamado duas vezes. A primeira chamada foi realizada com os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original com resolução QCIF (video\_original\_176x144.yuv) e o arquivo do vídeo reconstruído com um QP igual a 28 especificado no arquivo de configuração da camada base (video\_decodificado\_I0.yuv). Na segunda chamada, foram passados os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original na resolução QCIF (video\_original\_176x144.yuv) e o vídeo reconstruído com um QP igual a 22 especificado na camada de enriquecimento (video\_decodificado\_I1.yuv).

#### **4.2.4 Avaliações com o uso dos três tipos de escalabilidade**



Nestas avaliações, foram consideradas duas camadas de escalabilidade, misturando a escalabilidade espacial, a escalabilidade temporal e a escalabilidade de qualidade. Para tanto, a camada base foi definida com uma resolução espacial QCIF, com uma resolução temporal de 15 quadros por segundo e com um QP de 28. A camada de enriquecimento foi definida com uma resolução espacial CIF, com uma resolução temporal de 30 quadros por segundo e com um QP de 22.

Para realizar estas avaliações, foi necessário desenvolver os arquivos de configuração **main.cfg** e os arquivos de configuração de camada **layer0.cfg** e **layer1.cfg**, para cada um dos cinco vídeos utilizados. Os arquivos de configuração de camadas são específicos para o uso das três escalabilidades juntas (ver exemplo no Apêndice A).

Após criados os arquivos de configuração, o executável *H264AVCEncoderLibtestStatic* foi chamado na linha de comando com o parâmetro **main.cfg**. Quando esta etapa acaba, um arquivo com o vídeo comprimido é gerado (*video\_comprimido.264*).

Após a codificação, é executado o *BitStreamExtractorStatic* duas vezes. Na primeira vez ele tem como parâmetros o *video\_comprimido.264*, o nome para o primeiro *bitstream* que será extraído (*video\_comprimido\_I0t0.264*), **-l** igual a **0** (indicando a camada base com resolução QCIF e QP igual a 28) e **-t** igual a **0** (indicando que a taxa de quadros é igual a 15 quadros por segundo). Na segunda vez passa-se o *video\_comprimido.264*, o nome para o quarto *bitstream* que será extraído (*video\_comprimido\_I1t1.264*), **-l** igual a **1** (indicando a camada de enriquecimento com resolução CIF e QP igual a 22) e **-t** igual a **1** (indicando que a taxa de quadros é igual a 30 quadros por segundo).

Com a etapa da extração dos sub-*bitstreams* encerrada, o executável *H264AVCDecoderLibTestStatic* é chamado duas vezes. A primeira vez com os parâmetros *video\_comprimido\_I0t0.264* (camada base) e o nome do arquivo do vídeo decodificado após esta etapa (*video\_decodificado\_I0t0.yuv*). A segunda chamada é realizada com os parâmetros *video\_comprimido\_I1t1.264* (camada de enriquecimento) e o nome do vídeo decodificado nesta etapa (*video\_decodificado\_I1t1.yuv*).

Por fim, os dados para as avaliações de qualidade foram gerados. Para obter esta informação foi utilizado o executável *PSNRStatic* duas vezes. A primeira vez com os parâmetros largura do vídeo (176), altura do vídeo (144), o arquivo do vídeo original (*video\_original\_176x144.yuv*) e o arquivo do vídeo reconstruído, que possui um QP igual a 28, uma taxa de quadros igual a 15 quadros por segundo e uma resolução espacial QCIF, tal qual foram especificados no arquivo de configuração da camada base (*video\_decodificado\_l0t0.yuv*). Na segunda chamada, foi passado os parâmetros largura do vídeo (352), altura do vídeo (288), o arquivo do vídeo original (*video\_original\_352x288.yuv*) e o vídeo reconstruído, que possui um QP igual a 22, uma taxa de quadros igual a 30 quadros por segundo e uma resolução espacial CIF, tal qual especificado no arquivo de configuração da camada base (*video\_decodificado\_l1t1.yuv*).

#### **4.2.5 Avaliações sem o uso de escalabilidade (AVC)**

As avaliações sem o uso de escalabilidade, ou seja, usando o codificador H.264/AVC convencional, foram divididas em quatro tipos. Estas avaliações foram realizadas para que fosse possível a geração de dados de comparação do padrão não escalável com as diferentes possibilidades de uso de escalabilidade no padrão escalável. Para cada tipo de avaliação do AVC, foram usados os cinco vídeos citados na Seção 4.2 e para cada um dos vídeos, foram desenvolvidos dois arquivos de configurações, um **main\_base.cfg** e outro **main\_enriquecimento.cfg**. Abaixo estão listados os quatro tipos de avaliações usando o AVC com suas respectivas configurações.

- **Avaliações do AVC para comparação com a escalabilidade espacial:**

No arquivo de configuração **main\_base.cfg** as configurações foram as seguintes: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual a 28. Já no arquivo de configuração **main\_enriquecimento.cfg** as configurações foram: resolução CIF, taxa de quadros igual a 15 quadros por segundo e QP igual a 28.

- **Avaliações do AVC para comparação com a escalabilidade temporal**

No arquivo de configuração **main\_base.cfg** as configurações foram: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual a 28. Já no arquivo de configuração **main\_enriquecimento.cfg** as configurações foram: resolução QCIF, taxa de quadros igual a 30 quadros por segundo e QP igual a 28.

- **Avaliações do AVC para comparação com a escalabilidade de qualidade**

No arquivo de configuração **main\_base.cfg** as configurações foram: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual 28. Já no arquivo de configuração **main\_enriquecimento.cfg** as configurações foram: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual a 22.

- **Avaliações de AVC para comparação com o uso das três escalabilidades juntas**

No arquivo de configuração **main\_base.cfg** as configurações foram: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual 28. Já no arquivo de configuração **main\_enriquecimento.cfg** as configurações foram: resolução CIF, taxa de quadros igual a 30 quadros por segundo e QP igual a 22.

Para cada uma das avaliações sobre o AVC citadas acima, os passos a seguir foram os mesmos.

Após terem sido criados os arquivos de configuração, o executável *H264AVCEncoderLibtestStatic* é chamado na linha de comando duas vezes, com o parâmetro **main\_base.cfg** na primeira vez, e o **main\_enriquecimento.cfg** na segunda vez. Quando esta etapa acaba, dois arquivos com os vídeos comprimidos são gerados (*video\_comprimido\_base.264* e *video\_comprimido\_enriquecimento.264*).

Logo após, é executado o *H264AVCDecoderLibTestStatic* duas vezes. Os parâmetros passados na primeira vez são: *video\_comprimido\_base.264* (com configuração equivalente à da camada base do codificador escalável) e o nome do arquivo do vídeo decodificado após esta etapa (*video\_decodificado\_base.yuv*). Já os parâmetros passados na segunda chamada são: *video\_comprimido\_enriquecimento.264* (com configuração equivalente à da camada de enriquecimento do codificador escalável) e o nome do arquivo do vídeo decodificado após esta etapa (*video\_decodificado\_enriquecimento.yuv*).

Para gerar os dados de qualidade, foi utilizado o executável *PSNRStatic*, que foi chamado duas vezes. Na primeira vez, este executável tem como parâmetros a largura do vídeo (176), a altura do vídeo (144), o arquivo do vídeo original com resolução QCIF (*video\_original\_176x144.yuv*) e o arquivo do vídeo reconstruído (*video\_decodificado\_base.yuv*), que corresponde à configuração utilizada na camada base do codificador escalável. Na segunda chamada, foram passados os parâmetros largura do vídeo (352), altura do vídeo (288), o arquivo do vídeo original na resolução CIF (*video\_original\_352x288.yuv*) e o vídeo reconstruído (*video\_decodificado\_enriquecimento.yuv*), com configuração equivalente àquela utilizada na camada de enriquecimento da codificação escalável.

#### 4.2.6 Experimentos para a Análise da Qualidade

Para analisar a qualidade dos vídeos do padrão escalável de compressão em relação ao padrão sem escalabilidade, o critério adotado foi o *Peak Signal-to-Noise Ratio* (PSNR) (RICHARDSON, 2003), pois é um dos parâmetros de avaliação mais utilizados na literatura para comparar a qualidade objetiva de vídeos (BHASKARAN; KONSTANTINIDES, 1997).

A qualidade pode ser diretamente ligada ao valor do PSNR obtido. Valores altos indicam alta qualidade e valores baixos de PSNR indicam baixa qualidade.

A partir das avaliações citadas nas seções anteriores, foi possível descobrir o valor em dB do PSNR. No próximo capítulo serão apresentados todos os resultados obtidos com as execuções, bem como uma discussão destes resultados.

#### 4.2.7 Experimentos para a Análise da Eficiência de Codificação

Para a análise da eficiência da codificação o critério adotado foi o do tamanho do vídeo gerado após ser codificado com o uso da escalabilidade em relação à codificação sem o uso da escalabilidade.

Com as avaliações descritas nas seções anteriores, foi possível obter os tamanhos dos vídeos codificados com e sem o uso de escalabilidade, usando o executável *H264AVCEncoderLibTestStatic*.

Os resultados desta avaliação encontram-se no próximo capítulo deste Trabalho de Conclusão de Curso.

#### 4.2.8 Experimentos para a Análise da Complexidade

A análise da complexidade foi realizada comparando o tempo de execução da codificação com e sem o uso da escalabilidade. Para obter o tempo de execução da codificação, foi adicionado ao código do software de referência uma função para calcular o tempo de execução da codificação. Para esta análise ser mais precisa, todas as execuções foram rodadas em um mesmo computador, evitando ao máximo a execução paralela de outros processos. Os resultados desta comparação são apresentados no próximo capítulo.

As modificações no software de referência estão descritas abaixo.

```
//inlui-se a biblioteca time.h
```

```
#include <time.h>
```

```
//Declara-se algumas variáveis do tipo clock_t
```

```
clock_t  antes, depois, result, total=0;
```

```
//No lugar onde se quer começar a contar o tempo coloca essas chamadas  
srand(time(NULL));
```

```
antes = clock();
```

```
//Onde se quer parar de contar coloca essas chamadas
```

```
depois = clock();
```

```
result = depois - antes;  
total = total + result;  
  
//Pra mostrar o resultado  
printf("Tempo gasto: %d ms\n\n",total);
```

## 5 RESULTADOS

Este capítulo apresenta uma síntese dos resultados obtidos através das execuções com o software de referência usando os seguintes vídeos: *City*, *Foreman*, *Harbour*, *Ice* e *Soccer*. No total foram realizadas 350 execuções neste trabalho.

Os resultados foram divididos em seções. Na Seção 5.1, estão apresentados os resultados do tamanho dos vídeos comprimidos para a análise da eficiência, na Seção 5.2 estão apresentados os resultados para o valor do PSNR para a análise da qualidade e, por fim, na Seção 5.3, estão apresentados os resultados do tempo de execução da codificação para a análise da complexidade.

### 5.1 Resultados de Tamanho do vídeo

A seguir serão apresentados os resultados dos tamanhos dos vídeos comprimidos em KB (Kilo *bytes*), após a codificação dos cinco vídeos utilizando as avaliações citadas no capítulo anterior. Os resultados a seguir apresentam o tamanho do *stream* de vídeo gerado pela codificação com duas camadas de escalabilidade, conforme já discutido no capítulo 4 deste texto. Por outro lado, para a codificação sem escalabilidade, foram somados os tamanhos dos arquivos gerados por uma configuração equivalente à usada na camada base do codificador escalável e outro gerado por uma configuração equivalente à camada de enriquecimento do codificador escalável. Assim foi possível realizar uma comparação justa entre os dois processos de codificação.

A tab. 1 mostra os resultados dos tamanhos dos vídeos utilizando os resultados da codificação com escalabilidade espacial e sem escalabilidade (ou seja,

com codificação AVC convencional). Nesta tabela também são apresentados os ganhos trazidos pela versão escalável.

Tabela 1 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Espacial</b>	348	363	914	331	541
<b>AVC Sem Escalabilidade</b>	311	365	1056	332	542
<b>Ganhos (em %)</b>	-10,6	0,6	15,5	0,3	0,2

Os resultados da tab. 1, mostram que o uso da escalabilidade espacial gerou um ganho em relação à codificação sem escalabilidade na maior parte dos casos. O vídeo que alcançou uma taxa de compressão maior foi o *HARBOUR*, com uma compressão de 914KB, diminuindo 142KB em relação ao AVC convencional, que teve uma compressão de 1056KB. O pior caso foi obtido para o vídeo *CITY*, onde a compressão escalável gerou um arquivo maior do que o gerado pela codificação sem escalabilidade. Este foi um resultado inesperado e demonstra que o uso da escalabilidade espacial apenas não garante ganhos importantes na taxa de compressão. Além disso, é possível perceber que os ganhos em relação à codificação não escalável são dependentes do tipo de vídeo codificado.

Para as avaliações do uso da escalabilidade temporal e sem o uso de escalabilidade, os resultados são apresentados na tab. 2. A tab. 2 também apresenta os ganhos trazidos pela versão escalável.

Tabela 2 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Temporal</b>	74	103	156	102	139
<b>AVC Sem Escalabilidade</b>	132	194	318	196	268
<b>Ganhos (em %)</b>	78,4	88,3	103,8	92,2	92,8

Os ganhos gerados com o uso da escalabilidade temporal são mais significativos do que os ganhos obtidos para a escalabilidade espacial. Os dados da



Tabela 2 mostram que, com o uso da escalabilidade temporal, existem, em todos os casos, ganhos importantes na taxa de compressão quando comparado às taxas do codificador não escalável. Em termos absolutos, o vídeo que alcançou o maior ganho de compressão foi o *HARBOUR* novamente, atingindo um tamanho de 156KB com a codificação escalável, diminuindo 162KB em relação ao AVC convencional (que atingiu 318KB).

A tab. 3 apresenta os resultados para os tamanhos dos vídeos utilizando as avaliações com escalabilidade de qualidade e sem escalabilidade. Nesta tabela, também são apresentados os ganhos gerados com o uso da escalabilidade.

Tabela 3 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Escalabilidade Qualidade</b>	265	284	493	223	349
<b>AVC Sem Escalabilidade</b>	242	303	527	274	396
<b>Ganhos (em %)</b>	-8,8	6,7	6,9	22,9	13,5

Também nesta avaliação o vídeo *CITY* apresentou um resultado inesperado, com uma taxa de compressão inferior para a codificação com escalabilidade de qualidade do que com a codificação sem escalabilidade e esta perda chegou a 23 KB. Para todos os outros vídeos os resultados obtidos com o uso da escalabilidade de qualidade apresentaram uma maior taxa de compressão do que sem o uso de escalabilidade. Desta vez, o vídeo que alcançou uma taxa de compressão maior foi o *ICE*, com 223KB em relação aos 274KB do padrão AVC, obtendo um ganho de 51KB.

A tab. 4 mostra os resultados para os tamanhos dos vídeos utilizando as codificações com o uso das três escalabilidades em conjunto e sem o uso de escalabilidade. Nesta tabela também estão apresentados os ganhos obtidos pelo padrão escalável em relação ao AVC sem escalabilidade.

Tabela 4 - Comparação dos tamanhos (em KB) dos vídeos gerados pela codificação com as três escalabilidades em conjunto e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Três Escalabilidades juntas</b>	700	708	1511	531	962
<b>AVC Sem Escalabilidade</b>	749	762	1949	548	1009
<b>Ganhos (em %)</b>	7	7,6	29	3,2	4,9

Os resultados apresentados na tab. 4 mostram que o padrão escalável obteve uma melhora significativa na taxa de compressão dos vídeos em relação ao padrão AVC quando as três escalabilidades são utilizadas em conjunto. O maior ganho foi alcançado com o vídeo *HARBOUR*, com 438KB a menos do que o padrão AVC convencional. Os ganhos apresentados são inferiores daqueles obtidos com a soma dos ganhos das três escalabilidades quando executadas individualmente, como pode ser percebido através da análise das tabelas. 1, 2 e 3. Este resultado era esperado, já que as escalabilidades exploram de forma diferente e, por vezes, mutuamente exclusivas, as mesmas regiões de um vídeo e, deste modo, os ganhos do uso de escalabilidade nestas regiões não podem ser agregados.

## 5.2 Resultados de PSNR

Os resultados de PSNR (em dB) para as avaliações descritas no capítulo 4 são apresentadas nesta seção. Os resultados de PSNR foram gerados para a camada de enriquecimento das codificações escaláveis e para a codificação sem escalabilidade gerada por uma configuração equivalente à usada na camada de enriquecimento do codificador escalável. Assim foi possível realizar uma comparação justa entre os dois processos de codificação.

A tab. 5 apresenta os resultados do PSNR para a codificação com escalabilidade espacial e sem escalabilidade. Também estão apresentados os ganhos obtidos com o uso da escalabilidade.

Tabela 5 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Espacial</b>	35,5363	37,6207	35,4137	40,5229	36,4899
<b>AVC Sem Escalabilidade</b>	35,5872	37,5814	35,5965	40,5223	36,6619
<b>Ganhos (em %)</b>	-0,14	0,11	-0,52	0,002	-0,47

Na tab. 5, é possível notar que o uso de escalabilidade espacial não gerou ganhos em relação ao não uso de escalabilidade. Este resultado era esperado, já que os processos internos de codificação não são alterados substancialmente nas duas versões. Dentre os cinco vídeos, dois obtiveram ganhos sobre o padrão AVC convencional e outros três apresentaram perdas. De todo modo, mesmo que os dados estejam em dB, tanto os ganhos quanto as perdas não são expressivos.

Os resultados das avaliações do PSNR da codificação com escalabilidade temporal e sem escalabilidade são apresentados na tab. 6, assim como os ganhos gerados pelo uso de escalabilidade.

Tabela 6 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Temporal</b>	36,1148	37,6063	35,2635	39,6255	36,7836
<b>AVC Sem Escalabilidade</b>	35,4629	37,1656	34,7622	39,3908	36,5559
<b>Ganhos (em %)</b>	1,8	1,17	1,42	0,57	0,62

Os resultados da tab. 6 mostram que o uso da escalabilidade temporal gerou um pequeno ganho no PSNR para todos os vídeos analisados. O vídeo que alcançou o maior ganho foi o *CITY*, que teve um aumento de 0,6519 dB. Estes ganhos são mais expressivos do que os obtidos com o uso da escalabilidade espacial, mas ainda apresentam baixa relevância.

A tab. 7 apresenta os resultados obtidos com a codificação usando escalabilidade de qualidade e sem escalabilidade. Nesta tabela, também estão apresentados os ganhos que o uso da escalabilidade alcançou sobre o codificador AVC convencional.

Tabela 7 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Escalabilidade Qualidade</b>	35,4684	37,2186	34,6416	39,3927	36,4714
<b>AVC Sem Escalabilidade</b>	35,4629	37,1655	34,7622	39,3908	36,5559
<b>Ganhos (em %)</b>	0,02	0,14	-0,35	0,005	-0,23

Os resultados da tab. 7 mostram que o uso da escalabilidade de qualidade alcançou melhores resultados que o AVC convencional, mas com ganhos muito inexpressivos. Para três vídeos a codificação escalável apresentou melhor PSNR, enquanto a codificação convencional atingiu PSNR melhor em dois casos. O vídeo que obteve maior ganho foi o *FOREMAN* com 0,0531 dB.

Na tab. 8, estão apresentados os resultados obtidos com o uso das três escalabilidades e sem o uso de escalabilidade. Novamente os ganhos com uso da escalabilidade são apresentados.

Tabela 8 - Comparação do PSNR (em dB) dos vídeos gerados pela codificação com as três escalabilidades e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Três Escalabilidades juntas</b>	40,3976	41,1878	39,8105	43,9167	40,8769
<b>AVC Sem Escalabilidade</b>	40,5523	41,5726	40,7250	44,2851	41,2159
<b>Ganhos (em %)</b>	-0,38	-0,93	-2,3	-0,84	-0,83

Os resultados da tab. 8 mostraram que o uso das três escalabilidades juntas, para todos os casos, gerou uma leve perda de qualidade em relação ao AVC convencional. Novamente, estes dados são pouco expressivos e apenas mostram que, na análise de qualidade, não são gerados ganhos ou perdas importantes com o uso da escalabilidade.

### 5.3 Resultados de Tempo de Execução

O tempo de execução utilizado na codificação dos vídeos é interessante para apresentar um indicativo da diferença de complexidade entre a codificação escalável e a codificação não escalável. Para viabilizar esta comparação, foi medido o tempo de execução da codificação com duas camadas de escalabilidade e foram somados os tempos de execução de duas codificações com o codificador não escalável: uma com configuração equivalente à usada na camada base do codificador escalável e outra com configuração equivalente à usada na camada de enriquecimento do codificador escalável.

A tab. 9 apresenta os resultados de tempo de execução para as codificações com escalabilidade espacial e sem escalabilidade. Nesta tabela também estão os ganhos que o padrão escalável obteve em relação ao padrão AVC.

Tabela 9 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Escalabilidade Espacial</b>	296	430,4	391	431,7	515,8
<b>AVC Sem Escalabilidade</b>	143,8	229,1	190	206,9	265,5
<b>Ganhos (em %)</b>	-51,42	-46,77	-51,41	-52,07	-48,53

Os resultados da tab. 9 mostram que o tempo de execução da codificação da escalabilidade espacial é maior que o tempo de execução do AVC convencional, para todos os vídeos testados. Este resultado era esperado, já que a codificação escalável utiliza todo o processo de predição inter camadas (como já foi explicado no capítulo 3), que não está presente no codificador não escalável. As perdas em tempo de execução são expressivas e a codificação escalável, na maioria dos casos, utiliza mais do que o dobro do tempo utilizado na codificação não escalável.

Na tab. 10 estão apresentados os resultados de tempo de execução para as avaliações de escalabilidade temporal e sem escalabilidade, assim como os ganhos obtidos com o uso do padrão SVC.

Tabela 10 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade temporal e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Temporal</b>	44,9	69,5	47,1	62,7	84,6
<b>AVC Sem Escalabilidade</b>	49,1	76,8	54,9	73,9	88,6
<b>Ganhos (em %)</b>	9,4	10,5	16,56	17,86	4,73

Os resultados da tab. 10 mostram que o uso da escalabilidade temporal atingiu melhores resultados de tempo de execução do que o AVC convencional. O vídeo que obteve maior ganho foi o *ICE*, diminuindo 11,2 segundos no tempo de execução da codificação. Estes ganhos não eram esperados e são um ponto positivo a favor do uso do padrão escalável, pelo menos quando apenas a escalabilidade temporal é usada.

Os resultados da codificação com escalabilidade de qualidade são apresentados na tab. 11, como também os ganhos obtidos pelo padrão escalável em relação ao AVC convencional.

Tabela 11 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com escalabilidade de qualidade e sem escalabilidade.

	<b>CITY</b>	<b>FOREMAN</b>	<b>HARBOUR</b>	<b>ICE</b>	<b>SOCCER</b>
<b>Escalabilidade Qualidade</b>	21	31,1	22,8	30,4	36,8
<b>AVC Sem Escalabilidade</b>	52	78,5	56,7	74	90,6
<b>Ganhos (em %)</b>	147,62	152,41	148,68	143,42	146,2

Os resultados da tab. 11 mostram que a escalabilidade de qualidade alcançou melhores resultados de tempo de execução do em relação ao AVC convencional. Todos os cinco vídeos obtiveram os menores tempos de execução quando codificados com escalabilidade. O vídeo *SOCCER* foi o que teve o maior ganho, diminuindo 53,8 segundos em relação ao AVC equivalente. Os resultados da codificação escalável de qualidade, em termos de tempo de execução, foram os que apresentaram maiores ganhos. Novamente, este não era um resultado esperado e é um estímulo a mais para o uso do padrão escalável.

Os resultados das comparações do uso das três escalabilidades juntas com o uso do codificador não escalável são apresentados na tab. 12, que também apresenta os ganhos do SVC em relação ao AVC.

Tabela 12 - Comparação do tempo de execução (em segundos) dos vídeos gerados pela codificação com as três escalabilidades juntas e sem escalabilidade.

	<i>CITY</i>	<i>FOREMAN</i>	<i>HARBOUR</i>	<i>ICE</i>	<i>SOCCER</i>
<b>Três Escalabilidades juntas</b>	502,9	788,3	643,6	701,8	950,1
<b>AVC Sem Escalabilidade</b>	156,2	242,9	209,1	214,5	281,7
<b>Ganhos (em %)</b>	-68,94	-69,19	-67,51	-69,44	-70,35

Os resultados da tab. 12, mostram que o uso das três escalabilidades juntas causou uma importante perda no tempo de execução quando comparado ao padrão não escalável. Em todos os casos, o uso da codificação com as três escalabilidades utilizou mais do que o dobro do tempo de execução utilizado na codificação não escalável. A explicação para este consumo excessivo de tempo deve-se à necessidade do padrão escalável testar todas as possibilidades de escalabilidades antes de decidir qual será utilizada em cada bloco do vídeo.

## 6 CONCLUSÕES

Este trabalho de conclusão de curso apresentou uma análise da qualidade, da complexidade e da eficiência de codificação do padrão escalável de compressão de vídeos H.264/SVC, em relação ao padrão de compressão de vídeos não escalável H.264/AVC. Para realizar estas análises, foi utilizado o software de referência JSVM e cinco amostras de vídeos, garantindo que a análise fosse realizada com uma amostragem representativa. Para cada vídeo de entrada, foram utilizados cinco tipos de avaliações. Ao total, foram realizadas 350 execuções. Todos os critérios adotados para a avaliação do padrão escalável foram apresentadas e os resultados obtidos foram discutidos.

Também foram apresentados, neste trabalho, alguns conceitos de compressão de vídeos, assim como uma introdução ao padrão H.264/AVC e à sua extensão escalável H.264/SVC. Foi descrito, também, o software de referência e seus executáveis, dando um maior enfoque nos executáveis utilizados neste trabalho.

Para a análise da qualidade do padrão escalável, o critério adotado foi o PSNR. As avaliações feitas mostraram que o padrão escalável SVC não obteve um ganho expressivo em relação ao padrão não escalável AVC. Na comparação do PSNR (em dB) dos vídeos gerados pela codificação com escalabilidade espacial e sem escalabilidade, o SVC gerou ganhos apenas em dois dos cinco vídeos analisados. Já na comparação entre a escalabilidade temporal e sem escalabilidade, o uso da escalabilidade gerou um pequeno ganho de PSNR em todos os vídeos analisados. Na comparação entre escalabilidade de qualidade e sem escalabilidade, o uso da escalabilidade alcançou melhores resultados que o AVC, mas mesmo assim não gerou ganhos expressivos sobre o padrão sem escalabilidade. Para a comparação do PSNR no uso das três escalabilidades e sem escalabilidade, o uso da escalabilidade gerou uma pequena perda em todos os vídeos em relação ao



padrão AVC. Através destas comparações foi possível concluir que o padrão escalável não gera ganhos nem perdas expressivas em termos de PSNR e, por consequência, na qualidade dos vídeos comprimidos.

Na análise da complexidade do padrão escalável o critério utilizado foi o tempo de execução da codificação com e sem o uso da escalabilidade. As avaliações feitas mostraram que o padrão escalável perde expressivamente no tempo de codificação na comparação do uso da escalabilidade espacial e das três escalabilidades juntas, em relação ao padrão sem escalabilidade (a maioria dos casos usa mais do que o dobro de tempo utilizado na codificação não escalável). Já para as comparações com uso da escalabilidade temporal e de qualidade, obteve-se um ganho em relação ao padrão sem escalabilidade. Este ganho não era esperado, sendo um ponto positivo para o padrão escalável.

Na análise da eficiência da codificação do padrão escalável em relação ao padrão não escalável, as avaliações feitas mostraram que o padrão escalável obteve um ganho significativo em relação ao padrão não escalável. Na comparação com escalabilidade espacial e sem escalabilidade, quatro dos cinco vídeos alcançaram uma taxa de compressão maior com o uso da escalabilidade. Já na comparação com escalabilidade temporal e sem escalabilidade, o ganho foi mais expressivo e todos os vídeos obtiveram uma taxa de compressão maior que o sem escalabilidade. Para a comparação com escalabilidade de qualidade e sem escalabilidade, o uso da escalabilidade também alcançou melhores taxas de compressão. Por fim, na comparação com as três escalabilidades juntas e sem escalabilidade, novamente o uso da escalabilidade obteve um ganho significativo, onde todos os vídeos alcançaram taxas de compressão maior.

Com isto, pode-se concluir que para a análise da qualidade e da complexidade o padrão escalável não obteve ganhos expressivos em relação ao padrão não escalável. Já para a análise da eficiência da codificação do padrão escalável, foi possível concluir que os ganhos foram significativos quando comparado ao padrão não escalável.

Como trabalhos futuros pretende-se utilizar vídeos com resoluções mais elevadas nas avaliações, passando de resoluções QCIF (176x144 pixels) e CIF

(352x288 pixels) para resoluções de alta definição como HDTV 1080 (1920x1080 pixels).

Também estão previstos o aumento no número de amostras de vídeos utilizados, aplicando as mesmas avaliações realizadas neste trabalho, para gerar um conjunto de resultados ainda mais representativo.

## Referências

AGOSTINI, Luciano Volcan. **Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas à Compressão de Vídeo Segundo o Padrão H.264/AVC**. Tese (Doutorado em Microeletrônica). Universidade Federal do Rio Grande do Sul. Porto Alegre, 2007.

BHASKARAN, V.; KONSTANTINIDES, K. **Image and Video Compression Standards: Algorithms and Architectures**. 2. ed. Boston: Kluwer Academic Publishers, 1997.

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**. São Paulo: Edgard Blücher, 2003.

HUANG, HSIANG-CHUN; Peng, Wen-Hsiao; Chiang, Tihao; Hang, Hsueh-Ming. **Advances in the Scalable Amendment of H.264/AVC**. Communications Magazine, IEEE. Vol. 45, No. 1, pp. 68-76, Jan. 2007.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Home**. Disponível em: <[www.itu.int/ITU-T/](http://www.itu.int/ITU-T/)>. Acesso em: maio 2007.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO - International Organization for Standardization. Disponível em: <<http://www.iso.org>>. Acesso em: maio 2007.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 14496-2 - MPEG-4 Part 2 (01/1999)**: coding of audio visual objects – part 2: visual. [S.l.], 1999.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 11172 - MPEG-1 (11/1993)**: coding of moving pictures and associated audio for digital storage media up to about 1.5Mbit/s – part 2: video. [S.l.], 1993.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Recommendation H.264/AVC (03/05)**: advanced video coding for generic audiovisual services. [S.l.], 2005.

INTERNATIONAL TELECOMMUNICATION UNION. **JVT-L050: H.264/AVC AVC Fidelity Range Extensions**. [S.l.], 2004.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Recommendation H.264/AVC (05/03)**: advanced video coding for generic audiovisual services. [S.I.], 2003.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Recommendation H.263 v3 (11/00)**: video coding for low bit rate communication. [S.I.], 2000.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Recommendation H.262 (11/94)**: generic coding of moving pictures and associated audio information – part 2: video. [S.I.], 1994.

INTERNATIONAL TELECOMMUNICATION UNION. **ITU-T Recommendation H.261 v1 (11/90)**: video codec for audiovisual services at px64 kbit/s. [S.I.], 1990.

INTERNATIONAL TELECOMMUNICATION UNION. **Joint Video Team (JVT)**. Disponível em: <<http://www.itu.int/ITU-T/studygroups/com16/jvt/>>. Acesso em: maio 2007a.

ITU-T e ISO/IEC JTC1.”**Joint Scalable Video Model JSVM-9.12.2**”. Abril, 2008.

JVT Editors (T. Wiegand, G. Sullivan, A. Luthra), **Draft ITU-T Recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264|ISO/IEC 14496-10 AVC)**, 2003.

JVT Editors (T. Wiegand, G. Sullivan, A. Luthra), **Draft ITU-T Recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264|ISO/IEC 14496-10 AVC)**, 2007.

PURI, A.; et al. **Video Coding Using the H.264/MPEG-4 AVC Compression Standard**. Elsevier **Signal Processing: Image Communication**. [S.I.], n. 19, p.793–849, 2004.

RICHARDSON, I. **Video Codec Design – Developing Image and Video Compression Systems**. Chichester: John Wiley and Sons, 2002.

RICHARDSON, I. **H.264/AVC and MPEG-4 Video Compression – Video Coding for Next-Generation Multimedia**. Chichester: John Wiley and Sons, 2003.

SCHWARZ, H.; MARPE, D.; WIEGAND, T. **Hierarchical B Pictures, Joint Video Team, Doc. JVT-P014**, Jul. 2005.

SCHWARZ, H.; MARPE, D.; WIEGAND, T. “**Analysis of hierarchical B-pictures and MCTF**,” in Proc. ICME, Toronto, ON, Canada, Jul. 2006, pp. 1929-1932.

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, pp. 1103-1120, Sept. 2007, invited paper.

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.

SHI, Y.; SUN, H. **Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms and Standards**. Boca Raton: CRC Press, 1999.

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

SULLIVAN, G. et al. **The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions**. In: CONFERENCE ON APPLICATIONS OF DIGITAL IMAGE PROCESSING, SPIE, 2004. **Proceedings...** Denver: SPIE, 2004.

TNT - Institut fuer Informationsverarbeitung - Institute os Signal Processing, Universitaet Hannover. "**Test Sequences**." Disponível em: <ftp.tnt.uni-hannover.de/pub/svc/testsequences/>. Acesso em: junho, 2008.

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.

## **Apêndices**

## APÊNDICE A – Arquivos de Configuração

Este apêndice apresenta um exemplo de todos os arquivos de configuração usados nas avaliações apresentadas neste trabalho. Todos os arquivos aqui apresentados consideram que o vídeo de entrada é o *CITY*. Para cada novo vídeo, todos os arquivos devem ser modificados.

### A.1 Configuração para a Codificação Com Escalabilidade Espacial

Para a avaliação do uso da escalabilidade espacial, utiliza-se um arquivo de configuração geral e dois arquivos de configurações, um para cada camada definida. Nos arquivos de configurações de camadas deve se manter a mesma taxa de quadros e o mesmo QP do arquivo geral, mudando apenas o valor da resolução de uma camada para outra. Os três arquivos de configuração necessários são apresentados abaixo.

```
# Arquivo de configuração geral para a codificação com
# escalabilidade espacial – MAIN.CFG
AVCMode          0          # SVC: 0, AVC:1
OutputFile        stream.264 # arquivo de saída
FrameRate         15.0       # taxa de quadro [Hz]
FramesToBeEncoded 105       # número de quadros
GOPSize           1          # tamanho do GOP
NumLayers         2          # número de camadas
LayerCfg          city_layer0.cfg # arquivo de config. da camada base
LayerCfg          city_layer1.cfg # arquivo de config. da camada de
enriquecimento
```

```
# Arquivo de configuração da primeira camada da codificação com
# escalabilidade espacial – CITY_LAYER0.CFG
SourceWidth    176                # largura do vídeo de entrada
SourceHeight   144                # altura do vídeo de entrada
FrameRateIn    15                 # taxa de quadro de entrada [Hz]
FrameRateOut   15                 # taxa de quadro de saída [Hz]
InputFile      city_176x144_15.yuv # vídeo de entrada
ReconFile      city_layer0.yuv     # Reconstructed file
QP             28.0                # QP da camada base
```

```
# Arquivo de configuração da primeira camada da codificação com
# escalabilidade espacial – CITY_LAYER1.CFG
SourceWidth    352                # largura do vídeo de entrada
SourceHeight   288                # altura do vídeo de entrada
FrameRateIn    15                 # taxa de quadro de entrada [Hz]
FrameRateOut   15                 # taxa de quadro de saída [Hz]
InputFile      city_352x288.yuv   # vídeo de entrada
ReconFile      city_layer1.yuv    # Reconstructed file
QP             28.0                # QP da camada de enriquecimento
```

## A.2 Configuração para a Codificação Com Escalabilidade Temporal

Para a avaliação do uso de escalabilidade temporal, também utiliza-se um arquivo de configuração geral e um arquivo de configuração adicional por camada. Os arquivos de configuração de camada devem ter taxa de quadro igual a 30 quadros por segundo, resolução QCIF e QP igual a 28. Os três arquivos de configuração necessários são apresentados abaixo.



# Arquivo de configuração geral para a codificação com  
# escalabilidade temporal – MAIN.CFG

```

AVCMode          0          # SVC: 0, AVC:1
OutputFile       stream.264 # arquivo de saída
FrameRate        30.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
GOPSize          2         # tamanho do GOP
NumLayers        1         # número de camadas
LayerCfg         city_layer0.cfg # arquivo de config.da camada de
enriquecimento

```

# Arquivo de configuração da primeira camada da codificação com  
# escalabilidade temporal – CITY\_LAYER0.CFG

```

SourceWidth      176          # largura do vídeo de entrada
SourceHeight     144          # altura do vídeo de entrada
FrameRateIn      30           # taxa de quadro de entrada [Hz]
FrameRateOut     30           # taxa de quadro de saída [Hz]
InputFile        city_176x144.yuv # video de entrada
ReconFile        city_layer0.yuv # Reconstructed file
QP               28.0         # QP da camada base

```

### **A.3 Configuração para a Codificação Com Escalabilidade de Qualidade**

Para a avaliação utilizando escalabilidade de qualidade, novamente é utilizado um arquivo de configuração geral e dois arquivos de configurações de camadas. Nos arquivos de configurações de camadas deve-se manter a mesma resolução (QCIF) e a mesma taxa de quadros (15 quadros por segundo), mudando apenas o valor do QP de 28 para 22 de uma camada para outra. Os três arquivos usados nesta avaliação estão listados abaixo.

# Arquivo de configuração geral para a codificação com

# escalabilidade de qualidade – MAIN.CFG

```

AVCMode          0          # SVC: 0, AVC:1
OutputFile       stream.264 # arquivo de saída
FrameRate        15.0       # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
GOPSize          1          # tamanho do GOP
NumLayers        2          # número de camadas
LayerCfg         city_layer0.cfg # arquivo de config. da camada base
LayerCfg         city_layer1.cfg # arquivo de config. da camada de
enriquecimento

```

# Arquivo de configuração da primeira camada da codificação com

# escalabilidade de qualidade – CITY\_LAYER0.CFG

```

SourceWidth      176          # largura do vídeo de entrada
SourceHeight     144          # altura do vídeo de entrada
FrameRateIn      15           # taxa de quadro de entrada [Hz]
FrameRateOut     15           # taxa de quadro de saída [Hz]
InputFile        city_176x144.yuv # video de entrada
ReconFile        city_layer0.yuv  # Reconstructed file
QP               28.0         # QP da camada base

```

# Arquivo de configuração da primeira camada da codificação com

# escalabilidade de qualidade – CITY\_LAYER0.CFG

```

SourceWidth      176          # largura do vídeo de entrada
SourceHeight     144          # altura do vídeo de entrada
FrameRateIn      15           # taxa de quadro de entrada [Hz]
FrameRateOut     15           # taxa de quadro de saída [Hz]
InputFile        city_176x144.yuv # vídeo de entrada
ReconFile        city_layer1.yuv  # Reconstructed file
QP               22.0         # QP da camada de enriquecimento

```

#### A.4 Configuração para a Codificação Com as Três Escalabilidades Juntas

Para a avaliação utilizando as três escalabilidades juntas, também é necessário um arquivo de configuração geral e dois arquivos de configurações de camadas. Os arquivos de configurações de camadas devem conter diferentes especificações de resolução, de taxa de quadros e de QP. Os três arquivos usados nesta avaliação estão listados abaixo.

# Arquivo de configuração geral para a codificação com

# as três escalabilidades juntas – MAIN.CFG

```

AVCMode          0          # SVC: 0, AVC:1
OutputFile       stream.264 # arquivo de saída
FrameRate        15.0       # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
GOPSize          2          # tamanho do GOP
NumLayers        2          # número de camadas
LayerCfg         city_layer0.cfg # arquivo de config. da camada base
LayerCfg         city_layer1.cfg # arquivo de config. da camada de
enriquecimento

```

# Arquivo de configuração da primeira camada da codificação com

# as três escalabilidades juntas – CITY\_LAYER0.CFG

```

SourceWidth      176          # largura do vídeo de entrada
SourceHeight     144          # altura do vídeo de entrada
FrameRateIn      15          # taxa de quadro de entrada [Hz]
FrameRateOut     15          # taxa de quadro de saída [Hz]
InputFile        city_176x144.yuv # vídeo de entrada
ReconFile        city_layer0.yuv # Reconstructed file
QP               28.0         # QP da camada base

```

```

# Arquivo de configuração da primeira camada da codificação com
# as três escalabilidades juntas – CITY_LAYER0.CFG
SourceWidth    352                # largura do vídeo de entrada
SourceHeight   288                # altura do vídeo de entrada
FrameRateIn    30                 # taxa de quadro de entrada [Hz]
FrameRateOut   30                 # taxa de quadro de saída [Hz]
InputFile      city_352x288.yuv   # vídeo de entrada
ReconFile      city_layer1.yuv    # Reconstructed file
QP             22.0               # QP da camada de enriquecimento

```

### **A.5 Configuração para a Codificação Sem Escalabilidade com Configuração Equivalente ao Codificador com Escalabilidade Espacial**

Para a codificação sem escalabilidade, foi necessário criar dois arquivos de configuração, um com configuração equivalente a camada base do codificador escalável e o outro com configuração equivalente a camada de enriquecimento do codificador escalável.

No arquivo de configuração do codificador AVC sem escalabilidade equivalente à camada base usada com a escalabilidade espacial, são necessárias as seguintes especificações: resolução QCIF, taxa de quadros igual a 15 quadros por segundo e QP igual a 28. Para o arquivo de configuração equivalente à camada de enriquecimento, apenas muda-se o valor da resolução para a CIF. Os dois arquivos de configuração estão apresentados abaixo.

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada base
# do codificador com escalabilidade espacial
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_176x144.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      176        # largura do vídeo de entrada
SourceHeight     144        # altura do vídeo de entrada
FrameRate        15.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          28        # QP

```

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada de enriquecimento
# do codificador com escalabilidade espacial
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_352x288.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      352        # largura do vídeo de entrada
SourceHeight     288        # altura do vídeo de entrada
FrameRate        15.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          28        # QP

```

Para o arquivo de configuração da camada base no formato AVC compatível com a escalabilidade temporal, as especificações devem ser: resolução QCIF, taxa de quadros igual a 15 e QP igual a 28. Para o arquivo de configuração da camada de enriquecimento, troca-se o valor da taxa de quadro para 30.

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada base
# do codificador com escalabilidade temporal
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_176x144.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      176        # largura do vídeo de entrada
SourceHeight     144        # altura do vídeo de entrada
FrameRate        15.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          28        # QP

```

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada de enriquecimento
# do codificador com escalabilidade temporal
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_352x288.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      176        # largura do vídeo de entrada
SourceHeight     144        # altura do vídeo de entrada
FrameRate        30.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          28        # QP

```

No arquivo de configuração da camada base no formato AVC compatível com a escalabilidade de qualidade, deve ter as seguintes especificações: resolução QCIF, taxa de quadros igual a 15 e QP igual a 28. Para o arquivo de configuração da camada de enriquecimento, muda-se apenas o valor do QP.

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada base
# do codificador com escalabilidade de qualidade
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_176x144.yuv  # vídeo de entrada
OutputFile       stream.264 # arquivo de saída
SourceWidth      176        # largura do vídeo de entrada
SourceHeight     144        # altura do vídeo de entrada
FrameRate        15.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          28        # QP

```

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada de enriquecimento
# do codificador com escalabilidade de qualidade
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_352x288.yuv  # vídeo de entrada
OutputFile       stream.264 # arquivo de saída
SourceWidth      176        # largura do vídeo de entrada
SourceHeight     144        # altura do vídeo de entrada
FrameRate        15.0      # taxa de quadro [Hz]
FramesToBeEncoded 105      # número de quadros
BasisQP          22        # QP

```

Já para o arquivo de configuração da camada base no formato AVC compatível com o uso das três escalabilidades juntas, deve ter as seguintes especificações: resolução QCIF, taxa de quadros igual a 15 e QP igual a 28. Para o arquivo de configuração da camada de enriquecimento, o valor da resolução passa a ser CIF, a taxa de quadros passa a ser 30 e o QP passa para 22.

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada base
# do codificador com as três escalabilidades juntas
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_176x144.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      176         # largura do vídeo de entrada
SourceHeight     144         # altura do vídeo de entrada
FrameRate        15.0       # taxa de quadro [Hz]
FramesToBeEncoded 105       # número de quadros
BasisQP          28         # QP

```

```

# Arquivo de configuração para a codificação sem escalabilidade
# com configuração equivalente à camada de enriquecimento
# do codificador com as três escalabilidades juntas
AVCMode          1          # SVC: 0, AVC:1
InputFile        city_352x288.yuv  # vídeo de entrada
OutputFile       stream.264  # arquivo de saída
SourceWidth      352         # largura do vídeo de entrada
SourceHeight     288         # altura do vídeo de entrada
FrameRate        30.0       # taxa de quadro [Hz]
FramesToBeEncoded 105       # número de quadros
BasisQP          22         # QP

```



## APÊNDICE B – Publicações durante a graduação

### B.1 Trabalhos completos em anais de eventos

SAMPAIO, Felipe; REDIESS, Fabiane; FONSECA, Carolina Marques; SUSIN, Altamiro; BAMPI, Sérgio; AGOSTINI, Luciano. **Architectural Design for Forward and Inverse Transforms of H.264/AVC Standard Focusing in the Intra Frame Coder**. In: VLSI-SoC 2008 - The 16th IFIP/IEEE International Conference on Very Large Scale Integration, 2008, Ilha de Rhodes. VLSI-SoC 2008 - The 16th IFIP/IEEE International Conference on Very Large Scale Integration. Piscataway: IEEE, 2008.

SAMPAIO, Felipe; FONSECA, Carolina; REDIESS, Fabiane; BAMPI, Sérgio; AGOSTINI, Luciano. **Desenvolvimento de uma Arquitetura Paralela para a Hadamard 2-D 2x2 do Padrão H.264/AVC**. In: Iberchip 2008 - XIV Workshop IBERCHIP, 2008, Puebla - México.

REDIESS, Fabiane; SAMPAIO, Felipe Martin; FONSECA, Carolina; BAMPI, Sérgio; SUSIN, Altamiro; AGOSTINI, Luciano. **Architectural Design for Forward and Inverse 4x4 Transforms of H.264/AVC Standard Focusing in the Intra Frame Coder**. In: SIM 2008 - XXIII Simpósio Sul de Microeletrônica, 2008, Bento Gonçalves. XXIII Simpósio Sul de Microeletrônica. Porto Alegre, 2008. p. 123-126.

SAMPAIO, Felipe; REDIESS, Fabiane; FONSECA, Carolina; BAMPI, Sérgio; SUSIN, Altamiro; AGOSTINI, Luciano. **Architecture Design and Prototyping of a Fully Parallel H.264/AVC 2x2 Hadamard Transform Targeting Real Time in Very High Resolution Videos**. In: SIM 2008 - XXIII Simpósio Sul de Microeletrônica, 2008, Bento Gonçalves. XXIII Simpósio Sul de Microeletrônica. Porto Alegre, 2008. p. 131-134.

## **B.2 Resumos simples em anais de eventos**

FONSECA, Carolina; AGOSTINI, Luciano. **Investigação em Software sobre o Padrão H.264 Escalável de Compressão de Vídeos**. In: XVII Congresso de Iniciação Científica – UFPel, 2008, Pelotas.

FONSECA, Carolina; SAMPAIO, Felipe; REDIESS, Fabiane; AGOSTINI, Luciano. **Prototipação em FPGAs Xilinx da Arquitetura Paralela da Transformada Hadamard 2-D 2x2 do Padrão H.264/AVC de Compressão de Vídeos**. In: CIC 2007 - XVI Congresso de Iniciação Científica - UFPel, 2007, Pelotas.

SAMPAIO, Felipe; FONSECA, Carolina; REDIESS, Fabiane; AGOSTINI, Luciano. **Desenvolvimento de um Protótipo em FPGAs Altera da Arquitetura Serial da Transformada Hadamard 2-D 2x2 do Padrão H.264/AVC com Foco em Vídeos de Alta Resolução**. In: CIC 2007 - XVI Congresso de Iniciação Científica - UFPel, 2007, Pelotas.