UNIVERSIDADE FEDERAL DE PELOTAS

Centro de Desenvolvimento Tecnológico (CDTec)

Curso de Bacharelado em Ciência da Computação

Trabalho Acadêmico

Desenvolvimento e Implementação em Hardware de Heurísticas Rápidas Para o Módulo de Decisão de Modo do Padrão H.264/AVC de Codificação de Vídeo

Daniel Munari Vilchez Palomino

Pelotas, 2010

DANIEL MUNARI VILCHEZ PALOMINO

Desenvolvimento e Implementação em Hardware de Heurísticas Rápidas Para o Módulo de Decisão de Modo do Padrão H.264/AVC de Codificação de Vídeo

Trabalho acadêmico 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: Dr. Luciano Volcan Agostini
Co-orientador: MSc. Guilherme Corrêa

PELOTAS
2010

Dados de catalogação na fonte:
Maria Beatriz Vaghetti Vieira – CRB-10/1032
Biblioteca de Ciência & Tecnologia - UFPel

P181d Palomino, Daniel Munari Vilchez
Desenvolvimento e implementação em hardware de heurísticas rápidas para o módulo de decisão de modo do padrão H.264/AVC de codificação de vídeo / Daniel Munari Vilchez Palomino. - 91f. : fig. - Monografia (Conclusão de curso). Universidade Federal de Pelotas. Centro de Desenvolvimento Tecnológico. Curso de Bacharelado em Ciência da Computação. Pelotas, 2011. – Orientador Luciano Volcan Agostini; co-orientador Guilhewrme Ribeiro Corrêa.

1.Informática. 2.Codificação de vídeo. 3.Decisão de modo. 4.H.264/AVC. 5.Projeto de hardware. I.Agostini, Luciano Volcan. II.Corrêa, Guilherme Ribeiro. III.Título.

CDD: 006.72

Banca examinadora:

.....

Prof. Dr. Luciano Volcan Agostini (Orientador)

.....

MSc. Cláudio Diniz

.....

Prof. Dr. Ricardo Matsumura Araújo

.....

Prof. Dr. Leomar da Rosa Jr.

AGRADECIMENTOS

Agradeço primeiramente a todos os colegas, professores e amigos que descobri durante a graduação no curso de ciência da computação. Agradeço a todos os colegas que trabalharam comigo no GACI, compartilhando muitas xícaras de café e muitas idéias.

Um agradecimento especial aos colegas Felipe Sampaio, e Robson Dornelles, que desde o inicio do curso trabalharam comigo tanto nas atividades de pesquisa quanto nos trabalhos de disciplina. Valeu caras, vocês foram muito importantes.

Dedico outro agradecimento especial ao professor Luciano Volcan Agostini, que, além de meu orientador neste trabalho e chefe nas atividades de pesquisa, é um grande amigo e foi muito importante no meu desenvolvimento como aluno, como pessoa e como profissional. Ao Guilherme Corrêa que foi meu co-orientador neste trabalho e me ajudou muito na realização desta tarefa.

Por fim, um agradecimento mais que especial aos meus pais, que sempre trabalharam priorizando que, eu junto com meus irmãos, sempre estudássemos. Graças aos esforços e ensinamentos deles me tornei a pessoa que sou e com certeza, não só este trabalho, mas como tudo de produtivo que eu fiz na minha vida eu devo a eles.

“Nunca deixe ninguém dizer
que você não merece o que você quer”

RESUMO

PALOMINO, Daniel M. **Desenvolvimento e Implementação em Hardware de Heurísticas Rápidas Para o Módulo de Decisão de Modo do Padrão H.264/AVC de Codificação de Vídeo.** 2010. 91f. Trabalho acadêmico (Graduação) - Bacharelado em Ciência da Computação. Universidade Federal de Pelotas, Pelotas.

O padrão H.264/AVC define vários modos de predição para codificar um macrobloco (bloco de tamanho 16x16 pixels). A predição sobre um macrobloco pode ser intra-quadro, ou seja, quando o macrobloco atual é representado através de uma interpolação dos pixels vizinhos a ele, ou inter-quadros, quando o macrobloco atual é representado pelo macrobloco presente nos quadros de referência que seja mais semelhante a ele. Considerando a predição intra-quadro para amostras de luminância, existem dois tamanhos de bloco possíveis: (1) I16MB (intra 16x16), com 4 modos de codificação possíveis sobre blocos 16x16 e (2) I4MB (intra 4x4), com 9 modos de codificação possíveis sobre blocos 4x4. Já na predição inter-quadros, sete tamanhos de bloco são suportados (16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4), gerando um grande número de possibilidades de particionamento para o bloco 16x16. O módulo de decisão de modo presente em codificadores H.264/AVC é responsável por escolher, entre os vários modos de predição citados acima, qual será utilizado para codificar um determinado macrobloco. Essa escolha não é uma tarefa trivial, já que a quantidade de dados disponível após as etapas de predição não é definitiva. A técnica de RDO (*Rate-Distortion Optimization*) gera os melhores resultados em termos de relação taxa-distorção em um vídeo, pois a escolha do melhor modo é realizada somente após o final do processo de codificação. Entretanto, o custo computacional para completar o processo de codificação para todos os modos de predição possíveis é muito alto, dificultando a codificação de vídeos de alta resolução em tempo real (24 a 30 quadros por segundo). Este trabalho apresenta um conjunto de heurísticas e suas respectivas implementações em hardware, desenvolvidas para realizar a decisão de modo do padrão H.264/AVC da forma mais rápida possível. As heurísticas foram desenvolvidas com o objetivo de eliminar completamente o uso da técnica RDO no processo de decisão de modo. Foram geradas duas soluções para a decisão de modo de macroblocos codificados pela decisão intra-quadro e uma solução para a decisão de modo inter-quadros. Os resultados mostraram uma redução de 13 vezes no número de cálculos no processo de decisão intra-quadro e 33 vezes para a decisão inter-quadros, quando comparadas com a técnica RDO, apesar de pequenas perdas em taxa de compressão e qualidade do vídeo codificado. Além disso, este trabalho apresentou a maior redução no número de cálculos quando comparado com trabalhos da literatura.

Palavras-chave: Codificação de vídeo, decisão de modo, H.264/AVC, projeto de hardware.

ABSTRACT

PALOMINO, Daniel M. **Implementation and Hardware Design of Fast Heuristics for the Mode Decision of the H.264/AVC Video Coding Standard.** 2010. 91f. Trabalho acadêmico (Graduação) - Bacharelado em Ciência da Computação. Universidade Federal de Pelotas, Pelotas.

The H.264/AVC standard defines several prediction modes to encode one macroblock (16x16 block size). The macroblock prediction type can be intra-frame, when the current macroblock is predicted as an interpolationof the neighbor pixels, or inter-frames, when the current macroblock is predicted using reconstructed macroblocks from other frames. In the intra-frame prediction, there are two possible block sizes to encode a macroblock: (1) I16MB, with 4 possible prediction modes applied to 16x16 blocks and (2) I4MB, with 9 possible prediction modes applied to 4x4 blocks. Considering the inter-frames prediction, there are seven block sizes (16x16, 16x8, 8x16, 8x8, 8x4, 4x8 and 4x4), which generates a large number of possibilities to encode one 16x16 block. This choice is not an easy task, since the data available after the prediction steps are not conclusive. The RDO (Rate-Distortion Optimization) technique generates the best results considering the rate-distortion relation in a video, since the best mode choice is performed only after the whole encoder process. However, the computational cost of completing the encoding process for every possible prediction mode is too high, making difficult the encoding of high resolution videos in real time (24 to 30 frames per second). This work presents a set of heuristics and its hardware implementation developed to perform the H.264/AVC mode decision as fast as possible. The heuristics were developed aiming the complete elimination of the RDO technique in the H.264/AVC mode decision process. Two solutions were developed for the intra-frame decision and another one for the inter-frame decision. When compared with RDO, the results have shown a reduction of 13 times in number of operations for intra-frame decision and 33 times for the inter-frame decision. Besides, this work has shown the largest reduction in number of operations when compared with related works.

Keywords: Video coding, mode decision, H.264/AVC, hardware design.

LISTA DE FIGURAS

Figura 1 - Diagrama em blocos do codificador H.264/AVC. 23
Figura 2 - Partições de macrobloco definidas no padrão H.264/AVC.....27
Figura 3 - Sub-partições de macrobloco definidas no padrão H.264/AVC. 27
Figura 4 - Tamanhos de bloco utilizados na predição intra-quadro do padrão H.264/AVC para amostras de luminância. 29
Figura 5 - Amostras usadas na predição intra-quadro considerando macroblocos I4MB. 29
Figura 6 - Modos de predição intra-quadro para I4MB. 30
Figura 7 - Modos de predição intra-quadro para I16MB. 30
Figura 8 - Modos de predição intra-quadro para crominância 8x8. 32
Figura 9 - Curva típica da relação entre taxa de bits e distorção.....35
Figura 10 - Calculo do custo taxa-distorção (*RDcost*). 36
Figura 11 - Gráfico da quantidade de modos escolhidos pela técnica RDO em relação à diferença de distorção baseada em SAD. 47
Figura 12 - Hierarquia para decisão de particionamento e sub-particionamento na predição inter-quadros. 52
Figura 13 - Valores de H para macrobloco codificados sem sub-particionamento através da decisão RDO. 56
Figura 14 - Valores de H para macrobloco codificados com sub-particionamento através da decisão RDO. 56
Figura 15 - Possível particionamento sobre uma imagem. 58
Figura 16 - Amostras (em cinza) utilizadas na detecção de arestas vertical e horizontal para decisão de particionamento de macrobloco. 59

Figura 17 - Amostras (cinza) utilizadas para o cálculo das bordas verticais e horizontais para a decisão de tamanho de sub-particionamento de macrobloco.....62
Figura 18 - Arquitetura do Calculador de SAD para oito amostras.....70
Figura 19 - Arquitetura que gera os coeficientes da transformada DCT 16x16 parcial necessários para o cálculo da métrica de heterogeneidade.....71
Figura 20 - Diagrama em blocos da arquitetura do módulo de decisão intra-quadro baseado em heterogeneidade.....72
Figura 21 - Diagrama de ciclos para a arquitetura de decisão intra-quadro baseada em heterogeneidade. 73
Figura 22 - Diagrama em blocos da arquitetura do módulo de decisão intra-quadro baseado em diferença de distorção. 75
Figura 23 - Diagrama de ciclos para a arquitetura de decisão intra-quadro baseada em diferença de distorção. 76
Figura 24 - Arquitetura para o cálculo de força de borda para decisão de tamanho de partição. 78
Figura 25 - Arquitetura para o cálculo de força de borda para decisão de tamanho de sub-partição. 78
Figura 26 - Diagrama em blocos da arquitetura de decisão de modo inter-quadros completa.....79
Figura 27 - Diagrama de ciclos para a arquitetura de decisão inter-quadros completa.....80

LISTA DE TABELAS

Tabela 1 - Ocorrência de modos intra-quadro e inter-quadros em quadros P de resolução CIF. 43

Tabela 2 - Comparação entre a codificação com predições intra e inter em quadros P e a codificação com apenas predição inter em quadros P. 43

Tabela 3 - Resultados obtidos com SAD, SSD e SATD. 46

Tabela 4 - Comparação de SAD com SSD e SATD. 46

Tabela 5 - Resultados do módulo de decisão intra-quadro completo (SAD & DD). 48

Tabela 6 - Resultados médios obtidos com a técnica RDO, Heurística baseada em Heterogeneidade e a Heurística DD. 49

Tabela 7 - Comparação entre RDO vs. Heterogeneidade e RDO vs. DD. 49

Tabela 8 - Ocorrência de modos de codificação em quadros P de resolução HD1080p. 51

Tabela 9 - Decisão de modo SKIP baseada em estacionariedade. 54

Tabela 10 - Resultados da detecção de sub-particionamento em comparação com o RDO. 57

Tabela 11 - Resultados gerados pela decisão de particionamento baseada em força de borda. 61

Tabela 12 - Resultados gerados pela decisão de tamanho de sub-partição baseada em força de borda. 64

Tabela 13 - Resultados gerados pelo módulo completo de decisão rápida inter-quadros. 65

Tabela 14 - Comparação com trabalhos da literatura. 67

Tabela 15 - Resultados de síntese da arquitetura de decisão intra-quadro baseada em heterogeneidade. 74

Tabela 16 - Resultados de síntese da arquitetura de decisão intra-quadro baseada em diferença de distorção. 76
Tabela 17 - Resultados de síntese da arquitetura de decisão inter-quadros completa. 81

LISTA DE ABREVIATURAS E SIGLAS

ALUT	Adaptative Look-up Table
AVC	Advanced Video Coding
BH	Borda Horizontal
BV	Borda Vertical
CABAC	Context-Based Adaptive Arithmetic Coding
CIF	Common Intermediate Format
dB	Decibel
DCT	Discrete Cosine Tranform
DD	Diferença de Distorção
DLR	Dedicated Logic Register
FPGA	Field-Programmable Gate Array
GOP	Group of Pictures
HD	High Definition
ISO/IEC	Joint Technical Committee of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC)
ITU-T	Telecommunication Standardization Sector
JVT	Joint Video Team
Kbps	Kilo bits por segundo
MB	Macrobloco
MC	Motion Compensation
ME	Motion Estimation
MHz	Mega Hertz
MPEG	Motion Picture Experts Group
MSE	Mean Squared Error
PSNR	PeakSignal-to-Noise Rate

RDO	Rate-Distortion Optimization
RD	Rate-Distortion
RGB	Red Green Blue
SAD	Sum of Absolute Differences
SATD	Sum of Absolute Transformed Differences
SBTVD	Sistema Brasileiro de Televisão Digital
SBH	Sub Borda Horizontal
SBV	Sub Borda Vertical
SSD	Sum of Squared Differences
VCEG	Video Coding Experts Group
YCbCr	Luminância, Crominância Azul, Crominância Vermelha

SUMÁRIO

1 INTRODUÇÃO 17

2 CODIFICAÇÃO DE VÍDEO E O PADRÃO H.264/AVC 19
2.1 Características dos Vídeos Digitais 19
2.2 Redundância de Dados na Representação de Vídeos 20
2.3 O Padrão H.264/AVC 21

3 PREDIÇÃO INTER-QUADROS E INTRA-QUADRO NO PADRÃO H.264/AVC 25
3.1 Predição Inter-Quadros 25
3.2 Predição Intra-Quadro 28
3.2.1 Intra 4x4 (I4MB) 29
3.2.2 Intra 16x16 (I16MB) 30
3.2.3 Predição para Crominância 31

4 MODO DE DECISÃO 33
4.1 Métricas de Distorção 33
4.2 Otimização Taxa-Distorção 35

5 TRABALHOS RELACIONADOS 38
5.1 Sun et al. (SUN, 2008) 38
5.2 Fengqin et al. (FENGQIN e YANGYU, 2008) 38
5.3 Kim et al. (KIM, 2004) 39
5.4 Lee et al. (LEE, 2009) 39
5.5 Jeon et al. (JEON, 2009) 40
5.6 Wu et al. (WU, 2005) 40
5.7 Shen et al. (SHEN, 2008) 41

6 HEURÍSTICAS PARA A DECISÃO RÁPIDA DE MODO DE CODIFICAÇÃO ... 42
6.1 Decisão de Modo em Quadros I e em Quadros P 42
6.2 Decisão de Modo Rápida para Macroblocos Intra-Quadro 44
6.2.1 Heurística Baseada em Distorção (SAD) para a Decisão entre Mesmos Tamanhos de Bloco 45
6.2.2 Heurísticas para a Decisão entre Tamanhos de Bloco Diferentes 46
6.2.3 Decisões Intra-Quadro e Análise de Complexidade 49
6.3 Decisão de Modo Rápida para Macroblocos Inter-Quadros 50
6.3.1 Decisão 1: Detecção de Modo *SKIP* Baseada em Estacionariedade 52
6.3.2 Decisão 2: Detecção de Sub-Particionamento Baseada em Heterogeneidade do Macrobloco 55
6.3.3 Decisão 3: Tamanho do Particionamento Baseado em Força de Borda 58
6.3.4 Decisão 4: Tamanho do Sub-Particionamento Baseado em Força de Borda 61
6.3.5 Decisões para Macroblocos Inter-Quadros Integradas e Análise de Complexidade 65
6.4 Comparação com Trabalhos Relacionados..... 66

7 ARQUITETURAS PARA DECISÃO RÁPIDA DE MODO DE CODIFICAÇÃO .. 69
7.1 Arquiteturas para Decisão de Modo Intra-Quadro 69
7.1.1 Arquitetura Para o Cálculo de Distorção Baseada em SAD 69
7.1.2 Arquitetura Para o Cálculo da Métrica de Heterogeneidade 70
7.1.3 Arquitetura de Decisão Intra-Quadro Baseada em Heterogeneidade 72
7.1.4 Arquitetura de Decisão Intra-Quadro Baseada em Diferença de Distorção ... 74
7.2 Arquitetura para Decisão de Modo Inter-Quadros 77
7.2.1 Arquiteturas Para Decisão de Tamanho de Partição e Sub-Partição 77
7.2.2 Arquiteturas Integradas Para a Decisão de Modo Inter-Quadros 79

8 CONCLUSÕES E TRABALHOS FUTUROS..... 82

APENDICE A – Publicações e prêmios durante a graduação 88

INTRODUÇÃO

A codificação de vídeo é de extrema importância nos dias atuais, por conta da grande quantidade de informação presente em vídeos de alta resolução e a quantidade de aparelhos eletrônicos que manipulam este tipo de mídia, como computador pessoal e portátil, aparelho celular, televisão digital de alta resolução, *Blu-ray* players, câmeras e filmadoras digitais portáteis, entre muitos outros.

Além disso, é importante destacar a transmissão de televisão com sinal digital, que está se tornando uma realidade no Brasil com o Sistema Brasileiro de Televisão Digital (SBTVD, 2003), que tem como objetivo viabilizar a recepção de canais transmitidos em alta resolução.

A codificação de vídeo é essencial para o sucesso das aplicações que manipulam vídeos digitais, pois um vídeo não comprimido utiliza uma grande quantidade de bits para ser representado. Isso implica em custos muito elevados de armazenamento e transmissão destas informações, que acabam dificultando o desenvolvimento de produtos para esta área, caso a codificação não seja utilizada.

Apesar da grande quantidade de informação presente nas sequências de vídeos digitais, estas sequências possuem, em geral, uma importante propriedade: apresentam elevado grau de redundância de informação. As técnicas de codificação de vídeo buscam, portanto, a máxima eliminação destas redundâncias para que o vídeo original possa ser representado com a menor quantidade de informação possível.

Vários conjuntos de técnicas e algoritmos vêm sendo agregados para a criação de diversos padrões de codificação de vídeo, com o objetivo de padronizar a transmissão deste tipo de mídia e, consequentemente, aumentar a exploração comercial de produtos relacionados.

Neste contexto, o H.264/AVC (JVT, 2003), que é foco deste trabalho, surge como o mais novo padrão de codificação de vídeo e foi desenvolvido com o objetivo

de dobrar a taxa de compressão quando comparado com padrões anteriores, como o MPEG-2 (RICHARDSON, 2003).

Para atingir tal objetivo, o padrão H.264/AVC reuniu várias ferramentas eficientes de codificação, aumentando consideravelmente a complexidade computacional dos codificadores em relação aos padrões anteriores. Este aumento de complexidade faz com que sejam difíceis, com a tecnologia atual, implementações em software que atinjam processamento em tempo real (24 a 30 quadros por segundo), impulsionando, assim, a academia e a indústria a buscarem soluções em hardware para os módulos do codificador H.264/AVC.

O objetivo deste trabalho é desenvolver heurísticas rápidas para o módulo de decisão de modo do padrão H.264/AVC diminuindo a complexidade computacional do processo de decisão de modo. Além disso, também é objetivo implementar em hardware as heurísticas desenvolvidas com o objetivo de gerar arquiteturas que possibilitem a codificação de vídeos de alta resolução em tempo real (24 a 30 quadros por segundo).

O trabalho está organizado como segue: o capítulo 0 apresenta os princípios básicos de codificação de vídeo e o padrão H.264/AVC. O capítulo 0 mostra os dois tipos de predição presentes em codificadores H.264/AVC. O capítulo 0 mostra os conceitos básicos para a compreensão do funcionamento do módulo de decisão de modo no padrão H.264/AVC. O capítulo 0 mostra alguns trabalhos da literatura que também têm por objetivo diminuir a complexidade computacional do processo de decisão de modo do padrão H.264/AVC. O capítulo 0 apresenta as heurísticas rápidas propostas, tanto para a decisão intra-quadros quanto para a decisão inter-quadros, bem como os resultados gerados pelas mesmas. O capítulo 0 apresenta as arquiteturas desenvolvidas para cada heurística. Finalmente, o capítulo 0 apresenta as conclusões e alguns trabalhos futuros.

CODIFICAÇÃO DE VÍDEO E O PADRÃO H.264/AVC

Neste capítulo serão apresentados os princípios básicos de codificação de vídeo que serão importantes para contextualização deste trabalho. Além disso, o capítulo irá apresentar o padrão H.264/AVC, foco deste trabalho.

Características dos Vídeos Digitais

Um vídeo é formado por uma sequência de imagens estáticas (quadros) que são exibidas em uma determinada frequência (normalmente 24 a 30 quadros por segundo), dando impressão de movimento (RICHARDSON, 2003). Cada quadro é formado por uma matriz de blocos, os quais são formados por uma matriz de pontos chamados pixels. Quanto maior a quantidade de pixels utilizados para representar um quadro, maior é a resolução do vídeo, o que proporciona uma maior facilidade na percepção dos detalhes de uma determinada cena.

Vídeos digitais são representados por espaços de cores que têm o objetivo de separar os quadros dos vídeos em componentes de cores. Diversos espaços de cores, como RGB e YCbCr (RICHARDSON, 2003), vêm sendo utilizados para representar imagens digitais. O RGB, bastante conhecido e utilizado, divide as informações de cores em três componentes: R (*red* – vermelho), G (*green* – verde) e B (*blue* – azul). O YCbCr também é composto por três componentes: o componente de luminância (Y), que representa o nível de luminosidade na imagem, e os componentes de crominância azul e vermelha (Cb e Cr), que representam o desvio de uma cor a partir do cinza em direção ao azul ou ao vermelho, respectivamente (RICHARDSON, 2003).

O sistema YCbCr foi desenvolvido buscando tratar a informação de luminosidade de forma independente da informação de cor, já que o sistema visual humano percebe melhor as mudanças na intensidade de brilho do que as mudanças na intensidade de cor. Assim, os dois tipos de informação podem ser manipulados separadamente pelos codificadores de imagem e de vídeo. No H.264/AVC, por exemplo, o sistema YCbCr permite representar os componentes de crominância usando me-

nos amostras em relação aos componentes de luminância (RICHARDSON, 2003). Esta técnica é chamada de sub-amostragem de cores e é muito utilizada nos padrões de codificação de vídeo atuais.

A eficiência da codificação aumenta consideravelmente quando a técnica de sub-amostragem é utilizada, pois parte dos dados é simplesmente descartada sem causar nenhum impacto perceptível ao olho humano. Existem diversas relações de sub-amostragem entre os componentes de luminância e crominância, sendo mais comuns os formatos 4:4:4, 4:2:2 e 4:2:0. O 4:2:0 é o formato mais utilizado em codificadores H.264/AVC. Neste caso, para cada quatro amostras de luminância tem-se uma amostra de crominância azul e uma de crominância vermelha. O formato deveria ser chamado, portanto, de 4:1:1, mas, por motivos históricos, esta não é a nomenclatura utilizada (RICHARDSON, 2003).

Redundância de Dados na Representação de Vídeos

O principal objetivo das técnicas de codificação de vídeo é diminuir a quantidade de informação redundante presente nas sequências de vídeos digitais. O dado redundante é aquele que não contribui com novas informações para a representação do vídeo. Basicamente, existem três tipos de redundância que são exploradas por algoritmos de codificação de vídeo: redundância espacial, redundância temporal e redundância entrópica.

- Redundância Espacial – A redundância espacial, também chamada de redundância intra-quadro (WIEGAND, SULLIVAN, *et al.*, 2003) advém da correlação existente entre pixels em um quadro. Esta correlação é facilmente visualizada quando são observados pixels vizinhos em um mesmo quadro, os quais tendem a possuir valores semelhantes. Neste caso, a redundância pode ser diminuída através da codificação intra-quadro, presente em alguns padrões de codificação de vídeo atuais. No domínio das frequências, a operação utilizada para reduzir este tipo de redundância é a quantização. A quantização é uma divisão inteira aplicada sobre cada coeficiente gerado pela etapa de transformação e reduz grande parte dos coeficientes a zero.
- Redundância Temporal – A redundância temporal, também chamada de redundância inter-quadros (WIEGAND, SULLIVAN, *et al.*, 2003), existe devido à correlação entre quadros temporalmente próximos. Co-

mo um vídeo é uma sequência de imagens estáticas, e são necessários de 24 a 30 quadros por segundo para se ter a noção de movimento, muitos blocos de pixels simplesmente não mudam seu valor de um quadro para o outro, enquanto outros blocos apresentam uma pequena variação ocasionada, por exemplo, por uma mudança de iluminação. Ainda é possível que o bloco de pixels tenha se deslocado em relação ao quadro anterior, como por exemplo, o movimento de um objeto em uma cena. A eficiente redução da redundância temporal conduz a grandes taxas de compressão.

- Redundância Entrópica – A redundância entrópica refere-se à quantidade de símbolos semelhantes que são transmitidos em um vídeo codificado. As técnicas e algoritmos utilizados para diminuir este tipo de redundância têm como principal objetivo transmitir a maior quantidade de informação possível por símbolo.

O Padrão H.264/AVC

O padrão H.264/AVC, que é o foco deste trabalho, é o resultado de uma iniciativa de padronização colaborativa entre o ITU-T VCEG e o ISO/IEC MPEG (HANZO, CHERRIMAN e STREIT, 2007). O codificador H.264/AVC oferece um aumento significativo na eficiência de compressão, sem perda de qualidade quando comparado ao MPEG-4 (ISO/IEC, 1998) ou H.263 (ITU-T, 1996). Para atingir esse aumento na eficiência de compressão, o padrão H.264/AVC introduziu diversos recursos no processo de codificação, tais como o uso de tamanhos de bloco menores que 8x8 pixels, a predição intra-quadro realizada no domínio espacial, a transformada Hadamard, a utilização de múltiplos quadros de referência para a predição inter-quadros, entre outros. O uso dessas novas ferramentas proporcionou um aumento significativo nas taxas de compressão, consequentemente, ocasionando um aumento na complexidade computacional dos codificadores.

As técnicas de codificação empregadas no codificador H.264/AVC são similares às técnicas empregadas com sucesso em padrões anteriores. Elementos básicos como predição, transformadas, quantização e codificador de entropia, encontrados em padrões tais como MPEG-1 (ISO/IEC, 1993), MPEG-2 (ISO/IEC, 1995), MPEG-4 (ISO/IEC, 1998), H.261 (ITU-T, 1993) e H.263 (ITU-T, 1996) são facilmente identificados.

O padrão H.264/AVC define alguns perfis de codificação (*Baseline*, M*ain*, *Extended*) que são escolhidos dependendo da aplicação alvo. O *Baseline* é um perfil voltado para aplicações móveis que exijam melhor desempenho, já que este perfil possui o menor subconjunto de ferramentas de codificação definidos no padrão. O perfil *Main* inclui alguns recursos a mais que o *Baseline*, como múltiplos quadros de referência na predição inter-quadros e codificação de entropia utilizando codificação aritmética adaptativa ao contexto (CABAC). Já o perfil *Extended,* que é mais voltado para *streaming* de vídeo, agrega modos para habilitar uma troca eficiente entre *bitstreams* codificados (através de *slices* do tipo SP e SI) e melhora a robustez a erros de transmissão (através do particionamento de dados) (AGOSTINI, 2007).

O espaço de cores utilizado nos três perfis é o YCbCr com subamostragem 4:2:0, como já mencionado na seção 0. Isto significa que para quatro elementos de luminância existe um elemento de crominância azul e um elemento de crominância vermelha associados. A unidade de trabalho utilizada nos codificador H.264/AVC é o macrobloco. Um macrobloco é formado por um bloco de 16x16 amostras de luminância (Y), um bloco de 8x8 amostras de crominância azul (Cb) e um bloco 8x8 de crominância vermelha (Cr). Nos três perfis definidos no padrão, são utilizados 8 bits para a representação de cada amostra. Existem outros perfis que foram definidos para dar suporte a vídeos profissionais de qualidade elevada, já que os perfis *Baseline, Main* e *Extended* foram criados para entretenimento.

No padrão H.264/AVC um grupo de macroblocos é chamado de *slice*. Existem 3 tipos básicos de *slices*: (1) tipo I, que contém macroblocos codificados somente com a predição intra-quadro, (2) tipo P, onde os macroblocos podem ser codificados tanto pela predição inter-quadros quanto pela predição intra-quadro e ainda podem ser *skipped macroblocks* (RICHARDSON, 2003) e (3) tipo B, onde são usados mais de um quadro de referência para realizar a predição inter-quadros de cada macrobloco, o que é chamado de bi-predição.

O diagrama em blocos de um codificador H.264/AVC é mostrado na Fig. 1. O codificador consiste de dois caminhos de dados, um caminho direto de codificação orientado da esquerda para a direita e um caminho de reconstrução orientado da direita para a esquerda.

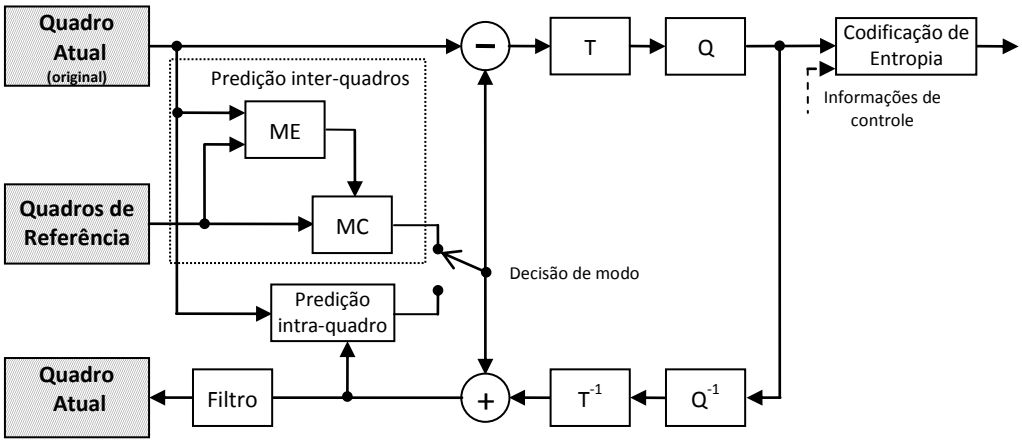

Figura 1 - Diagrama em blocos do codificador H.264/AVC.

O padrão H.264/AVC utiliza a codificação residual, que é a técnica que apresenta os melhores resultados em termos de compressão (SULLIVAN e WIEGAND, 1998). O fluxo de dados em codificadores H.264/AVC é basicamente o seguinte: primeiramente são executadas as etapas de predição, que podem ser intra-quadro ou inter-quadros. Nesta etapa, o bloco atual é codificado usando blocos de referência que podem ser do mesmo quadro, no caso da predição intra-quadro, ou de quadros vizinhos, no caso da predição inter-quadros. A escolha entre os dois tipos de predição é realizada pela chave seletora representada na Fig. 1. Em seguida, é realizada uma subtração do bloco atual pelo bloco predito escolhido. O resultado desta subtração é chamado de bloco de resíduos, o qual é processado pelas transformadas diretas (módulo T) e pela quantização direta (módulo Q), gerando os coeficientes quantizados que são enviados junto com outras informações (dados de controle, modos de predição e vetores de movimento utilizados) para a etapa de codificação de entropia. A codificação de entropia gera uma sequência de *bits* chamada *bitstream* que forma o vídeo codificado.

A predição inter-quadros é formada pelos módulos de estimação de movimento (ME) e compensação de movimento (MC). A predição inter-quadros é responsável por reduzir a redundância temporal em um vídeo, utilizando blocos de quadros vizinhos no tempo para representar o bloco que está sendo codificado. A estimação de movimento é responsável por buscar, no quadro de referência, o bloco mais semelhante ao bloco atual. Quando este bloco é encontrado, é gerado um vetor de movimento que indica a posição deste bloco mais semelhante em relação à posição do bloco atual.

A predição intra-quadro é responsável por reduzir a redundância espacial em um mesmo quadro. O bloco atual é representado utilizando pixels de blocos vizinhos previamente codificados. O capitulo 0 explicará com mais detalhes, tanto a predição inter-quadros quanto a predição intra-quadro.

Os módulos de quantização inversa (Q^{-1}) e de transformadas inversas (T^{-1}) são responsáveis pela reconstrução dos blocos que serão usados como referência nas próximas predições. Este caminho de reconstrução, que está presente também no decodificador, é necessário, pois o módulo de quantização gera perdas. Isso significa que as referências utilizadas no processo de decodificação serão diferentes dos blocos originais. Sendo assim, o objetivo deste caminho de reconstrução é garantir que as referências usadas no processo de codificação sejam as mesmas usadas no processo de decodificação, evitando erros de casamento (em inglês *matching*).

PREDIÇÃO INTER-QUADROS E INTRA-QUADRO NO PADRÃO H.264/AVC

As etapas de predição do padrão H.264/AVC são realizadas usando, como referência, informações previamente codificadas. Amostras dentro de um macrobloco do tipo I são preditas a partir de amostras do mesmo quadro que já foram codificadas, decodificadas e reconstruídas. A predição inter sobre um macrobloco é realizada a partir de informações de quadros previamente codificados, decodificados e reconstruídos (RICHARDSON, 2003). Neste capitulo, serão detalhados os processos de predições intra-quadro e inter-quadros no padrão H.264/AVC.

Predição Inter-Quadros

A predição inter-quadros é responsável por eliminar a redundância temporal que existe entre sucessivos quadros, consequentemente, provendo uma codificação eficiente sobre sequências de vídeo (HANZO, CHERRIMAN e STREIT, 2007). O módulo de predição inter-quadros, no padrão H.264/AVC, é composto por dois módulos: o módulo de estimação de movimento (ME) e o módulo de compensação de movimento (MC).

O módulo de estimação de movimento está presente apenas nos codificadores. Este módulo é o que apresenta maior complexidade computacional entre todos os módulos do codificador H.264/AVC (PURI, CHENB e LUTHRA, 2004). Este grande custo computacional é função das inovações inseridas neste módulo pelo padrão H.264/AVC. A ME e a MC são as maiores responsáveis pelos ganhos em compressão do padrão em relação aos demais padrões de codificação de vídeo (RICHARDSON, 2003).

O processo de estimação de movimento sobre um bloco é realizado da seguinte maneira: inicialmente, uma busca é realizada em um determinado quadro de referência com o objetivo de encontrar o bloco mais similar ao bloco atual. Então,

quando este é encontrado, um vetor de movimento é gerado a partir da posição do bloco atual até a posição do bloco mais similar no quadro de referência.

Existem vários critérios para definir o quanto dois blocos são similares. O SAD (*Sum of Absolute Differences*) é o critério de similaridade mais conhecido e mais simples. O cálculo do SAD é realizado da seguinte forma: uma subtração pixel a pixel é realizada entre o bloco atual e o bloco de referência no momento da realização da busca. Logo depois, um somatório do valor absoluto dessas diferenças é realizado, gerando o valor de SAD. O bloco que possuir o menor SAD é considerado o mais similar. Existem outros critérios de similaridade conhecidos, como SATD (*Sum of Absolute Transformed Differences*), onde é aplicada uma transformada após a realização das diferenças pixel a pixel entre os blocos para depois ser realizado o somatório, e o SSD (*Sum of Squared Differences*), que aplica uma potência de dois sobre as diferenças antes de realizar o somatório. A escolha do critério de similaridade é muito importante e afeta diretamente o resultado em termos de qualidade e quantidade de bits necessários para representar o vídeo codificado.

A busca pelo bloco mais similar no quadro de referência pode ser realizada de várias maneiras. Algoritmos de busca para a estimação de movimento são muito estudados na academia e grande parte da pesquisa realizada sobre algoritmos de codificação de vídeo são relacionadas a algoritmos de busca para a estimação de movimento.

A estimação de movimento é aplicada apenas sobre amostras de luminância de um macrobloco. Como os componentes de crominância possuem metade da resolução horizontal e vertical do componente de luminância, o vetor de movimento para os blocos de crominância são gerados a partir de uma divisão escalar por dois sobre o vetor do bloco de luminância.

A principal inovação do H.264/AVC no ponto de vista da estimação de movimento está na possibilidade de utilização de vários tamanhos de bloco para realizar a codificação sobre um macrobloco. Ao invés de usar um macrobloco inteiro na estimação de movimento, o padrão H.264/AVC permite o uso de partições e subpartições de macroblocos. As partições de macrobloco possuem os tamanhos de 16x16, 16x8, 8x16 e 8x8, como está apresentado na Fig. 2 (WIEGAND, SULLIVAN, *et al.*, 2003).

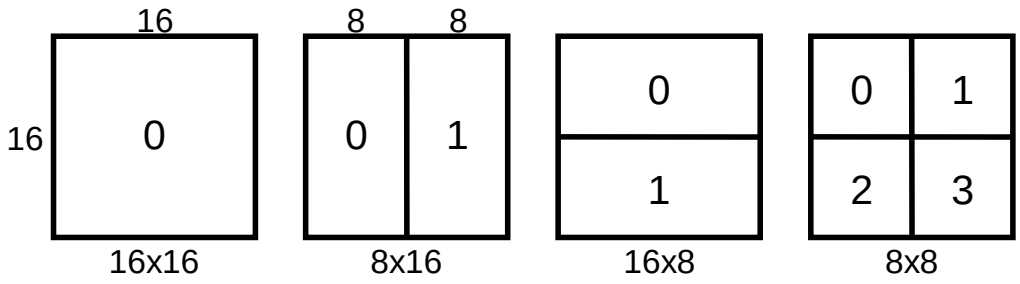

Figura 2 - Partições de macrobloco definidas no padrão H.264/AVC.

Quando o tamanho da partição escolhida é do tipo 8x8, ainda é possível dividir o macrobloco em sub-partições. As sub-partições possuem os tamanhos de 8x8, 8x4, 4x8 e 4x4, como mostrado na Fig. 3.

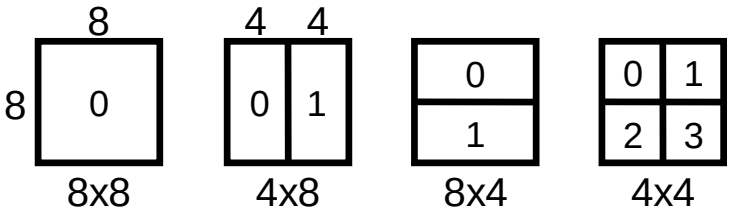

Figura 3 - Sub-partições de macrobloco definidas no padrão H.264/AVC.

Os blocos de crominância são divididos do mesmo modo que os blocos de luminância, mas com metade da resolução horizontal e vertical. Por exemplo, se o tamanho de bloco de luminância utilizado for 16x8 então o tamanho do bloco de crominância será de 8x4.

Após a etapa de estimação de movimento, é realizada a compensação de movimento (MC). O objetivo do módulo de compensação de movimento é gerar os resíduos que serão enviados para os próximos módulos no caminho de codificação. A compensação de movimento recebe os vetores gerados pelo módulo de estimação de movimento e busca, no quadro de referência, o bloco apontado por este vetor. Então, o bloco atual é subtraído do bloco apontado pelo vetor, gerando o resíduo.

A informação residual e o vetor de movimento são as informações necessárias para que o decodificador reconstrua o bloco. O bloco de compensação de movimento está presente, também, no decodificador.

Para cada partição ou sub-partição de macrobloco é necessário um vetor de movimento e cada vetor de movimento gerado é codificado e transmitido no *bitstre-*

am. A escolha de partições maiores (16x16, 16x8, 8x16 e 8x8) significa um número menor de bits para representar os vetores, já que menos vetores são utilizados. Entretanto, a escolha de partições maiores pode acabar gerando uma maior quantidade de resíduo após a compensação de movimento, principalmente, se utilizada em áreas com grande quantidade de detalhes. A escolha de partições menores (8x4, 4x8 e 4x4) acaba gerando uma quantidade menor de resíduo após a compensação de movimento, mas a necessidade de bits para representar os vetores será maior, pois serão necessários mais vetores.

Considerando as várias possibilidades de particionamento, a escolha do tamanho da partição utilizada na estimação de movimento tem um impacto significativo no desempenho de compressão do codificador. Em geral, partições maiores são apropriadas para áreas mais homogêneas do quadro, enquanto partições menores podem beneficiar áreas com muitos detalhes (RICHARDSON, 2003).

Predição Intra-Quadro

A predição intra-quadro é responsável por eliminar a redundância espacial dentro de um mesmo quadro. A predição intra-quadro é aplicada tanto para amostras de luminância quanto para amostras de crominância. A predição de um macrobloco é realizada a partir da cópia de amostras vizinhas previamente codificadas e reconstruídas.

Existem dois tipos possíveis de predição intra-quadro, no caso das amostras de luminância: intra 4x4 (I4MB) e intra 16x16 (I16MB). No tipo I4MB, o bloco 16x16 de amostras de luminância é dividido em 16 blocos de tamanho 4x4 e, então, cada um desses blocos é predito separadamente. No tipo I16MB, o bloco 16x16 de amostras de luminância é predito por inteiro. A Fig. 4 mostra as divisões de amostras de luminância utilizada nos dois tipos possíveis de predição. Os componentes de crominância são codificados sempre com o mesmo tamanho (8x8), quando a subamostragem 4:2:0 é usada.

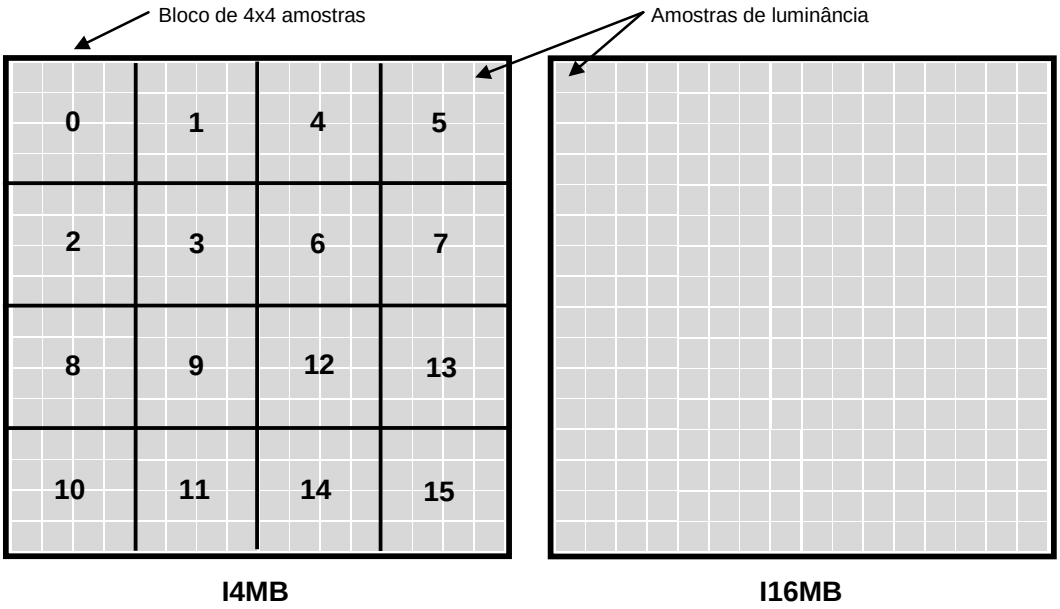

Figura 4 - Tamanhos de bloco utilizados na predição intra-quadro do padrão H.264/AVC para amostras de luminância (DINIZ, 2009).

1.1.1 Intra 4x4 (I4MB)

Considerando o modo I4MB, existem nove modos possíveis para gerar a predição de um bloco 4x4. A Fig. 5 ilustra as amostras vizinhas que podem ser utilizadas para gerar a predição de um bloco 4x4. As amostras de **A** a **M** na Fig. 5 foram codificadas e reconstruídas antes deste bloco ser processado. As amostras de **a** a **p** representam o bloco 4x4 que será predito.

M	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				

Figura 5 - Amostras usadas na predição intra-quadro considerando macroblocos I4MB.

A Fig. 6 mostra os nove modos de predição possíveis para macroblocos I4MB. O modo 0 faz uma simples cópia das amostras que compõem a borda superior (A, B, C e D) para todas as posições do bloco. Da mesma forma, o modo 1 faz a cópia das amostras que compõem a borda lateral esquerda (I, J, K e L). O modo 2

(DC) calcula a média aritmética entre as amostras de A à D e I à L e copia para todas as posições do bloco. Os modos 3, 4, 5, 6, 7 e 8 realizam uma interpolação linear das amostras vizinhas conforme a direção das setas apresentadas na Fig. 6 e copiam estes valores para o bloco o bloco predito (DINIZ, 2009).

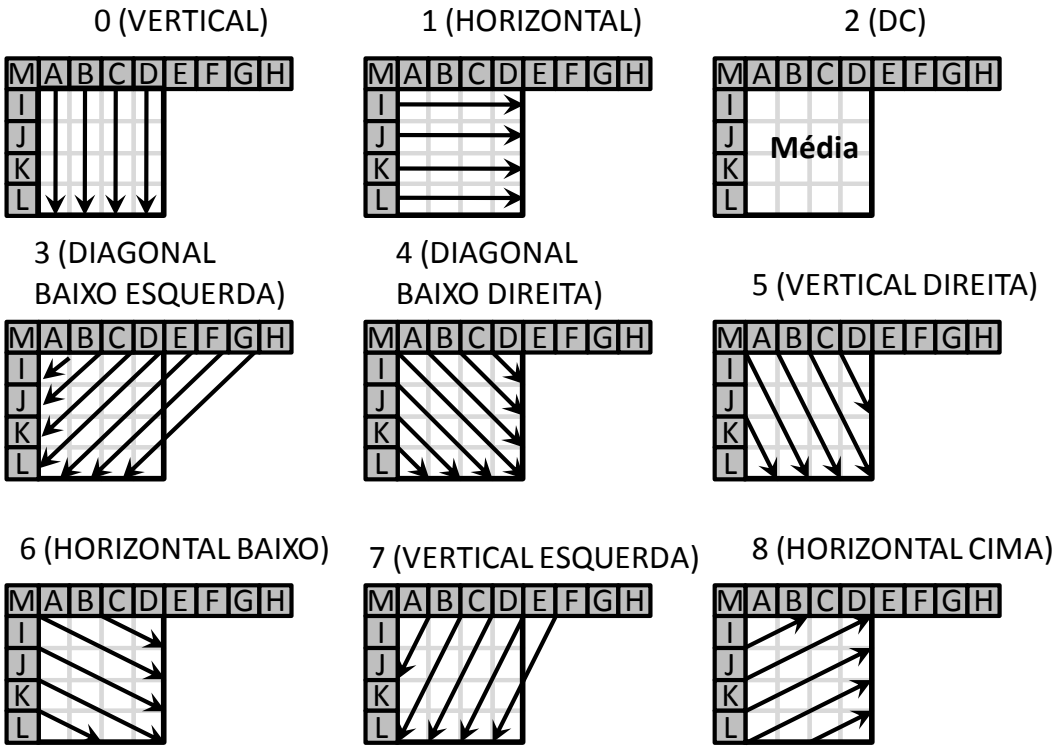

Figura 6 - Modos de predição intra-quadro para I4MB.

1.1.2 Intra 16x16 (I16MB)

Quando a predição realizada é do tipo I16MB, ou seja, sobre blocos de tamanho 16x16, existem quatro modos de predição que podem ser utilizados. Todos os modos utilizam as 32 amostras de luminância vizinhas ao bloco predito (H e V na Fig. 7). A Fig. 7 ilustra os modos de cópia possíveis para o modo I16MB.

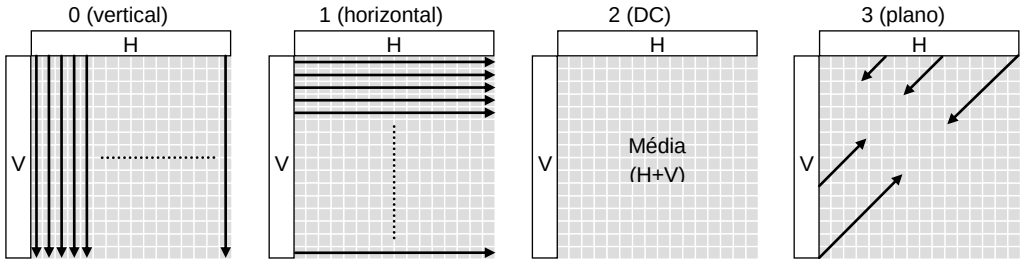

Figura 7 - Modos de predição intra-quadro para I16MB.

Nos modos 0 e 1, as amostras de H e V são copiadas no sentido vertical e horizontal, respectivamente, para todo o bloco predito, assim como na predição I4MB. No modo 2, as amostras preditas são calculadas através de uma média simples entre as amostras H e V, conforme (1).

$$DC = (\sum_{x'=0}^{15} p[x',-1] + \sum_{y'=0}^{15} p[-1,y'] + 16) >> 5 \quad (1)$$

Em (1), **p[y,x]** é a amostra predita localizada na posição indicada pela linha **y** e pela coluna **x** do bloco. A amostra [0,0] localiza-se no canto superior esquerdo, de forma que, quando o índice é igual a -1, a amostra está no bloco vizinho (esquerdo ou superior), já codificado e reconstruído. O valor resultante (**DC**) é o valor que vai representar todas as amostras do bloco predito.

O modo 3 (plano) é o mais complexo, pois seu cálculo envolve uma função linear que utiliza três parâmetros (**a**, **b** e **c**) pré-calculados para cada amostra do bloco predito, como definido em (2).

$$P[y,x] = Clip((a + b\cdot(x-7) + c\cdot(y-7) + 16) >> 5) \quad (2)$$

Em (2), **P[y,x]** é a amostra predita, ***Clip*** é uma função de truncagem (ou arredondamento) para limitar os valores na faixa de representação da imagem e os parâmetros **a**, **b** e **c** são definidos em (3), (4) e (5), respectivamente. Os valores de H e V estão definidos em (6) e (7), respectivamente.

$$a = 16\cdot(p[-1,15] + p[15,-1]) \quad (3)$$

$$b = (5\cdot H + 32) >> 6 \quad (4)$$

$$c = (5\cdot V + 32) >> 6 \quad (5)$$

$$H = \sum_{x'=0}^{7} (x'+1)\cdot(p[-1,8+x'] - p[-1,6-x']) \quad (6)$$

$$V = \sum_{y'=0}^{7} (y'+1)\cdot(p[8+y',-1] - p[6-y',-1]) \quad (7)$$

1.1.3 Predição para Crominância

As amostras de crominância são preditas de forma similar as amostras de luminância codificadas no modo I16MB, com exceção do número de pixels envolvidos. Também são usados quatros modos com os mesmo nomes, mas com numeração trocada em relação à predição I16MB, como mostra a Fig. 8.

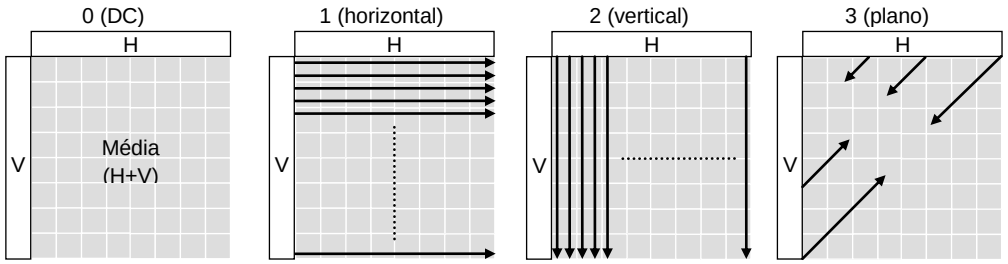

Figura 8 - Modos de predição intra-quadro para crominância 8x8.

Considerando a amostragem 4:2:0, os blocos de crominância possuem o tamanho de 8x8. Desta forma existem oito amostras vizinhas à esquerda e acima que são consideradas para realizar a predição para blocos de crominância.

MODO DE DECISÃO

O padrão H.264/AVC incorpora muitas técnicas do estado-da-arte em codificação de vídeo com o objetivo de atingir altas taxas de compressão. Estas técnicas estão inseridas, principalmente, nos módulos de predição, como por exemplo, os vários tamanhos de bloco possíveis de serem utilizados na predição inter-quadros. Considerando todos os modos em que um macrobloco pode ser predito, fica a cargo do módulo de decisão de modo escolher qual modo é o mais adequado.

O RDO (*Rate-Distortion Optimization* – Otimização Taxa-Distorção) (SULLIVAN e WIEGAND, 1998) é a técnica de decisão de modo que gera os melhores resultados para a relação entre qualidade e quantidade de bits do vídeo codificado. Entretanto, o uso desta técnica causa um grande aumento na complexidade computacional dos codificadores H.264/AVC. Sendo assim, é importante o uso de técnicas para acelerar a decisão de modo de predição, acelerando o processo de codificação como um todo.

Este capítulo apresentará conceitos básicos para a compreensão do papel do modo de decisão em codificadores H.264/AVC.

Métricas de Distorção

A técnica RDO é utilizada para encontrar o modo de codificação que apresenta o menor custo taxa-distorção (RD), sendo assim, é necessário definir uma métrica para o cálculo de distorção do vídeo codificado. Devido à elevada complexidade do sistema visual humano, é muito complicado medir a quantidade de distorção percebida em um conteúdo visual. Mesmo assim, métricas como SSD (*Sum of Squared Differences*), MSE (*Mean-Squared Error*) e PSNR (*Peak Signal-to-Noise Ratio*), são utilizadas na maioria das comparações objetivas (RICHARDSON, 2003). Tais métricas se baseiam na diferença entre os pixels da imagem original e da imagem após a decodificação. Existem, também, métricas subjetivas para avaliação de qualidade (RICHARDSON, 2003), entretanto, elas não serão abordadas neste trabalho.

O critério mais conhecido e mais utilizado pela comunidade é o PSNR (GHANBARI, 2003), definido na equação (8).

$$PSNR_{dB} = 20 \cdot \log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \tag{8}$$

Em (8), *MAX* é o máximo valor de representação de uma amostra (2^n – 1, sendo *n* o número de bits para representar cada amostra), e MSE é o erro médio quadrático (*Mean Squared Error*), que está definido em (9), onde *mn* é a dimensão da imagem em *pixels, O* e *R* são a imagem original e reconstruída, respectivamente.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left(R_{i,j} - O_{i,j}\right)^2 \tag{9}$$

O SAD (*Sum of Absolute Differences*), além de ser usado como critério de similaridade nas predições (como explicado na seção 0), também é usado como métrica para avaliar distorção entre duas imagens. O cálculo do SAD é realizado a partir do somatório das diferenças absolutas entre pixels da imagem original e da imagem reconstruída (RICHARDSON, 2003). A equação do SAD é mostrada em (10), onde *m* e *n* são as dimensões da imagem a ser comparada e *O* e *R* são as amostras do quadro original e reconstruído, respectivamente. Este critério é amplamente usado, devido à sua baixa complexidade computacional em relação ao MSE (RICHARDSON, 2003).

$$SAD = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|R_{i,j} - O_{i,j}\right| \tag{10}$$

Estas medidas, em geral, são usadas somente sobre o componente de luminância de uma imagem durante o processo de otimização, mas um melhor desempenho pode ser obtido aplicando-as sobre os três componentes de cor.

Otimização Taxa-Distorção

Existe uma relação direta entre a quantidade de bits utilizada no vídeo codificado (taxa) e a sua distorção. Em geral, quanto maior a taxa de bits, melhor será a qualidade do vídeo representado (e vice-versa). Caso o codificador esteja configurado para utilizar um menor número de bits no vídeo codificado, isso resultará em um vídeo muito distorcido; caso o objetivo seja obter um vídeo codificado com uma qualidade melhor, mais bits serão necessários para representá-lo. A Fig. 9 mostra a curva típica que relaciona a taxa de bits (em Kbps) e a distorção do vídeo (medida com PSNR em decibéis).

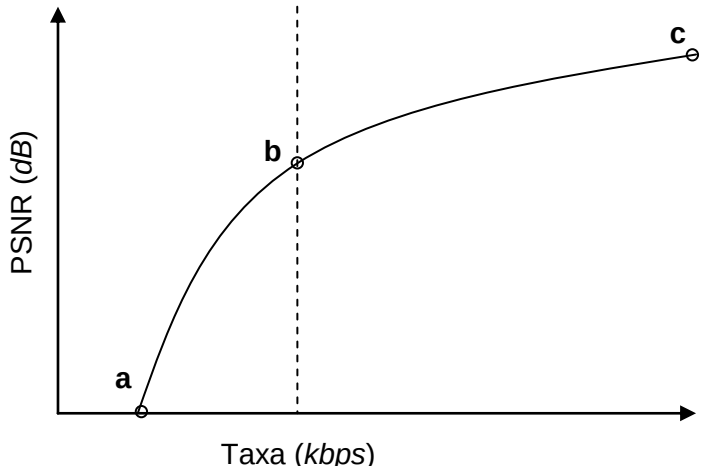

Figura 9 - Curva típica da relação entre taxa de bits e distorção.

No gráfico apresentado na Fig. 9, pode-se notar que do ponto **b** ao ponto **a** há um grande aumento na distorção do vídeo e uma pequena diminuição na taxa de bits utilizados para representá-lo, o que é bastante desvantajoso. Por outro lado, do ponto **b** ao ponto **c** há um grande aumento na taxa de bits, e ao mesmo tempo, uma pequena diminuição na distorção do vídeo codificado, o que também não é vantajoso. O ponto **b**, então, se mostra o ideal se considerarmos a relação entre taxa de bits e distorção.

O papel da técnica RDO é justamente escolher, para cada macrobloco a ser codificado, a representação (segmentação, modo de predição, vetores de movimento, níveis de codificação, etc.) mais eficiente no sentido taxa-distorção. Esta escolha é muito complicada devido às várias opções de codificação, que mostram eficiência variável quando taxas de bits diferentes são utilizadas e quando conteúdos de cenas

diferentes são codificados (SULLIVAN e WIEGAND, 1998). Por exemplo, a codificação de um macrobloco por um dos modos inter-quadros seguidos pela quantização é bastante eficiente para codificar áreas de uma imagem onde existe muito movimento. Por outro lado, a escolha de algum modo intra-quadro é mais apropriado para a codificação de áreas que não possuem movimento contínuo.

O cálculo do RDO é resolvido através da utilização da otimização de Lagrange, onde um termo de distorção é utilizado como ajuste com um termo de taxa de bits, a fim de encontrar-se o custo taxa-distorção (*RDCost*). A equação (11) mostra a função de custo *RDCost*, onde o funcional taxa-distorção *J* de Lagrange é minimizado para um determinado valor do multiplicador de Lagrange λ. O multiplicador de Lagrange está relacionado ao parâmetro de quantização e está apresentado na equação (12).

$$RD\cos t = J = D + \lambda \times R \tag{11}$$

$$\lambda = 0{,}85 \times 2^{(QP-12)/3} \tag{12}$$

Em (11), *D* representa um valor de distorção calculado a partir do bloco atual A e do bloco reconstruído A', obtido através de uma das medidas de distorção apresentadas na seção anterior. O parâmetro *R* é o número de bits utilizados para representar o bloco codificado. Após o cálculo de todos os custos taxa-distorção para todos os modos de predição, o modo com menor custo é escolhido.

A Fig. 10 mostra como é realizado o cálculo de *RDcost*. É possível perceber a elevada complexidade computacional do cálculo de *RDcost*, já que é necessário realizar o caminho completo de codificação e de reconstrução (chamado de laço de codificação) para cada modo de predição possível para cada bloco.

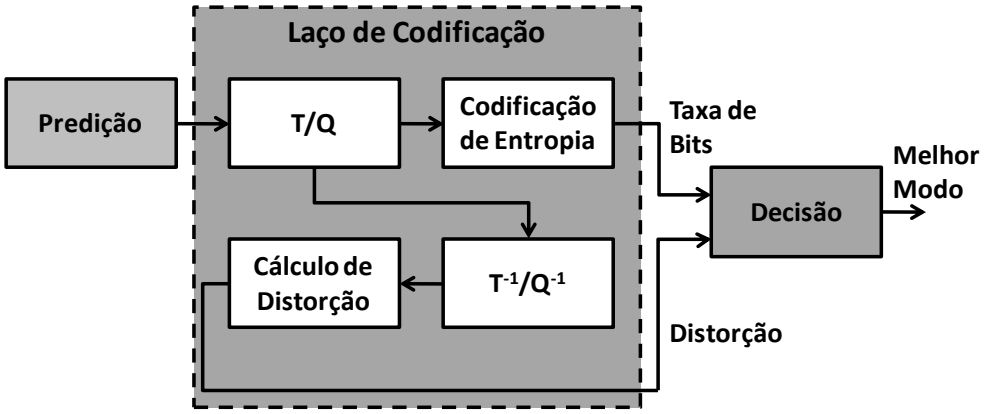

Figura 10 - Calculo do custo taxa-distorção (*RDcost*).

A técnica RDO apresenta os melhores resultados na relação taxa-distorção. Entretanto, o tempo de computação necessário para avaliar todas as possibilidades de codificação para todos os blocos é um fator limitante no desempenho dos codificadores H.264/AVC, praticamente impossibilitando, com a tecnologia atual, a codificação de vídeos de alta resolução em tempo real.

Desse modo, o principal objetivo deste trabalho é desenvolver um módulo de decisão rápido de modo, baseada somente em heurísticas, bem como as suas respectivas implementações em hardware, visando a eliminação completa do uso da técnica RDO do processo de codificação, reduzindo drasticamente o tempo necessário para codificar um vídeo.

TRABALHOS RELACIONADOS

Neste capítulo serão apresentados trabalhos encontrados na literatura, que visam reduzir a complexidade do processo de decisão de modo. É importante salientar que, apesar das heurísticas propostas nestes trabalhos apresentarem redução do custo computacional no processo de decisão, todas elas, em algum passo, ainda fazem uso da técnica RDO.

Sun et al. (SUN, 2008)

Este trabalho propõe uma decisão rápida voltada apenas para a decisão intra-quadro. Os autores utilizam a métrica de distorção SATD para determinar se o custo taxa-distorção sobre um modo deve ser calculado. Modos que apresentam um SATD alto são descartados da avaliação de melhor modo, pois têm tendência a apresentar uma alta taxa de bits.

O algoritmo proposto é dividido em três etapas. A primeira etapa calcula o SATD para todos os modos intra-quadro e seleciona os modos com menor SATD. A segunda etapa calcula a variância dos coeficientes gerados após o cálculo do SATD, e elimina aqueles modos onde os coeficientes variam muito. A última etapa realiza o processo RDO para os módulos restantes.

A decisão de modo é capaz de reduzir o número de cálculos em até 79%, com uma perda de 0,136 dB em PSNR e um aumento na taxa de bits de 0,69%.

Fengqin et al. (FENGQIN e YANGYU, 2008)

Neste trabalho, os autores propõem uma decisão rápida baseada na característica dos coeficientes transformados. Segundo os autores, a aplicação de uma transformada nos resíduos após a predição intra, gera um disposição dos coeficientes diferentes para cada modo intra-quadro. Considerando isso, a probabilidade de cada modo ser escolhido é calculada através de uma série de equações propostas

no trabalho. Cada modo é classificado de acordo com sua probabilidade e, então, modos com probabilidade baixa são descartados. Os modos não descartados são avaliados pelo processo RDO e, então, é realizada a decisão. Para evitar a difusão de erro de decisão, um método preventivo é aplicado através da inserção de quadros periódicos com decisão RDO sobre todos os modos.

A aplicação desta técnica gerou uma redução de 5,92 vezes no número de cálculos necessários para realizar a decisão em comparação com o RDO. Houve uma redução de 0,02 dB na qualidade da imagem e um aumento médio de 3% na taxa de bits do vídeo codificado.

Kim et al. (KIM, 2004)

Este trabalho utiliza a mesma técnica de (SUN e YINYI, 2008), que visa reduzir o número de modos candidatos através de uma avaliação sobre os coeficientes gerados pelo cálculo do SATD. O trabalho também propõe a utilização de uma estimativa de taxa de bits baseada na quantidade de coeficientes quantizados para zero.

Neste caso a avaliação é realizada sobre os coeficientes transformados e quantizados, sendo necessária a aplicação da transformada e quantização apresentados na Fig. 10 para cada um dos modos selecionados pela avaliação SATD.

A aplicação da técnica proposta, gerou uma redução de 0,070 dB na qualidade de imagem. O trabalho não apresenta nenhuma informação com relação à taxa de bits ou número de cálculos taxa-distorção. Entretanto, a redução em tempo de execução foi de aproximadamente 85%.

Lee et al. (LEE, 2009)

O trabalho propõe a diminuição da complexidade da decisão de modo intra-quadro através da análise dos componentes de baixa frequência de um bloco de coeficientes transformados. A proposta se baseia na análise de alguns coeficientes gerados por uma transformada DCT, para determinar o nível de homogeneidade de um bloco para decidir qual tamanho de bloco será utilizado. Contudo, o trabalho realiza somente a decisão entre tamanhos de bloco diferentes, sendo que a decisão entre mesmos tamanhos de bloco é realizada pela técnica RDO. Para reduzir o número de modos analisados, o trabalho propõe a utilização do valor do SATD de um bloco para calcular a probabilidade da escolha de cada modo.

A redução no número de cálculos ficou, em média, em torno de 25%, com a qualidade da imagem sofrendo uma redução de 0,088 dB, enquanto que a taxa de bits apresentou uma aumento de 0,09% em relação ao uso do RDO.

Jeon et al. (JEON, 2009)

Este trabalho se baseia no fato de que pixels que compõem uma mesma borda local possuem valores semelhantes. O algoritmo proposto realiza a decisão de melhor modo intra-quadro de acordo com os gradientes direcionais do bloco de acordo com cada modo. Os gradientes são calculados com base em uma série de equações apresentadas no trabalho. Os valores de gradiente são, então, comparados com valores limite pré-calculados através de simulações. Os dois modos com menor gradiente são selecionados como melhores e, junto com o modo DC, são processados pelo RDO para realizar a escolha.

A utilização desta técnica reduz o tempo de computação em até 77%, com uma perda de 0,148 dB na qualidade da imagem e um aumento de 4% na taxa de bits.

Wu et al. (WU, 2005)

O trabalho apresenta um conjunto de heurísticas para a decisão de modo completa, ou seja, intra-quadro e inter-quadros. Os autores observam que as regiões homogêneas e estáticas, em geral, são maioria nas sequências de vídeo e são codificadas com os modos *SKIP* e *P16x16*. A detecção de bordas na imagem através da utilização do operador de Sobel é utilizada para determinar a homogeneidade dos blocos e a detecção de estacionariedade é realizada através do cálculo do SAD para determinar o modo *SKIP*.

Um conjunto de condições especificadas é utilizado para realizar uma finalização prévia do processo RDO, diminuindo a complexidade computacional do modo de decisão. Caso nenhuma das condições pré-definidas seja alcançada, todos os modos são checados com o RDO.

Os resultados mostraram uma redução média de 30% no processo de decisão inter-quadros, com uma perda de 0,030 dB na qualidade da imagem e um aumento de 0,6% na taxa de bits. No pior caso, ou seja, quando nenhuma condição é satisfeita, o algoritmo não apresenta nenhuma redução no número de cálculos.

Shen et al. (SHEN, 2008)

O trabalho se baseia na continuidade espacial do campo de movimento para realizar a decisão inter-quadros. Através de uma série de simulações, foi possível observar que regiões com movimento contínuo têm maior chance de serem codificadas com os modos *P16x16, P16x8* ou *P8x16*. Regiões estáticas são, em geral, codificadas com o modo *SKIP* e regiões com pouco movimento contínuo são codificados com o modo *P8x8*. Regiões sem movimento continuo tendem a usar sub-partições. O operador de Sobel também é utilizado neste trabalho para identificar bordas de objetos.

A aplicação da heurística proposta resultou em uma queda de 0,070 dB na qualidade da imagem e um aumento de 0,27% na taxa de bits. Simulações demonstraram que a redução no custo computacional foi aproximadamente de 50%.

HEURÍSTICAS PARA A DECISÃO RÁPIDA DE MODO DE CODIFICAÇÃO

A técnica RDO apresenta os melhores resultados na relação taxa-distorção. Entretanto, o tempo de computação necessário para avaliar todas as possibilidades de codificação para todos os blocos é um fator limitante no desempenho dos codificadores H.264/AVC. Nesse contexto, o desenvolvimento de algoritmos de decisão rápida para a escolha do modo de predição que será utilizado na codificação de um macrobloco se faz necessário, a fim de possibilitar a codificação de vídeos de alta resolução em tempo real.

Diversos trabalhos têm explorado características dos vídeos digitais, com o objetivo de diminuir a complexidade do processo de decisão de modo. Entretanto, a grande maioria destes trabalhos é direcionada a aplicações em software, desconsiderando o fluxo de dados do codificador H.264/AVC, quando implementado em hardware.

Este capítulo apresenta uma série de algoritmos heurísticos que foram desenvolvidos com o objetivo de diminuir o tempo gasto na escolha do modo de predição que será utilizada para codificar um determinado macrobloco. Primeiramente, são apresentadas heurísticas para decisão de modo intra-quadro, focando em uma futura implementação em hardware. Em seguida são apresentados os algoritmos heurísticos para a decisão de modo inter-quadros. Após a apresentação dos algoritmos de decisão, são apresentados os resultados em termos de PSNR e taxa de bits e uma comparação com a decisão baseada na técnica RDO. Todos os resultados apresentados neste capítulo foram obtidos através da execução do software de referência (VCEG) com seqüências de vídeos tipicamente utilizadas pela academia.

Decisão de Modo em Quadros I e em Quadros P

O software de referência do padrão H.264/AVC permite que a codificação dos macroblocos em quadros do tipo I seja realizada utilizando a predição intra-quadro e

que a codificação dos macroblocos em quadros do tipo P seja realizada tanto pela predição inter-quadros quanto pela predição intra-quadro. Entretanto, análises estatísticas sobre um conjunto de vídeos revelaram que a quantidade de macroblocos codificados pela predição intra-quadro em quadros P é muito pequena em relação à quantidade de macroblocos codificados pela predição inter-quadros. A Tab. 1 apresenta a ocorrência média de cada modo em quadros P. Oito sequências diferentes com resolução CIF (352x288 *pixels*) foram utilizadas nesta avaliação.

Tabela 1 - Ocorrência de modos intra-quadro e inter-quadros em quadros P de resolução CIF.

Modo	Ocorrência (Número de MBs)	Frequência (%)
Intra-Quadro	11.438	4,5
Inter-Quadros	242.285	95,5

Como é mostrado na Tab. 1, somente 4,5% dos macroblocos são codificados utilizando o modo intra-quadro. Além disso, outra simulação, utilizando as mesmas sequências, foi realizada com o objetivo de comparar a qualidade (PSNR) e a taxa de bits dos vídeos quando são utilizados macroblocos codificados com a predição intra e inter em quadros P e quando são utilizados somente macroblocos codificados com predição inter em quadros P. A Tab. 2 mostra essa comparação.

Tabela 2 - Comparação entre a codificação com predições intra e inter em quadros P e a codificação com apenas predição inter em quadros P.

Vídeo	Intra em P		Somente Inter em P		Comparação	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR (dB)	+ Taxa de Bits (%)
FOREMAN	37,613	530.984	37,624	535.424	0,011	0,84
BUS	35,936	1.315.648	35,959	1.326.232	0,023	0,80
CITY	35,788	768.808	35,794	769.888	0,006	0,14
FOOTBALL	36,035	1.647.808	36,038	1.697.496	0,003	3,02
HARBOUR	35,562	1.525.488	35,564	1.528.824	0,002	0,22
ICE	40,152	464.944	40,173	471.656	0,021	1,44
SOCCER	36,118	910.360	36,131	915.464	0,013	0,56
MÉDIA	**36,743**	**1.023.434**	**36,755**	**1.034.998**	**0,011**	**1,13**

Os resultados apresentados na Tab. 2 mostram um crescimento de apenas 1,13% na taxa de bits quando os quadros P são compostos apenas por macroblocos codificados com a predição inter-quadros. A qualidade da imagem sofreu um acréscimo de 0,011 dB. Desse modo, este trabalho foi desenvolvido considerando que macroblocos codificados pela predição intra-quadro serão utilizados somente em quadros do tipo I, enquanto que quadros do tipo P utilizarão somente macroblocos codificados pela predição inter-quadros. Sendo assim o problema de decisão de modo foi dividido em dois: (1) decisão sobre macroblocos do tipo I e (2) decisão sobre macroblocos do tipo P.

Todos os resultados desta avaliação foram relatados em (CORRÊA, 2010), sendo que todo o trabalho de desenvolvimento foi realizado em parceria com o autor desta monografia.

Decisão de Modo Rápida para Macroblocos Intra-Quadro

A decisão para macroblocos codificados pela predição intra-quadro foi desenvolvida de forma hierárquica em dois passos:

Passo 1 – Decisão entre mesmos tamanhos de bloco:

- Decisão que escolhe o melhor modo para blocos de tamanho 4x4 (I4MB) de amostras de luminância;
- Decisão que escolhe o melhor modo para blocos de tamanho 16x16 (I16MB) de amostras de luminância;
- Decisão que escolhe o melhor modo para o bloco de tamanho 8x8 de amostras de crominância.

Passo 2 – Decisão entre tamanhos de bloco diferentes:

- Decisão que escolhe qual tamanho de bloco (I4MB e I16MB) será utilizado para codificar o bloco de amostras de luminância.

Todas as decisões citadas acima no Passo 1 são independentes entre si, ou seja, não definem uma ordem no fluxo de processamento e ainda podem ser realizadas concorrentemente. Neste trabalho, as escolhas de melhor modo citadas no Passo 1 são definidas de acordo com uma métrica de distorção. Considerando a decisão que escolhe qual o tamanho de bloco será utilizado na codificação do macrobloco, este trabalho propõe duas heurísticas: (1) decisão baseada em heterogeneidade e (2) decisão baseada em diferença de distorção. Os resultados referentes

à primeira heurística estão apresentados em (CORRÊA, 2010) e (CORRÊA, PALOMINO, *et al.*, 2010) e foram desenvolvidos em conjunto com os autores destes trabalhos.

1.1.4 Heurística Baseada em Distorção (SAD) para a Decisão entre Mesmos Tamanhos de Bloco

Para a decisão entre mesmos tamanhos de bloco, foram realizadas simulações considerando três métricas de distorção: Soma das Diferenças Absolutas (SAD), Soma das Diferenças Absolutas Transformadas (SATD) e Soma das Diferenças Quadráticas (SSD). Os resultados gerados pela utilização destas três métricas foram comparados com a utilização da técnica RDO. Além disso, foi realizada uma comparação entre as três métricas considerando a complexidade, medida em número de somas, e os resultados de qualidade e taxa de bits gerados pelo vídeo codificado. As simulações foram realizadas com vídeos de resolução CIF (352x288 pixels).

Os resultados obtidos com essas simulações mostram que a métrica SATD apresentou os melhores resultados considerando qualidade e taxa de bits do vídeo codificado. Entretanto, a complexidade computacional necessária para a realização do cálculo do SATD é muito maior do que as outras duas métricas, sendo SAD a métrica que possui a menor complexidade. Como o objetivo deste trabalho é o desenvolvimento de um módulo de decisão que realize a escolha do melhor modo o mais rápido possível, as operações envolvidas no módulo de decisão devem ter o menor custo computacional possível. Sendo assim, é necessária uma avaliação que considere o custo computacional necessário para realizar a decisão, além da qualidade e da taxa de bits do vídeo codificado. A Tab. 3 mostra os resultados médios em termos de PSNR e taxa de bits, com a adição do custo computacional, calculado em número de somas, que cada técnica utiliza para gerar a métrica de distorção. Neste caso, a multiplicação no SSD foi desdobrada em adições sucessivas. Os resultados de número de somas apresentados na Tab. 3 foram gerados considerando o cálculo de distorção sobre um bloco 4x4. Foram utilizados oito vídeos de resolução CIF (352x288 pixels).

Tabela 3 - Resultados obtidos com SAD, SSD e SATD.

SAD			SSD			SATD		
PSNR (dB)	Taxa de Bits	Nº Somas	PSNR (dB)	Taxa de Bits	Nº Somas	PSNR (dB)	Taxa de Bits	Nº Somas
37.073	5.349k	31	37.081	5.316k	143	37.151	5.289k	139

A Tab. 4 apresenta uma comparação envolvendo as três métricas, sendo o SAD o parâmetro de comparação.

Tabela 4 - Comparação de SAD com SSD e SATD.

SSD vs. SAD			SATD vs. SAD		
PSNR (dB)	Taxa de Bits (%)	Nº Somas (%)	PSNR (dB)	Taxa de Bits (%)	Nº Somas (%)
-0,008	-0,63	+361,29	-0,079	-1,13	+348,39

Apesar das métricas SSD e SATD apresentarem resultados melhores em termos de PSNR e taxa de bits, os ganhos não são suficientemente expressivos para justificar a utilização de alguma destas métricas, visto que o custo computacional da utilização das métricas SSD e SATD chega a ser 361,29% maior que a utilização da métrica SAD. Por isso, neste trabalho, a decisão entre mesmos tamanhos de bloco, considerando macroblocos codificados pela predição intra (crominância e luminância), é realizada utilizando a métrica SAD, que apresenta um custo computacional muito menor que as outras duas métricas avaliadas (SSD e SATD).

1.1.5 Heurísticas para a Decisão entre Tamanhos de Bloco Diferentes

Para realizar a decisão sobre qual tamanho de bloco será utilizado (I4MB ou I16MB) foram desenvolvidas, neste trabalho, duas heurísticas diferentes. A primeira utiliza uma métrica que serve para determinar o nível de heterogeneidade do macrobloco, enquanto que a segunda utiliza os níveis de distorção, já calculados pelo passo de decisão de mesmos tamanhos de bloco, para realizar a decisão sobre qual tamanho de bloco será utilizado. Os resultados referentes à primeira heurística estão apresentados em (CORRÊA, 2010) e (CORRÊA, PALOMINO, *et al.*, 2010) e não serão relatados neste trabalho. As duas heurísticas desenvolvidas, foram implementadas e integradas junto ao software de referência e os próximos parágrafos apre-

sentam a segunda heurística, bem como os resultados obtidos. As simulações foram realizadas sobre vídeos de resolução CIF (352x288) e os resultados gerados pela heurística foram comparados com a técnica RDO.

Como resultado do 1º passo de decisão, obtém-se a distorção gerada pelo melhor modo do bloco 16x16 e pelo melhor modo dos 16 blocos 4x4. Se for realizada uma simples comparação entre a distorção gerada pelo melhor modo I16MB com a distorção gerada pelos melhores modos I4MB, na grande maioria dos casos, notar-se-ia um valor de SAD menor para os modos I4MB em relação ao melhor modo I16MB, ocasionando um grande número de escolhas para o primeiro tamanho de bloco, caso a decisão fosse baseada apenas na análise da distorção. Entretanto, através de uma análise da diferença entre os dois valores é possível decidir qual tamanho de bloco será utilizado (I4MB ou I16MB). Uma série de simulações foi realizada, a fim de classificar as diferenças de distorção em relação ao modo escolhido quando a técnica RDO é utilizada. A Fig. 11 apresenta um gráfico onde é realizada uma análise entre a quantidade de modos escolhidos (I4MB ou I16MB) pela técnica RDO (em porcentagem) em relação à diferença de distorção medida em SAD. Foram utilizados três quadros do vídeo *ICE* de resolução CIF (352x288) na análise.

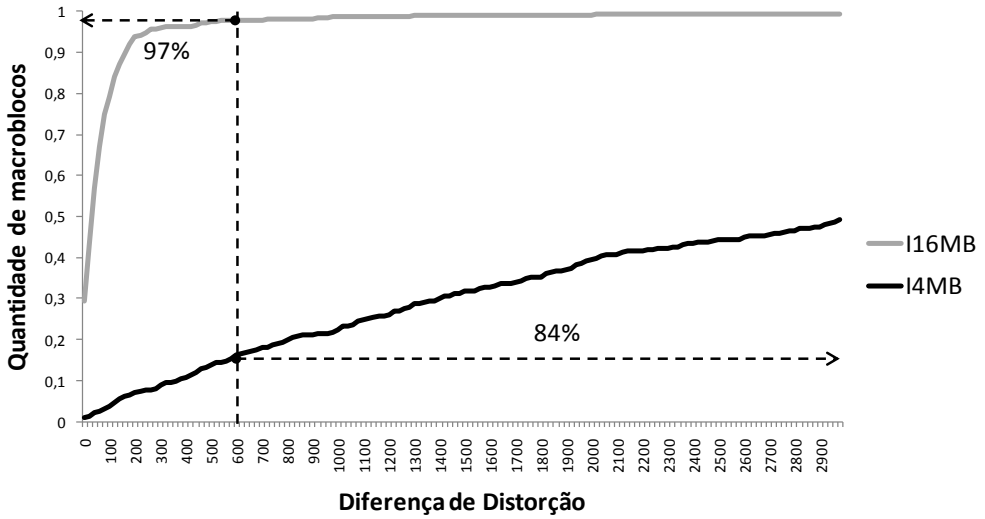

Figura 11 - Gráfico da quantidade de modos escolhidos pela técnica RDO em relação à diferença de distorção baseada em SAD.

Através de uma análise do gráfico acima, é possível notar que a maioria das vezes em que o modo escolhido é o I16MB a diferença de distorção é bem baixa, enquanto que a maioria das escolhas I4MB é realizada para diferenças de distorção altas. Desse modo é possível utilizar essa diferença de distorção, em comparação

com um valor limite, para decidir se um macrobloco será codificado com o modo I4MB ou com o modo I16MB. No gráfico acima, para um valor limite igual a 600, a quantidade de escolhas iguais aquelas realizadas pela técnica RDO chega a 97% considerando os modos I16MB e 84% para os modos I4MB

A metodologia utilizada para obter o valor limite que gera os melhores resultados em termos de PSNR e taxa de bits foi a seguinte: primeiramente os vídeos foram codificados utilizando a técnica RDO para realizar a decisão e os valores de distorção foram salvos junto com o modo escolhido. Em seguida, as diferença de distorção foram comparadas com o modo escolhido para definir o valor limite. Várias simulações foram realizadas com o valor limite variando de 0 a 1000, sendo o valor limite igual a 600 o que gerou os melhores resultados.

A Tab. 5 apresenta os resultados da avaliação em software para o módulo de decisão intra-quadro completo considerando o uso de distorção (SAD) para a decisão de modo para blocos de mesmo tamanho e a heurística de diferença de distorção (DD) para decisão de melhor tamanho de bloco (valor limite igual a 600). As simulações utilizaram vídeos de resolução HD1080p (1920x1080 pixels).

Tabela 5 - Resultados do módulo de decisão intra-quadro completo (SAD & DD).

Vídeo	RDO		SAD & DD		RDO vs. SAD&DD	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	- PSNR	+ Taxa de Bits (%)
STATION 2	39,643	34022k	39,338	35729k	0,305	5,02
SUNFLOWER	42,766	33030k	42,413	35391k	0,353	7,15
TRACTOR	39,400	58208k	39,035	60939k	0,365	4,69
TRAFFIC	39,643	50884k	39,308	53070k	0,335	4,30
MANINCAR	42,609	8818k	42,518	9018k	0,091	2,27
PEDISTRIAN AREA	40,878	24949k	40,699	26875k	0,179	7,72
RIBERBED	38,652	56735k	38,280	58962k	0,372	3,93
ROLLING TOMATOES	40,475	12170k	40,387	12537k	0,088	3,02
RUSHHOUR	41,694	20013k	41,484	21306k	0,210	6,46
MÉDIA	**40,640**	**33203k**	**40,385**	**34870k**	**0,255**	**5,02**

Os resultados apresentados na Tab. 5 com o módulo de decisão intra-quadro proposto (SAD & DD) mostram um pequeno crescimento médio de 5,02% na taxa de bits e uma pequena perda média de 0,255 dB na qualidade da imagem, em comparação à aplicação da técnica RDO. Novamente, essas perdas geradas pela utilização do módulo de decisão intra-quadro proposto não são tão significantes se as compararmos com a redução obtida em termos de complexidade computacional.

1.1.6 Decisões Intra-Quadro e Análise de Complexidade

Nesta seção será apresentada uma comparação considerando as duas heurísticas desenvolvidas (CORRÊA, 2010) para o módulo de decisão intra-quadro e a técnica RDO. Essa comparação é importante para avaliar o peso das perdas geradas pela heurística em relação ao RDO, levando em consideração a redução na complexidade computacional das heurísticas propostas.

A Tab. 6 apresenta os resultados médios gerados pela técnica RDO e pelas heurísticas, em termos de PSNR e taxa de bits, com a adição do custo computacional que cada solução necessita para realizar a decisão. O custo computacional é apresentado considerando o número de iterações necessárias para realizar a decisão, ou seja, o número de vezes que o laço de codificação (Fig. 10) necessita ser processado para que a decisão sobre um bloco esteja completa.

Tabela 6 - Resultados médios obtidos com a técnica RDO, Heurística baseada em Heterogeneidade e a Heurística DD.

RDO			Heurística Heterogeneidade			Heurística DD		
PSNR (dB)	Taxa de Bits	Nº Itera.	PSNR (dB)	Taxa de Bits	Nº Itera.	PSNR (dB)	Taxa de Bits	Nº Itera.
40,640	33203k	13	40,351	35713k	1	40,385	34870k	1

A Tab. 7 mostra uma comparação entre as duas heurísticas desenvolvidas em relação ao RDO.

Tabela 7 - Comparação entre RDO vs. Heterogeneidade e RDO vs. DD.

RDO vs. Heterogeneidade			RDO vs. DD		
PSNR (dB)	Taxa de Bits (%)	Nº Iterações (vezes)	PSNR (dB)	Taxa de Bits (%)	Nº Iterações (vezes)
+0,289	-7,02	+13	+0,255	-4,78%	+13

A comparação evidencia que, apesar dos ganhos em PSNR e taxa de bits existentes quando a técnica RDO é utilizada para a decisão de modo intra-quadro, a redução na complexidade computacional das soluções propostas neste trabalho é muito maior. O número de cálculos necessários para realizar a decisão baseada em RDO é aproximadamente 13 vezes maior quando comparado com as duas heurísticas apresentadas. Esses resultados mostram que é possível desenvolver um módulo de decisão rápido para os modos da predição intra-quadro, sem o uso de RDO, diminuindo consideravelmente o tempo necessário para realizar a codificação sobre um macrobloco.

Decisão de Modo Rápida para Macroblocos Inter-Quadros

A decisão de modo para macroblocos inter-quadros foi dividida, assim como a decisão para macroblocos intra-quadro, com a finalidade de simplificar um problema complexo em passos simples. As decisões propostas são, também, independentes entre elas e podem, portanto, ser realizadas em paralelo quando implementadas em hardware. No final de todas as decisões, os resultados são combinados para gerar o melhor modo. Novamente, a base para o desenvolvimento destas heurísticas está em (CORRÊA, 2010) e o trabalho foi realizado em colaboração com o autor deste trabalho.

A decisão inter-quadros consiste em decidir qual o melhor tamanho de partição e de sub-partição (explicados no capítulo 0 e apresentados na Fig. 2 e Fig. 3) para um determinado macrobloco. Considerando aplicações em software, a análise de ocorrência dos modos de codificação em quadros P auxilia em uma ordem preferencial de avaliação e decisão, diminuindo o tempo necessário para a decisão de modo. A partir de simulações realizadas com vídeos HD1080p codificados com base na decisão RDO, foram obtidas as frequências médias de cada modo para os quadros do tipo P, conforme apresentado na Tab. 8.

Tabela 8 - Ocorrência de modos de codificação em quadros P de resolução HD1080p.

Modo	Ocorrência	Frequência
SKIP	11.384	46,50%
P16x16	5.966	24,37%
P16x8	1.997	8,15%
P8x16	1.991	8,16%
P8x8	796	3,25%
Intra-quadro	2.346	9,57%

Como é possível ser observado na Tab. 8, os modos *SKIP* e *P16x16* somam aproximadamente 70% da totalidade de modos escolhidos. Em seguida, aparecem os modos *P16x8* e *P8x16*, com ocorrência semelhante (cerca de 8%). Conforme explicado anteriormente, quando o modo escolhido é o P8x8 ainda é realizado um subparticionamento, que totaliza, na análise realizada, 3,25% dos modos escolhidos. Sendo assim, é possível concluir que quanto maior a partição de macrobloco, maior é sua ocorrência, o que demonstra que os vídeos digitais de elevada resolução possuem uma grande quantidade de áreas homogêneas e estacionárias.

Considerando os resultados apresentados na Tab. 8 foi definida uma ordem hierárquica para a decisão de modo inter-quadros, conforme mostra a Fig. 12. Em uma implementação em software, essa ordem hierárquica pode diminuir o tempo necessário para a realização da decisão inter-quadros. Em casos de implementação em hardware, todos as decisões apresentadas na Fig. 12 podem ser realizadas em paralelo, pois, como explicado anteriormente, não há dependência entre elas.

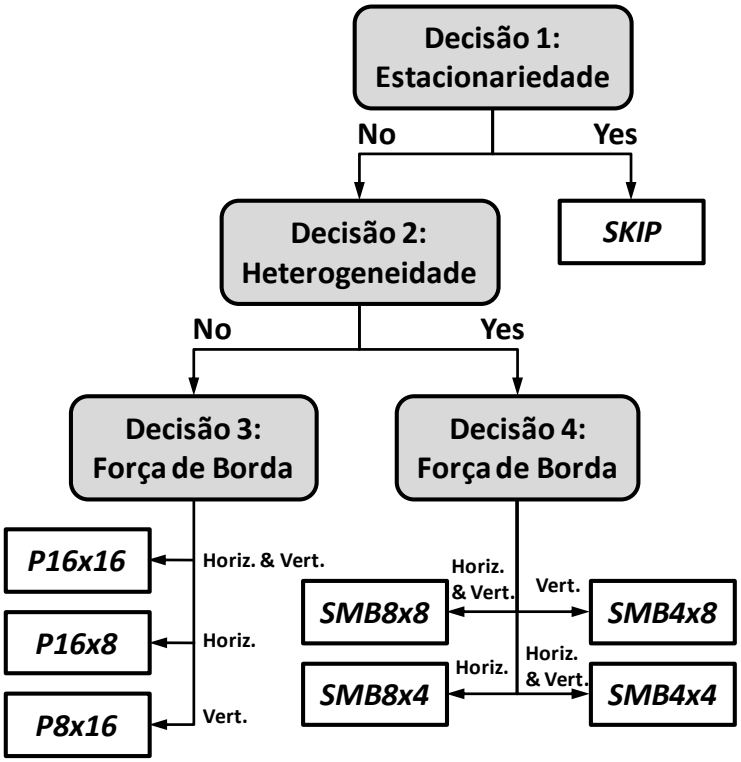

Figura 12 - Hierarquia para decisão de particionamento e sub-particionamento na predição inter-quadros.

As decisões apresentadas na Fig. 12 estão detalhadas nas próximas seções. A Decisão 1 realiza a primeira análise entre o macrobloco original e o macrobloco de referência, detectando regiões estacionárias entre estes macroblocos. A Decisão 2 é baseada em análise de heterogeneidade, de forma semelhante à decisão intra-quadro de melhor tamanho de bloco. Ela utiliza a métrica de heterogeneidade para decidir se serão utilizadas sub-partições de macrobloco ou não.

Todos os resultados apresentados nas próximas seções foram realizados com vídeos de resolução HD1080p (1920x1080 pixels).

1.1.7 Decisão 1: Detecção de Modo *SKIP* Baseada em Estacionariedade

Sequências de vídeos apresentam uma característica natural de serem compostas por movimentos constantes, e principalmente, por regiões temporalmente estacionárias. Esta característica é normalmente visível nas áreas presentes no fundo das imagens (WU, 2005). Existem vários trabalhos publicados na literatura que propõe métodos de detecção do modo *SKIP*, como cálculo adiantado do custo taxa-distorção (LIU, 2008) (MA, 2008), a previsão de blocos totalmente quantizados para

zero (WANG, 2007) (KIM, 2004) e a detecção de estacionariedade em uma sequência de quadros (WU, 2005) (YU, 2008).

Na decisão baseada em RDO, o modo *SKIP* é escolhido quando o custo taxa-distorção de não codificar nenhum resíduo é menor do que o custo de codificar utilizando qualquer um dos outros modos que codificam resíduo. Quando a decisão RDO não é utilizada, o padrão H.264/AVC define um conjunto de quatro condições que devem ser satisfeitas para que o modo *SKIP* seja utilizado. Contudo, de acordo com (YU, 2008) uma das condições já é, por si só, um bom indicador para a utilização do modo *SKIP*: se os coeficientes transformados do modo *SKIP* são todos quantizados para zero. Sabe-se que quanto menor for a amplitude dos valores dos resíduos gerados pelas etapas de predição, maior é a possibilidade de eles serem quantizados para zero. Sendo assim, alguns trabalhos vêm explorando a idéia de estacionariedade para a detecção do modo *SKIP*.

A estacionariedade é uma forma simples de identificar o grau de similaridade entre cada quadro de um vídeo e o seu quadro de referência. No trabalho de (WU, 2005), a estacionariedade é detectada através de uma comparação de quadros temporalmente vizinhos utilizando a métrica SAD. Caso o SAD seja menor que um valor limite, a região da imagem é considerada estacionária e o macrobloco é codificado com o modo *SKIP*. Esta ideia foi adotada neste trabalho, conforme será explicado a seguir.

O valor limite a ser usado como referência para a detecção de estacionariedade entre macroblocos foi obtido através de uma série de simulações em software realizadas com vídeos de resolução HD1080p. Os vídeos foram codificados com base na técnica RDO e os valores de SAD para cada macrobloco, tanto o localizado no quadro atual como o co-localizado no quadro de referência, foram salvos. Em seguida, os valores de SAD nos casos em que o modo escolhido foi o *SKIP* e os valores da SAD nos casos em que o modo foi diferente do *SKIP* foram agrupados e as respectivas médias foram calculadas. Com base nos valores médios obtidos, novas simulações foram realizadas para definir um valor limite de SAD para a detecção de macroblocos estacionários. O valor limite foi variado de 200 a 700, possibilitando a definição do valor 500 como o valor limite que gerou os melhores resultados. Em outras palavras, o valor SAD dos macroblocos é comparado com o valor limite 500 para decidir se o macrobloco é estacionário ou não. Caso o valor SAD seja menor que 500 então o macrobloco é definido como estacionário, caso contrário, o macrob-

loco é definido como não estacionário, e deve ser codificado utilizando algum dos outros modos de codificação inter-quadros.

Algumas rotinas foram inseridas no software de referência, para que a decisão sobre o modo *SKIP* fosse realizada com base na detecção de estacionariedade proposta neste trabalho. Se um macrobloco for definido como estacionário, o modo de codificação escolhido é o *SKIP* e nenhum outro modo é analisado. Entretanto, se o macrobloco for definido como não estacionário, então os outros modos são analisados com base na técnica RDO.

A Tab. 9 apresenta os resultados obtidos com a decisão de modo *SKIP* baseada em detecção de estacionariedade e os compara com a decisão RDO. Os resultados estão apresentados em termos de taxa de bits e PSNR.

Tabela 9 - Decisão de modo SKIP baseada em estacionariedade.

Vídeo	RDO		*SKIP* por Estacionariedade		RDO vs. Estacionariedade (Valor limite = 500)	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR	+ Taxa de Bits (%)
STATION 2	39,640	4.407k	39,700	4.586k	0,060	4,07
SUNFLOWER	42,580	7.105k	42,600	7.162k	0,020	0,80
TRACTOR	38,920	11.344k	38,940	11.393k	0,020	0,44
TRAFFIC	39,370	8.657k	39,400	8.752k	0,030	1,10
MANINCAR	42,430	2.138k	42,510	2.301k	0,080	7,63
RIVERBED	38,710	15.729k	38,710	15.585k	0,000	-0,92
ROLLING TOMATOES	40,360	3.427k	40,410	3.869k	0,050	12,90
RUSHHOUR	41,470	3.711k	41,530	3.859k	0,040	3,99
MÉDIA	**40,435**	**7.065k**	**40,475**	**7.188k**	**0,040**	**1,75**

A utilização do valor limite igual a 500 resultou em um aumento médio de apenas 1,75% na quantidade de bits do vídeo codificado e um aumento médio de 0,040 dB na qualidade da imagem gerada. É possível notar que para o vídeo *ROLLING TOMATOES* o resultado de taxa de bits ficou muito acima da média dos outros vídeos. Como a detecção de estacionariedade é uma heurística que gera resultados aproximados, em alguns casos pode acontecer de que o número de escolhas *SKIP*

seja muito menor em comparação ao uso de RDO, ou seja, uma quantidade maior de resíduo terá que ser codificada, aumentando a quantidade de bits do vídeo codificado. Se considerarmos o vídeo *RIVERBED*, é possível notar que o efeito citado anteriormente ocorre de forma contrária. Nesse caso, mais macroblocos foram definidos como estacionários e codificados com o modo *SKIP* em relação à decisão RDO, gerando uma diminuição na quantidade de resíduo codificado, consequentemente na quantidade de bits do vídeo codificado.

1.1.8 Decisão 2: Detecção de Sub-Particionamento Baseada em Heterogeneidade do Macrobloco

Além de serem compostas por regiões estacionárias, as sequências de vídeos digitais possuem outra característica importante: uma grande quantidade de regiões homogêneas. As regiões homogêneas tendem a apresentar movimento em direções similares, enquanto que regiões heterogêneas apresentam movimento desordenado (YU, 2004) (CHOI, 2006). Isso acontece porque as regiões homogêneas, em geral, compõem um mesmo objeto que possui partes que se movem juntas, continuamente, na mesma direção e sentido. Enquanto que as regiões heterogêneas são, em geral, compostas por objetos diferentes que se movem em direções e sentidos diferentes. Sendo assim, regiões homogêneas são geralmente codificadas em partições de macroblocos maiores, como *P16x16*, *P16x8* e *P8x16*. A prévia identificação de heterogeneidade de um macrobloco pode evitar a tarefa de analisar partições de tamanho menor em uma implementação sequencial.

Como apresentado em (CORRÊA, 2010) e (CORRÊA, PALOMINO, *et al.*, 2010) o nível de heterogeneidade de um macrobloco pode ser detectado através da análise de alguns coeficientes gerados após a aplicação de uma transformada, como a DCT. Essa mesma técnica foi utilizada como base para a detecção de sub-particionamento. Testes semelhantes foram realizados com vídeos de resolução HD 1080p para encontrar a métrica de heterogeneidade que gere os melhores resultados. Nas simulações apresentadas a seguir o vídeo *MAN IN CAR* foi codificado com a decisão RDO completa. Os macroblocos codificados com sub-particionamento foram separados daqueles que não utilizaram sub-particionamento e os valores de heterogeneidade (**H**) foram calculados para cada macrobloco, a fim de se obter o valor limite que melhor se encaixe com as decisões realizadas pelo RDO.

A Fig. 13 mostra um gráfico de dispersão que relaciona macroblocos codificados com os modos *P16x16*, *P16x8* e *P8x16* (ou seja, sem sub-particionamento) pela decisão RDO (eixo x) com os seus respectivos valores **H** (eixo y).

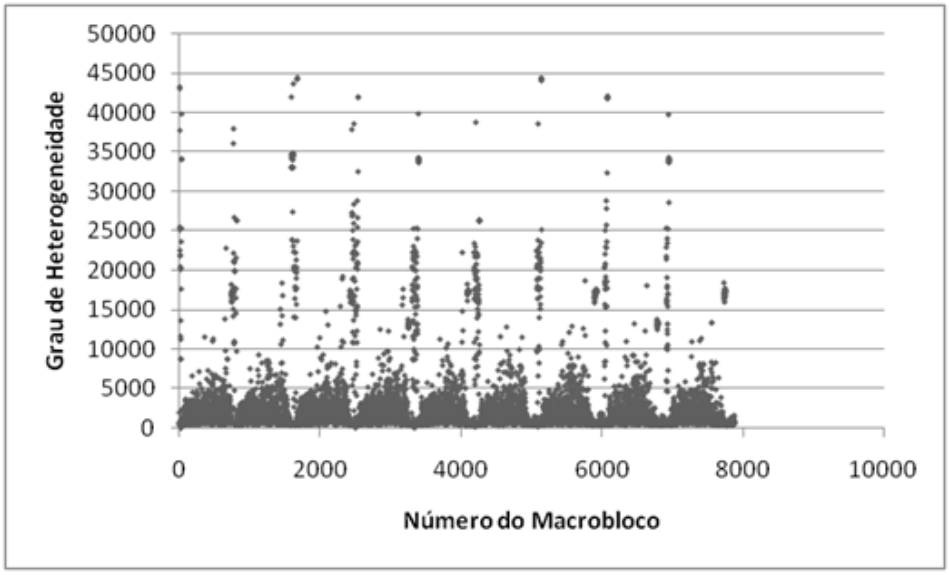

Figura 13 - Valores de **H** para macrobloco codificados sem sub-particionamento através da decisão RDO.

É possível perceber que a maioria dos macroblocos codificados sem sub-particionamento possuem um valor de **H** menor que 10.000. A Fig. 14 apresenta um gráfico semelhante para macroblocos da mesma seqüência. Entretanto, somente os macroblocos codificados com sub-particionamento pela decisão RDO são apresentados.

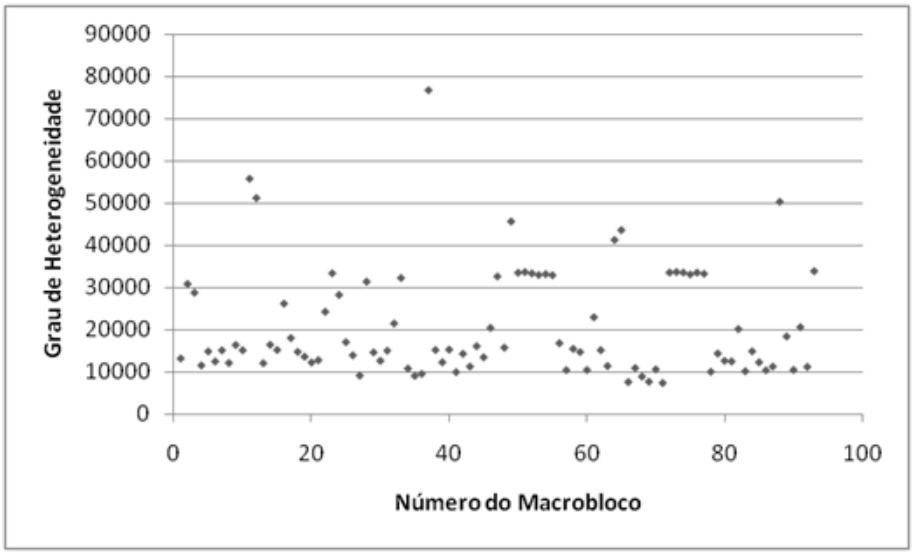

Figura 14 - Valores de **H** para macrobloco codificados com sub-particionamento através da decisão RDO.

Nesse caso a maioria dos macroblocos codificados com sub-particionamento teve seu valor **H** maior que 10.000. Esse comportamento, em geral, se repete para as oito sequências utilizadas nas simulações. Sendo assim, o valor limite em torno de 10.000 é um bom indicador para a detecção de sub-particionamento.

Resumindo a métrica de heterogeneidade **H** é calculada e comparada ao valor 10.000, se o valor de **H** for menor, então o sub-particionamento não é utilizado e o macrobloco é codificado com os modos *P16x16*, *P16x8* e *P8x16*. Se o valor de **H** for maior que 10.000, então o sub-particionamento é utilizado e para cada partição *P8x8* é gerada uma sub-partição que pode ser do tipo *SMB8x8, SMB8x4, SMB4x8* ou *SMB4x4*.

A Tab. 10 apresenta os resultados obtidos pela heurística de detecção de sub-particionamento em termos de taxa de bits e PSNR. É importante esclarecer que, nesta análise, somente a decisão sobre utilizar ou não sub-particionamento é realizada com a heurística proposta. A decisão sobre qual particionamento ou sub-particionamento será utilizado é realizada pela decisão RDO. A utilização do valor limite 10.000 gerou um aumento médio de 0,62% na taxa de bits do vídeo codificado e a diferença em qualidade foi quase desprezível.

Tabela 10 - Resultados da detecção de sub-particionamento em comparação com o RDO.

Vídeo	RDO		Sub-Part. por Heterog.		RDO vs. Sub-Part. por Heterog.	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR	+ Taxa de Bits (%)
STATION 2	39,640	4.407k	39,640	4.428k	0,000	0,49
SUNFLOWER	42,580	7.105k	42,590	7.149k	0,010	0,61
TRACTOR	38,920	11.344k	38,920	11.409k	0,000	0,58
TRAFFIC	39,370	8.657k	39,370	8.733k	0,000	0,89
MANINCAR	42,430	2.138k	42,440	2.135k	0,010	-0,14
RIVERBED	38,710	15.729k	38,720	15.837k	0,010	0,68
ROLLING TOMATOES	40,360	3.427k	40,360	3.425k	0,000	-0,04
RUSHHOUR	41,470	3.711k	41,470	3.749k	0,000	1,01
MÉDIA	**40,435**	**7.065k**	**40,438**	**7.109k**	**0,003**	**0,62**

1.1.9 Decisão 3: Tamanho do Particionamento Baseado em Força de Borda

Seguindo a ordem da decisão hierárquica proposta, após a detecção de sub-particionamento, o tamanho do particionamento é definido com base nas diferenças entre os pixels que compõem as linhas e as colunas do macrobloco. Conforme explicado anteriormente, as partes que compõem os objetos presentes em um vídeo se movem, em geral, na mesma direção e sentido. Assim, identificando as áreas limítrofes de um objeto é possível definir o tamanho da partição de macrobloco que agrupe as partes pertencentes a um mesmo objeto em partições únicas. A Fig. 15 mostra uma imagem e um possível particionamento sobre ela. É possível observar o particionamento *P16x16* nas áreas compostas por um único objeto ou, neste caso, o fundo da cena, enquanto que os particionamentos *P16x8* e *P8x16* são utilizados para distinguir partes que não pertencem ao mesmo objeto, que neste caso é a apresentadora de um programa de TV.

Figura 15 - Possível particionamento sobre uma imagem.

Neste trabalho, considera-se que a diferença entre as amostras de luminância é suficiente para distinguir dois objetos diferentes. Essa restrição pode gerar decisões incompatíveis em comparação com a técnica RDO, mas conforme explicado no final desta seção, a heurística proposta apresenta resultados bastante satisfatórios, apesar destas decisões incompatíveis.

Neste trabalho a detecção de arestas é realizada de uma forma semelhante à detecção de arestas utilizada no Filtro Redutor de Efeito de Bloco do padrão H.264/AVC. Este módulo é o responsável por amenizar o efeito de bloco gerado por

um elevado passo de quantização, através da detecção de arestas entre blocos. Ele ainda é capaz de distinguir uma aresta real da imagem (que não é filtrada) de uma aresta gerada pelo elevado passo de quantização (que necessita ser filtrada). Contudo, a detecção de arestas para decidir o tamanho do particionamento do macrobloco proposta neste trabalho é bem mais simples, pois, como a decisão ocorre em uma etapa da codificação em que a quantização ainda não foi aplicada e as bordas indesejáveis não foram geradas, somente as arestas reais da imagem serão detectadas.

A detecção é aplicada às bordas que podem gerar um possível particionamento. A Fig. 16 mostra as amostras que são utilizadas para a detecção de borda vertical ou horizontal de um macrobloco.

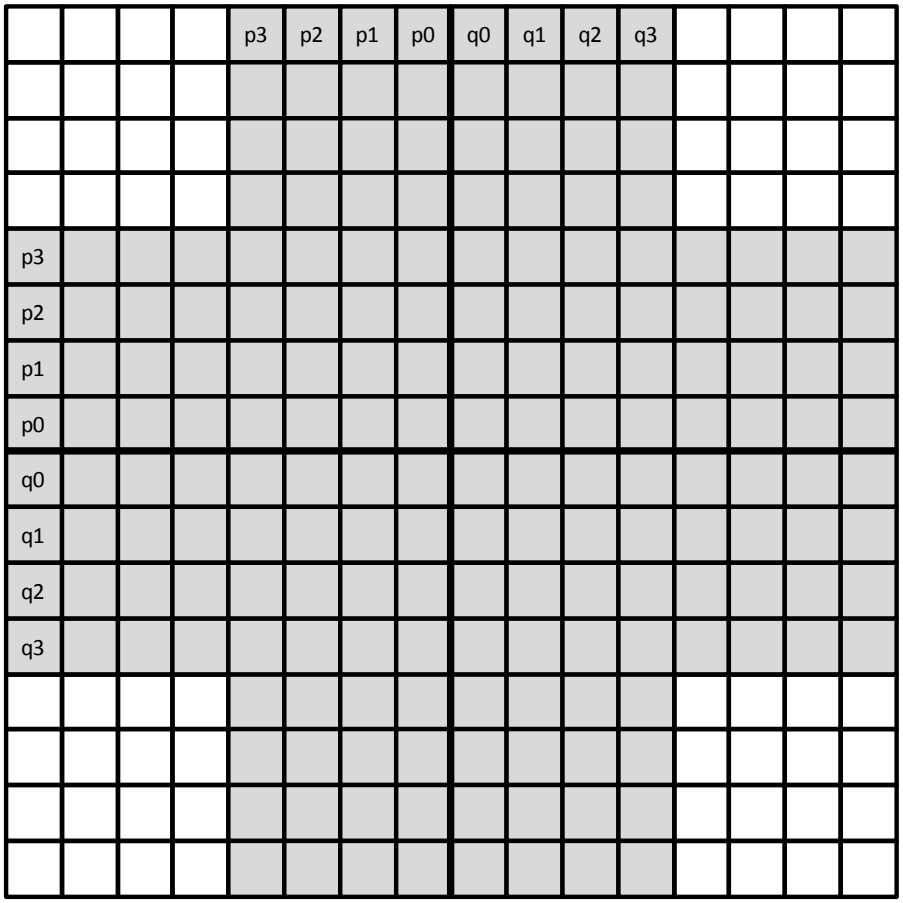

Figura 16 - Amostras (em cinza) utilizadas na detecção de arestas vertical e horizontal para decisão de particionamento de macrobloco.

São utilizadas oito amostras (**p0, p1, p2, p3, q0, q1, q2** e **q3** na Fig. 16) de cada linha ou coluna para a detecção de borda. A equação (13) mostra o cálculo

realizado para detectar a força de borda vertical e a equação (14) define o cálculo para detectar a força de borda horizontal, onde **O** é o bloco original, **i** é a linha do macrobloco e **j** é a coluna.

$$BV = \sum_{i=0}^{16} \sum_{j=4}^{7} \left(O_{i,j} - O_{i,(15-j)} \right) \tag{13}$$

$$BH = \sum_{j=0}^{16} \sum_{i=4}^{7} \left(O_{i,j} - O_{(15-i),j} \right) \tag{14}$$

Os pixels são comparados dois a dois de fora para dentro, ou seja, **p3** é comparado com **q3**, **p2** é comparado com **q2**, e assim por diante. Isso permite que, tanto mudanças abruptas quanto mudanças graduais sejam detectadas.

A partir de uma análise sobre os valores de força de borda **BH** e **BV** obtidos para cada macrobloco e a partir dos tamanhos de partição escolhidos, quando o modo RDO é utilizado, simulações, com o software de referência foram realizadas para determinar o valor limite a ser comparado para decidir qual partição será utilizada. Os melhores resultados em termos de taxa de bits e PSNR foram atingidos com o valor limite 80. O algoritmo abaixo descreve o processo de decisão baseado em força de borda para determinar o tamanho da partição.

Se |BH – BV| ≤ 80 Então
    partição = P16x16
Senão Se BH – BV > 80 Então
    partição = P16x8
Senão Se BV – BH > 80 Então
    partição = P8x16

A Tab. 11 mostra os resultados gerados pelo uso da heurística de detecção de força de borda para determinar qual partição será utilizada na codificação de um macrobloco inter-quadros. Uma comparação com a técnica RDO é realizada em termos de taxa de bits e qualidade do vídeo codificado.

Tabela 11 - Resultados gerados pela decisão de particionamento baseada em força de borda.

Vídeo	RDO		Partição por Força de Borda.		RDO vs. Part. Força de Borda	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR	+ Taxa de Bits (%)
STATION 2	39,640	4.407k	39,610	4.545k	0,030	3,13
SUNFLOWER	42,580	7.105k	42,590	7.532k	-0,010	6,01
TRACTOR	38,920	11.344k	38,960	12.220k	-0,030	7,72
TRAFFIC	39,370	8.657k	39,360	8.923k	0,010	3,07
MANINCAR	42,430	2.138k	42,400	2.138k	0,040	0,02
RIVERBED	38,710	15.729k	38,670	15.737k	0,040	0,05
ROLLING TOMATOES	40,360	3.427k	40,330	3.553k	0,030	3,67
RUSHHOUR	41,470	3.711k	41,440	3.899k	0,030	5,07
MÉDIA	**40,490**	**6.933k**	**40,480**	**7.187k**	**0,020**	**3,59**

Os resultados gerados pelo uso da técnica de detecção de força de borda para o particionamento mostram um crescimento médio na taxa de bits de 3,59% e um aumento de 0,020 dB na qualidade de imagem (PSNR). É possível notar que a qualidade da imagem e a taxa de bits varia de um vídeo para o outro. Como já explicado anteriormente, isso acontece devido à variação do número de particionamentos, maiores ou menores, escolhidos, já que a técnica é apenas uma aproximação da técnica RDO.

1.1.10 Decisão 4: Tamanho do Sub-Particionamento Baseado em Força de Borda

Após a aplicação da heurística baseada em heterogeneidade para decidir se será utilizado sub-particionamento ou não, é necessário definir o tamanho de sub-particionamento de cada partição *P8x8*. Isso significa que para cada partição *P8x8* será gerada uma informação de sub-particionamento, que pode ser: *SMB8x8*, *SMB8x4*, *SMB4x8* ou *SMB4x4*.

A técnica utilizada para definir o tamanho de sub-particionamento de macrobloco é similar à técnica utilizada na seção anterior. A única diferença está na quantidade de amostras necessárias para realizar o cálculo da detecção de força de bor-

da, que antes eram oito e agora são quatro por linha. Além disso, para que seja possível a identificação de sub-partições de tamanho 4x4, as bordas são calculadas em duas partes: (1) parte superior (ou esquerda) e (2) a parte inferior (ou direita). A Fig. 17 mostra a estrutura do macrobloco e quais as amostras (em cinza) são utilizadas no cálculo.

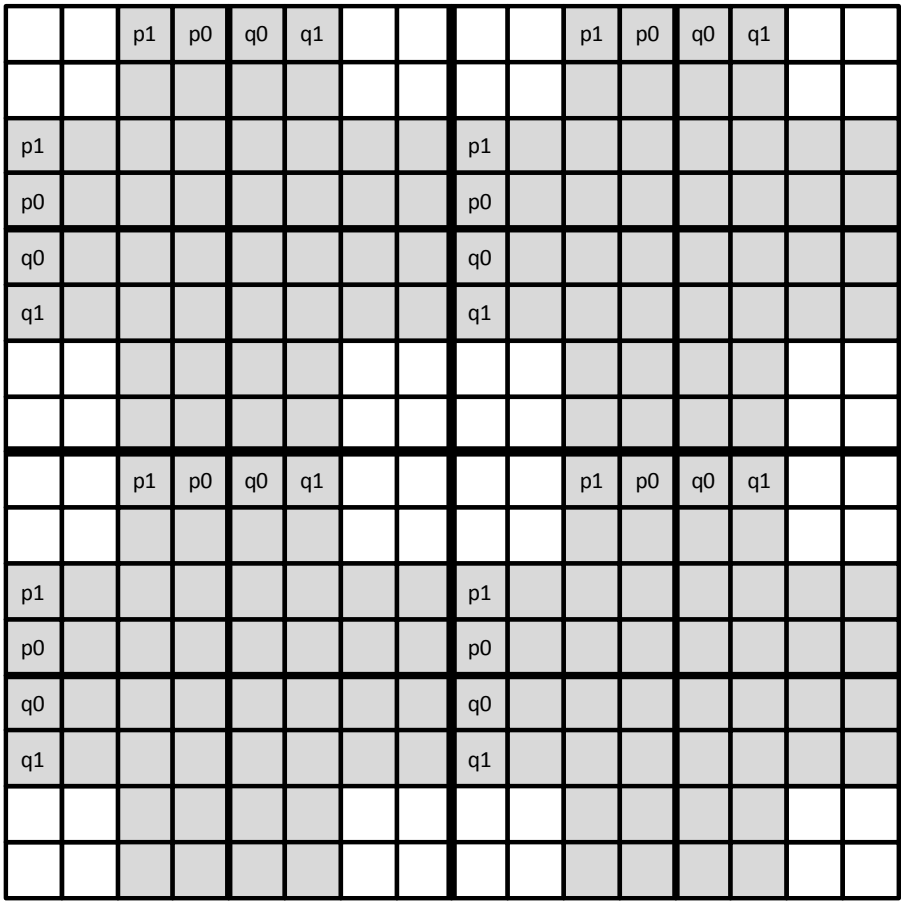

Figura 17 - Amostras (cinza) utilizadas para o cálculo das bordas verticais e horizontais para a decisão de tamanho de sub-particionamento de macrobloco.

As equações (19) e (20) apresentam os cálculos para as bordas verticais e horizontais de cada sub-macrobloco. **SBV** significa sub-borda vertical e **SBH** significa sub-borda horizontal, as quais representam as bordas verticais e horizontais de sub-macrobloco respectivamente. Como explicado anteriormente, as duas bordas (vertical e horizontal) são divididas em duas para possibilitar a detecção de sub-partições de tamanho 4x4. As equações (15), (16), (17) e (18) apresentam os cálculos para cada parte das bordas vertical (**SBV1** e **SBV2**) e horizontal (**SBH1** e **SBH2**).

$$SBV1 = \sum_{i=0}^{4} \sum_{j=2}^{3} \left(O_{i,j} - O_{i,(7-j)} \right) \quad (15)$$

$$SBV2 = \sum_{i=4}^{7} \sum_{j=2}^{3} \left(O_{i,j} - O_{i,(7-j)} \right) \quad (16)$$

$$SBH1 = \sum_{j=0}^{4} \sum_{i=2}^{3} \left(O_{i,j} - O_{(7-i),j} \right) \quad (17)$$

$$SBH2 = \sum_{j=4}^{7} \sum_{i=2}^{3} \left(O_{i,j} - O_{(7-i),j} \right) \quad (18)$$

$$SBV = SBV1 + SBV2 \quad (19)$$

$$SBH = SBH1 + SBH2 \quad (20)$$

De forma similar às outras heurísticas apresentadas neste trabalho, a partir dos valores de sub-borda vertical (**SBV**) e sub-borda horizontal (**SBH**), simulações foram realizadas para gerar um valor limite para ser utilizado como base para a decisão de qual tamanho de sub-particionamento será utilizado. O algoritmo apresentado a seguir explica o processo de detecção de tamanho de sub-particionamento.

Se |SBH – SBV| ≤ 40 Então
    tamanho_sub_partição = SMB8x8
Senão Se (SBH – SBV) > 40 Então
    Se SBV1 > 20 ou SBV2 > 20 Então
        tamanho_sub_partição = SMB4x4
    Senão
        tamanho_sub_partição = SMB8x4
Senão Se (SBV – SBH) >40 Então
    Se SBH1 > 20 ou SBH2 > 20 Então
        tamanho_sub_particão = SMB4x4
    Senão
        tamanho_sub_partição = SMB4x8

Caso a diferença entre **SBH** e **SBV** seja pequena não há muita diferença entre os pixels que compõem as duas metades do sub-macrobloco. Sendo assim, utiliza-se o maior tamanho de sub-partição possível *SMB8x8*. Caso essa diferença seja

significativa, então é necessário analisar qual das bordas é mais forte. Se for **SBH** então é feita uma análise sobre as partes de **SBV** (**SBV1** e **SBV2**) para detectar um possível sub-particionamento *SMB4x4*. Caso contrário, é utilizado o sub-particionamento SMB8x4, pois a borda horizontal é a mais forte. Entretanto, se **SBV** for mais forte que **SBH**, o teste sobre as partes de **SBH** (**SBH1** e **SBH2**) também é realizado para detectar uma possível sub-partição *SMB4x4*. Caso contrário, a sub-partição *SMB4x8* é a escolhida, pois a borda vertical é a mais forte.

A Tab. 12 apresenta os resultados gerados pela decisão de tamanho de sub-particionamento baseado em força de borda. De forma similar às outras simulações, somente a decisão sobre tamanho de sub-particionamento foi realizada com a heurística proposta nesta análise, sendo os outros passos realizados com base na técnica RDO.

Tabela 12 - Resultados gerados pela decisão de tamanho de sub-partição baseada em força de borda.

Vídeo	RDO		Sub-Partição por Força de Borda.		RDO vs. Sub-Part. Força de Borda	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR	+ Taxa de Bits (%)
STATION 2	39,640	4.407k	39,650	4.423k	-0,010	0,37
SUNFLOWER	42,580	7.105k	42,560	7.152.k	0,020	0,66
TRACTOR	38,920	11.344k	38,890	11.435k	0,030	0,81
TRAFFIC	39,370	8.657k	39,360	8.666k	0,010	0,11
MANINCAR	42,430	2.138k	42,440	2.096k	0,000	-1,97
RIVERBED	38,710	15.729k	38,630	15.747k	0,080	0,12
ROLLING TOMATOES	40,360	3.427k	40,360	3.460k	0,000	0,96
RUSHHOUR	41,470	3.711k	41,450	3.729k	0,030	0,48
MÉDIA	**40,440**	**6.933k**	**40,420**	**7.089k**	**0,020**	**0,34**

O uso da técnica de detecção de tamanho de sub-particionamento baseado em força de borda gerou um aumento médio de apenas 0,34% na taxa de bits e um aumento de 0,020 dB na qualidade da imagem (PSNR). A aplicação desta heurística gerou resultados muito próximos aos obtidos com a utilização do RDO, pois, como

pode ser visto na Tab. 8, a quantidade de macroblocos que são sub-particionados é muito pequena, diminuindo o impacto do erro gerado pela heurística.

1.1.11 Decisões para Macroblocos Inter-Quadros Integradas e Análise de Complexidade

Todas as decisões para macroblocos inter-quadros (estacionariedade, heterogeneidade, detecção de força de borda) foram integradas com o objetivo de gerar um módulo completo de decisão rápida para macroblocos inter-quadros. Todas as heurísticas foram implementadas junto ao software de referência do padrão H.264/AVC e as simulações utilizaram os mesmos vídeos de resolução HD 1080p. Os resultados gerados pelo módulo de decisão para macroblocos inter-quadros estão apresentados na Tab. 13.

Tabela 13 - Resultados gerados pelo módulo completo de decisão rápida inter-quadros.

Vídeo	RDO		Modo de Decisão Inter Proposto		RDO vs. Decisão Inter Proposta	
	PSNR (dB)	Taxa de Bits	PSNR (dB)	Taxa de Bits	+ PSNR	+ Taxa de Bits (%)
STATION 2	39,640	4.407k	39,600	4.630k	0,040	5,06
SUNFLOWER	42,580	7.105k	42,590	7.636k	-0,010	7,46
TRACTOR	38,920	11.344k	38,870	11.897k	0,060	4,88
TRAFFIC	39,370	8.657k	39,330	10.650k	0,030	23,03
MANINCAR	42,430	2.138k	42,400	2.171k	0,040	1,55
RIVERBED	38,710	15.729k	38,550	15.981k	0,160	1,60
ROLLING TOMATOES	40,360	3.427k	40,320	3.740k	0,040	9,14
RUSHHOUR	41,470	3.711k	41,400	4.066k	0,080	9,57
MÉDIA	**40,440**	**6.933k**	**40,380**	**7.596k**	**0,050**	**7,53**

Os resultados mostram um crescimento médio de 7,53% na taxa de bits e um crescimento médio de 0,050 dB na qualidade da imagem codificada (PSNR). A qualidade praticamente não se alterou em relação ao RDO, pois, ela está diretamente relacionada à quantidade de bits do vídeo codificado. Em geral, uma maior taxa de bits implica em uma maior qualidade de imagem.

Apesar do crescimento da taxa de bits, ele ainda é muito pequeno em comparação com a redução da complexidade computacional gerada pelo modo de decisão proposto comparado à técnica RDO. Conforme apresentado na Fig. 10 no capítulo 0, o cálculo do custo gerado pela técnica RDO é realizado apenas após a codificação e decodificação completa para todos os modos. Já a decisão inter-quadros proposta neste trabalho, depende somente do macrobloco original e do macrobloco de referência, eliminando completamente a necessidade de codificação e decodificação completa para realizar a decisão de modo.

A predição inter-quadros apresenta cinco modos de codificação diferentes (*SKIP, P16x16, P16x8, P8x16 e P8x8*). Se o modo *P8x8* é escolhido, existem ainda mais quatro modos para cada partição 8x8 (*SMB8x8, SMB8x4, SMB4x8* e *SMB4x4*). Além disso, a técnica RDO completa, ainda considera os macroblocos codificados com a predição intra-quadro (I16MB e I14MB) para os quadros P. Sendo assim, considerando o fluxo da Fig. 10, o número de iterações necessárias para a decisão sobre um macrobloco é uma para o *SKIP*, três para os tamanhos de partição, *P16x16*, *P16x8* e *P8x16*, quatro para cada partição *P8x8*, e mais 13 para os modos intra-quadro. Desse modo, são necessárias 33 iterações para realizar a decisão sobre um macrobloco de um quadro do tipo P.

No módulo de decisão de modo inter-quadros proposto neste trabalho, somente uma iteração, no laço apresentado na Fig. 10, é realizada. Esta iteração é necessária somente para a reconstrução do macrobloco, utilizado como referência para futuras predições, e para a geração do macrobloco codificado, não sendo necessária para realizar a decisão.

Como neste trabalho considerou-se que macroblocos codificados pela predição intra-quadro são utilizados somente em quadros I e não em quadros P, uma avaliação sobre as duas decisões integradas não foi realizada, pois dependeria fortemente do GOP (*Group of Pictures*) utilizado. O GOP define a frequência de quadros do tipo I utilizada na codificação de um vídeo. Um GOP com valor 5 define uma sequência do tipo IPPPP, ou seja, a cada cinco quadros, um é codificado somente com a predição intra-quadro.

Comparação com Trabalhos Relacionados

Nesta seção é apresentada uma comparação dos resultados gerados pelas decisões propostas neste trabalho com trabalhos publicados na literatura (apresen-

tados no capítulo 0) que, também, visam reduzir a complexidade do processo de decisão de modo. A Tab. 14 apresenta esta comparação em termos de taxa de bits, qualidade de imagem e redução na complexidade computacional medida em número de iterações no laço de codificação (Fig. 10). A decisão intra-quadro baseada em heterogeneidade (CORRÊA, 2010) utiliza uma transformada DCT para realizar o cálculo da métrica de heterogeneidade e então decidir qual tamanho de bloco utilizar, I4MB se for muito heterogêneo ou I16MB se for pouco heterogêneo.

Tabela 14 - Comparação com trabalhos da literatura.

Trabalhos	Perda PSNR (dB)	Aumento Taxa de Bits (%)	Redução no Número Iterações (vezes)
(SUN e YINYI, 2008)	0,136	0,69	4,7
(FENGQIN, 2008)	0,020	3,00	5,9
(LEE, 2009)	0,088	0,09	4,0
(JEON, 2009)	0,148	4,00	6,0
(WU, 2005)	-0,030	0,60	6,0
(SHEN, 2008)	0,070	0,27	0,0
Decisão Intra (Corrêa 2010)	**0,289**	**7,56**	**13,0**
Decisão Intra (DD)	**0,255**	**5,02**	**13,0**
Decisão Inter	**-0,050**	**7,53**	**33,0**

Os números apresentados para redução no número de iterações referem-se ao pior caso. É possível notar que a redução na complexidade computacional apresentada neste trabalho é muito maior em comparação aos trabalhos relacionados. Enquanto os outros trabalhos apresentam reduções que variam de 4 a 6 vezes (em relação à técnica RDO) no número de iterações, este trabalho apresentou uma redução de 33 vezes quando a decisão inter-quadros é considerada e de 13 vezes para as duas decisões intra-quadro propostas.

Entretanto, a expressiva redução em complexidade computacional gera uma perda média na compressão de 7,53% em relação ao uso de RDO quando considerada a decisão inter-quadros proposta. Apesar de esta perda ser bem maior quando comparada com os trabalhos relacionados, este valor perde o seu impacto quando considerada a enorme redução de complexidade computacional obtida. A qualidade

de imagem praticamente não se altera, com uma pequena diminuição de 0,289 dB no caso da decisão intra-quadro baseada em heterogeneidade. A decisão inter-quadros proposta, por sua vez, levou até mesmo a um pequeno acréscimo na qualidade da imagem.

ARQUITETURAS PARA DECISÃO RÁPIDA DE MODO DE CODIFICAÇÃO

Neste capítulo serão apresentadas as arquiteturas desenvolvidas com base nas heurísticas apresentadas no capítulo anterior. Do mesmo modo que o desenvolvimento das heurísticas foi dividido em dois: (1) decisão intra e (2) decisão inter, as arquiteturas foram desenvolvidas, também, separadamente. Todas as arquiteturas foram descritas em VHDL e sintetizadas para o dispositivo FPGA EP2S130F1508C3 da família Stratix II da Altera (ALTERA, 2010) e para *standard cells* com a tecnologia TSMC 0,18µm (COMPONENTS, 2001). As arquiteturas foram sintetizadas considerando tecnologias diferentes com o objetivo de permitir o seu uso em várias aplicações.

Arquiteturas para Decisão de Modo Intra-Quadro

Nesta seção são apresentadas as arquiteturas desenvolvidas para o módulo de decisão intra-quadro com base na heurística baseada em heterogeneidade (CORRÊA, PALOMINO, *et al.*, 2010) e na heurística apresentada na seção 0. As arquiteturas foram desenvolvidas em módulos, considerando os passos descritos na seção 0.

1.1.12 Arquitetura Para o Cálculo de Distorção Baseada em SAD

A arquitetura desenvolvida neste trabalho para o cálculo de SAD entre o bloco original e o bloco predito é consideravelmente simples. A Fig. 18 apresenta o diagrama em blocos da arquitetura do calculador de SAD utilizado neste trabalho, onde a letra **O** representa o bloco original e **P** representa o bloco predito.

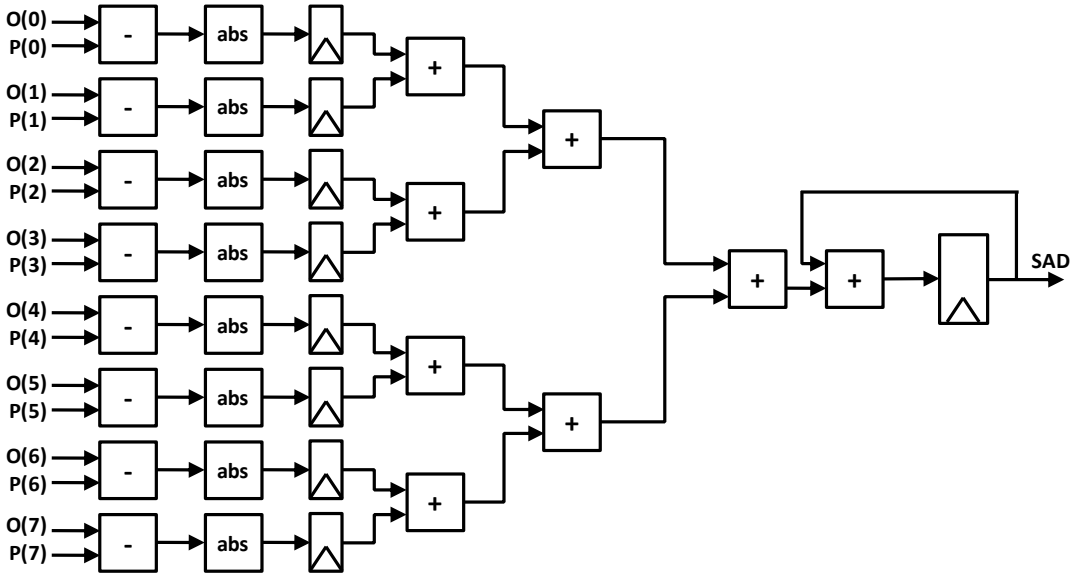

Figura 18 - Arquitetura do Calculador de SAD para oito amostras.

A arquitetura é composta por oito subtratores, oito calculadores de módulo, oito registradores e por uma árvore composta de sete somadores. Os registradores foram utilizados com o propósito de diminuir o caminho crítico do calculador de SAD, gerando um *pipeline* de dois estágios. Um registrador acoplado a um somador (isto é, um acumulador) é necessário para armazenar os resultados parciais, visto que a arquitetura consome apenas oito amostras por ciclo. Essa decisão foi tomada visando reduzir a quantidade de elementos lógicos da arquitetura final, pois a arquitetura para o cálculo de SAD será utilizada muitas vezes na composição da arquitetura de decisão de modo completa.

1.1.13 Arquitetura Para o Cálculo da Métrica de Heterogeneidade

Para realizar o cálculo da métrica de heterogeneidade utilizada na decisão entre tamanhos de bloco diferente, foi desenvolvida uma arquitetura que implementa a transformada DCT. Entretanto, como somente alguns coeficientes são necessários para o cálculo da métrica de heterogeneidade, uma arquitetura da DCT 16x16 parcial foi desenvolvida, desconsiderando o cálculo dos coeficientes que não são utilizados. Como apresentado em (CORRÊA, 2010), além da redução no número de coeficientes, a matriz de transformação da DCT foi aproximada para valores em potência de dois. Considerando uma implementação arquitetural, isso é muito útil, pois é possível implementar os cálculos utilizando apenas deslocadores para esquerda ou para a direita, eliminando possíveis multiplicações e divisões que aumentariam con-

sideravelmente o uso de hardware na arquitetura final. Desse modo a arquitetura para o cálculo dos coeficientes da DCT 16x16 foi composta, unicamente, por somadores, subtratores e deslocadores. Sendo assim, muitos dos cálculos utilizados para gerar os coeficientes apresentaram estrutura similar, possibilitando a utilização de mesmos operadores para realizar mais de um cálculo.

A Fig. 19 apresenta o diagrama em blocos da arquitetura que gera os coeficientes necessários para o cálculo da métrica de distorção.

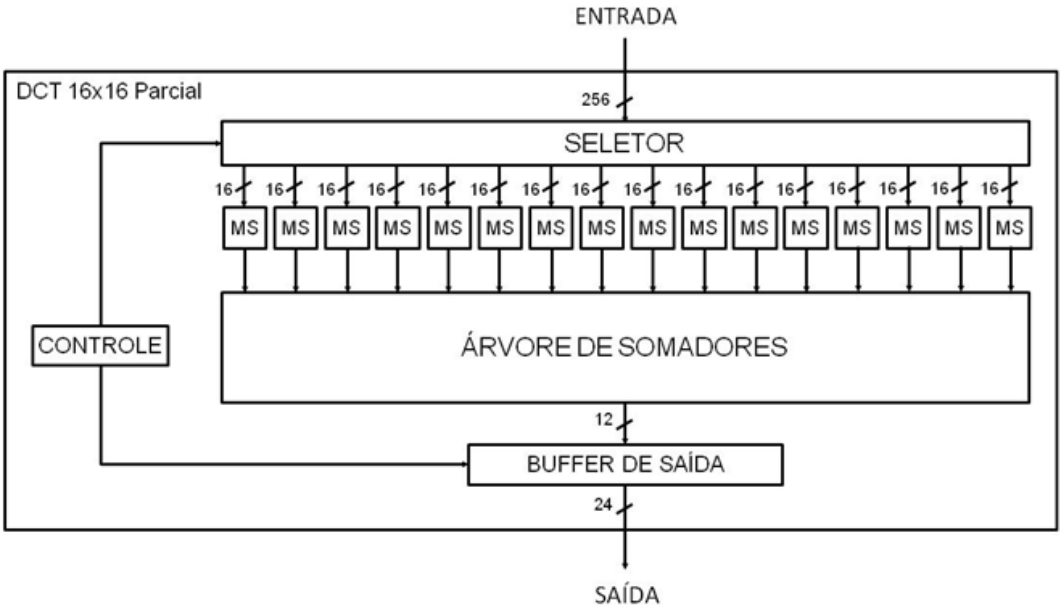

Figura 19 - Arquitetura que gera os coeficientes da transformada DCT 16x16 parcial necessários para o cálculo da métrica de heterogeneidade.

A arquitetura possui um paralelismo de 256 amostras, ou seja, o bloco original de dimensão 16x16 completo é necessário na entrada da arquitetura. Entretanto, considerando as restrições de acesso à memória, isso se torna inviável. Desse modo, foi desenvolvido um buffer na entrada da arquitetura que armazena as amostras do bloco original até que todas as amostras sejam lidas da memória, para, então, serem consumidas pela arquitetura. O módulo seletor é responsável por ordenar as amostras na entrada dos módulos *MS*. Os módulos *MS* são onde os cálculos de estrutura similar foram agrupados para poderem ser reutilizados no cálculo de mais de um coeficiente. Foram utilizados quatro estágios de *pipeline* para gerar um dos dados na saída dos calculadores *MS*. A árvore de somadores realiza os somatórios que não puderam ser agrupados e entrega em sua saída 12 coeficientes por ciclo. São necessários quatro ciclos para que um resultado válido seja gerado na saída da

Arvore de Somadores. O buffer de saída reordena os coeficientes e os entrega todos à saída. Considerando quatro ciclos para o término do cálculo do módulo *MS* e quatro ciclos para a árvore de somadores, são necessários oito ciclos para o cálculo de uma linha e mais oito para o cálculo da coluna, totalizando 16 ciclos para o cálculo dos coeficientes necessários para gerar a métrica de heterogeneidade.

1.1.14 Arquitetura de Decisão Intra-Quadro Baseada em Heterogeneidade

A arquitetura completa de decisão de modo intra-quadro baseada em heterogeneidade foi desenvolvida com base nas arquiteturas já apresentadas. A Fig. 20 mostra o diagrama em blocos da arquitetura completa.

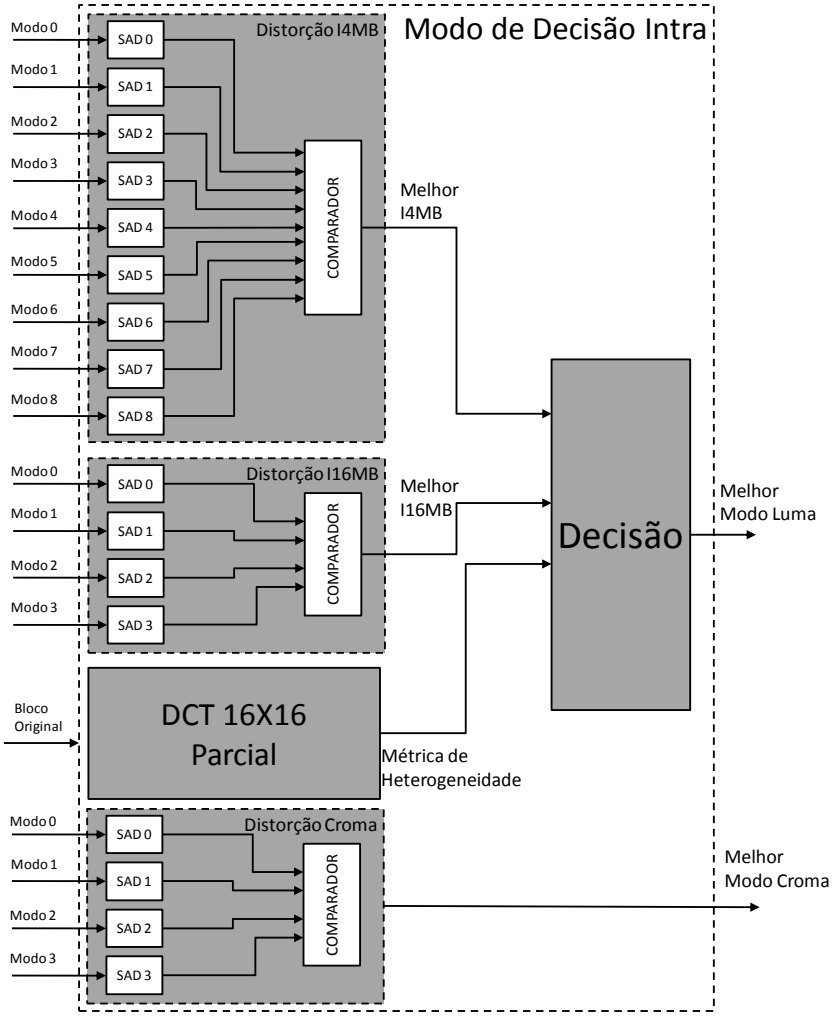

Figura 20 - Diagrama em blocos da arquitetura do módulo de decisão intra-quadro baseado em heterogeneidade.

O paralelismo da arquitetura é de oito amostras. Isso significa que somente oito amostras do bloco original e dos blocos gerados por cada modo de predição estão disponíveis por ciclo.

O cálculo de distorção entre mesmos tamanhos de bloco, considerando todos os modos, é realizado em paralelo, pois não há dependência entre esses módulos.

A Fig. 21 mostra um diagrama de ciclos que mostra o tempo necessário para a realização da decisão sobre um macrobloco.

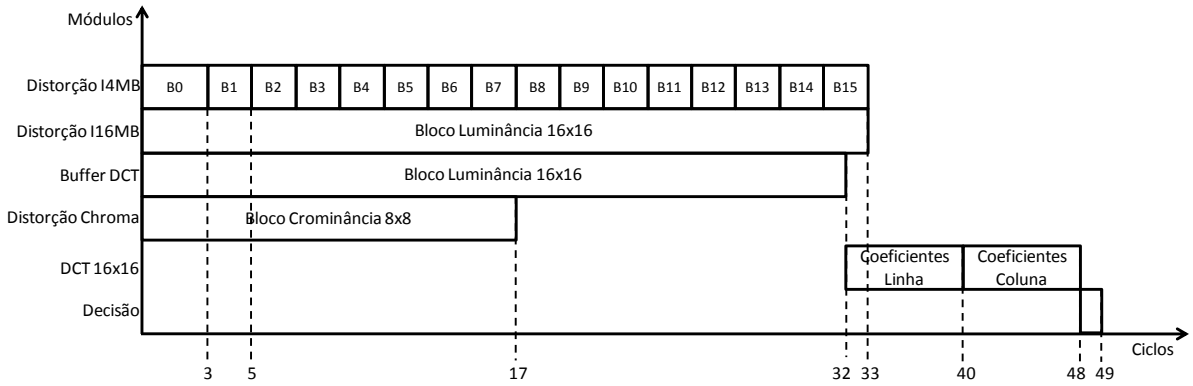

Figura 21 - Diagrama de ciclos para a arquitetura de decisão intra-quadro baseada em heterogeneidade.

Como somente oito amostras (duas linhas de um bloco 4x4) estão disponíveis por ciclo, são necessários três ciclos para o cálculo de distorção I4MB (SAD) para o primeiro bloco 4x4 (representado com B0 na Fig. 21), devido à barreira temporal presente no meio do calculador de SAD. Como a arquitetura de SAD foi desenvolvida com dois estágios de *pipeline*, os resultados do cálculo de SAD dos próximos blocos 4x4 estão disponíveis na saída a cada dois ciclos, totalizando 33 ciclos para o cálculo de distorção sobre todos os modos I4MB. O cálculo de distorção sobre os modos I16MB também é realizado em 33 ciclos. O cálculo de distorção para os blocos de crominância 8x8 é realizado em 17 ciclos. No mesmo momento que as amostras do bloco original estão sendo utilizadas para os cálculos de distorção, elas estão, também, sendo armazenadas no *buffer*. Após 32 ciclos, quando o *buffer* possui as 256 amostras do bloco 16x16, é inicializado o processamento da DCT, que como dito anteriormente, necessita de 16 ciclos para ser finalizado. Por último, um único ciclo é necessário para realizar a comparação entre a métrica de heterogeneidade, e então, escolher o melhor tamanho de bloco (I4MB ou I16MB) que será disponibilizado na saída. Sendo assim, são necessários 49 ciclos para realizar a decisão intra sobre um macrobloco.

A Tab. 15 apresenta os resultados de síntese da arquitetura considerando duas tecnologias (FPGA e *standard cell*). Os resultados são apresentados em termos de uso de recursos de hardware e frequência máxima de operação atingida.

Tabela 15 - Resultados de síntese da arquitetura de decisão intra-quadro baseada em heterogeneidade.

FPGA Stratix II			Standard Cell	
ALUTs	DLRs	Frequência (MHz)	Área (Número de *Gates*)	Frequência (MHz)
4.224	4.737	103,24	133.228	129,1

Considerando a síntese para FPGA, a arquitetura utilizou 4.224 ALUTs (*Adaptative Look-Up Tables*) e 4.737 DLRs (*Dedicated Logic Registers*), totalizando 6% do total de recursos lógicos do dispositivo. A frequência máxima de operação atingida foi de 129,1 MHz para a síntese em *standard cell*.

Através dos dados de frequência máxima atingida, número de ciclos para processar um macrobloco e a quantidade de macroblocos em um quadro, é possível calcular a taxa de processamento em quadros por segundo da arquitetura desenvolvida. Neste caso, temos 49 ciclos para processar um macrobloco e 129,1 MHz de frequência máxima de operação. Considerando um vídeo de resolução HD 1080p (1920x1080 pixels), temos 8160 macroblocos em um quadro. Sendo assim, a arquitetura é capaz de processar até 322 quadros de resolução HD 1080p por segundo, atingindo facilmente tempo real. Além disso, se considerarmos uma taxa de processamento de 30 quadros por segundo, a arquitetura pode operar a uma frequência de 12,0 MHz. Esse dado é muito importante, pois, além de mostrar que a arquitetura é capaz de operar a baixas frequências, economizando energia, é um dado independente de tecnologia.

1.1.15 Arquitetura de Decisão Intra-Quadro Baseada em Diferença de Distorção

A arquitetura para a decisão baseada em distorção é similar à arquitetura anterior. A Fig. 22 mostra o diagrama em blocos da arquitetura.

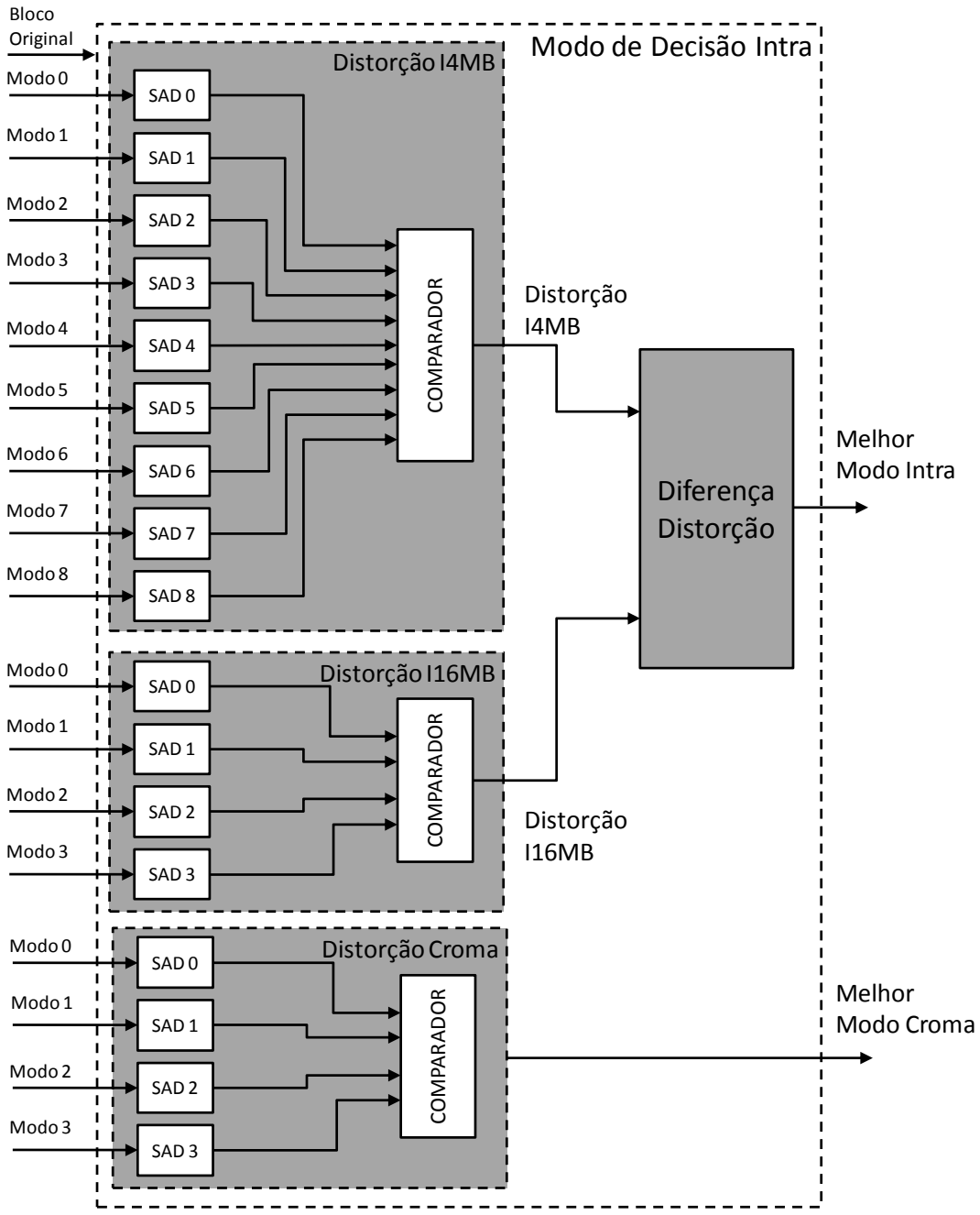

Figura 22 - Diagrama em blocos da arquitetura do módulo de decisão intra-quadro baseado em diferença de distorção.

A arquitetura possui as mesmas características que a anterior, entretanto, não há a necessidade da transformada DCT 16x16 parcial, pois a decisão é realizada com base nas distorções geradas pelos calculadores de SAD. O módulo de decisão (Diferença de Distorção) realiza uma subtração entre as diferenças para, então, compará-la com um valor limite e decidir qual o tamanho de bloco será utilizado (I4MB ou I16MB).

Sem a necessidade da transformada, o número de ciclos necessários para realizar a decisão sobre um macrobloco é reduzido em comparação com a arquitetura anterior. A Fig. 23 mostra o diagrama de ciclos para esta arquitetura.

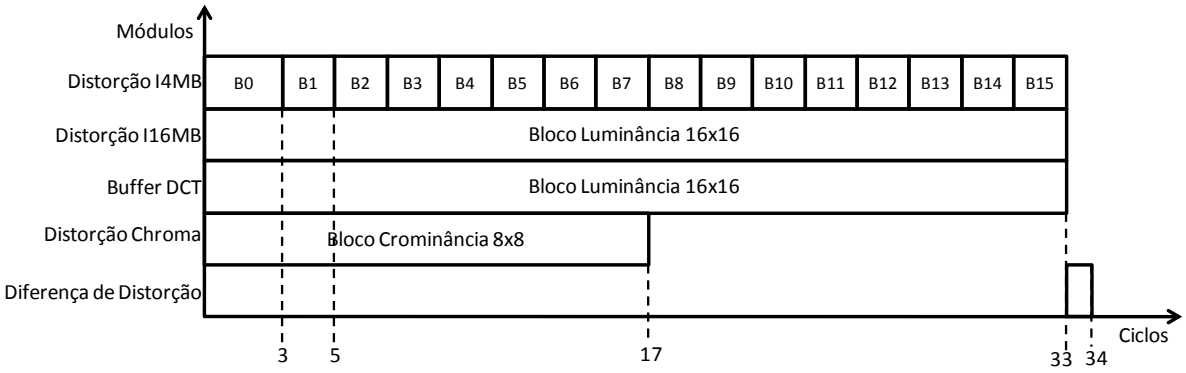

Figura 23 - Diagrama de ciclos para a arquitetura de decisão intra-quadro baseada em diferença de distorção.

O número de ciclos utilizados para realizar o cálculo das distorções são os mesmos apresentados pela arquitetura anterior. Como somente esse resultado é necessário para realizar a decisão final (pelo módulo Diferença de Distorção), somente um ciclo a mais é necessário, totalizando 34 ciclos para realizar a decisão intra-quadro para um macrobloco.

A Tab. 16 apresenta os resultados de síntese considerando as mesmas duas tecnologias já utilizadas (FPGA e *standard cell*). Os resultados são apresentados em termos de utilização de recursos de hardware e frequência máxima de operação atingida.

Tabela 16 - Resultados de síntese da arquitetura de decisão intra-quadro baseada em diferença de distorção.

FPGA Stratix II			Standard Cell	
ALUTs	DLRs	Frequência (MHz)	Área (Número de *Gates*)	Frequência (MHz)
3.267	2.312	98,43	28.518	129,1

Foram utilizadas 3.267 ALUTs e 2.312 DLRs do FPGA, totalizando 4% dos recursos do dispositivo. A maior frequência foi atingida pela síntese para *standard cell*, chegando a 129 MHz.

Levando em conta os resultados de síntese apresentados na Tab. 16, o número de ciclos que a arquitetura necessita para processar um macrobloco e um vídeo de resolução HD 1080p (1920x1080 pixels), foi realizado o cálculo da taxa de processamento da arquitetura em quadros por segundo, do mesmo modo que foi realizado para a arquitetura anterior. Sendo assim, a arquitetura é capaz de processar até 465 quadros por segundo quando a frequência máxima de operação é considerada. Para atingir uma taxa de processamento de 30 quadros por segundo, a arquitetura precisa operar com apenas 8,32 MHz de frequência.

Arquitetura para Decisão de Modo Inter-Quadros

Nesta seção, são apresentadas as arquiteturas desenvolvidas com base nas heurísticas apresentadas na seção 0. As arquiteturas foram desenvolvidas em módulos segundo a hierarquia proposta na Fig. 12, e integradas para formar um módulo de decisão inter-quadros completo.

A arquitetura baseada em detecção de estacionariedade para decisão de modo *SKIP* foi desenvolvida utilizando a arquitetura de SAD proposta na Fig. 18 e a arquitetura para realizar a decisão de quando utilizar sub-particionamento ou não é a mesma utilizada para detecção de heterogeneidade na decisão intra-quadro.

A seção seguinte apresenta as arquiteturas desenvolvidas para o cálculo de força de borda utilizada nas heurísticas de detecção de tamanho de particionamento e sub-particionamento.

1.1.16 Arquiteturas Para Decisão de Tamanho de Partição e Sub-Partição

A arquitetura desenvolvida para o cálculo de força de borda é similar ao cálculo do SAD, entretanto somente as amostras do bloco original (**p3**, **p2**, **p1**, **p0**, **q3**, **q2**, **q1** e **q0** na Fig. 16) são utilizadas. A Fig. 24 mostra a arquitetura para o cálculo de força de borda para decisão de tamanho de partição.

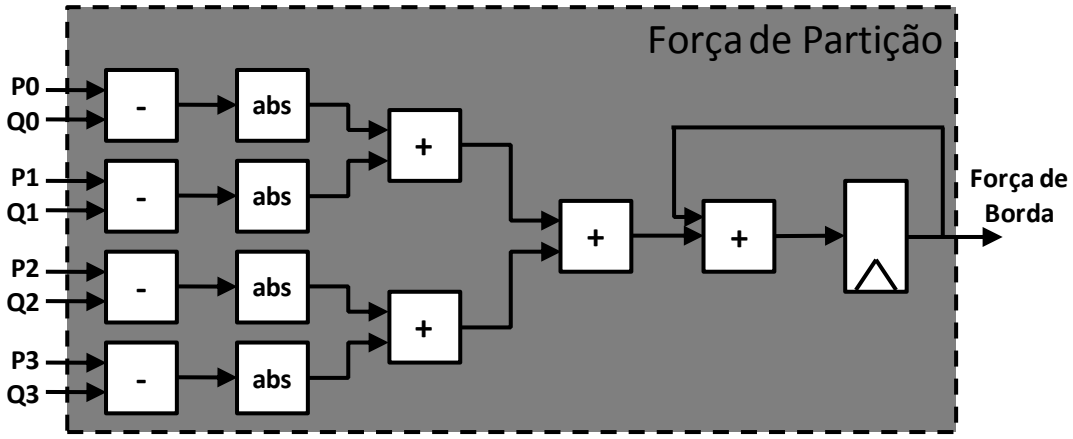

Figura 24 - Arquitetura para o cálculo de força de borda para decisão de tamanho de partição.

A arquitetura realiza o cálculo da força de borda de uma linha por ciclo. Isso significa que são necessários 16 ciclos para realizar o cálculo de cada borda (oito para a boda vertical e oito para a horizontal).

A arquitetura para o cálculo de força de borda para tamanho de sub-partição é similar à apresentada na Fig. 24, entretanto menos amostras são utilizadas para o cálculo (apenas **p0**, **q0**, **p1** e **q1** na Fig. 17). A Fig. 25 mostra esta arquitetura.

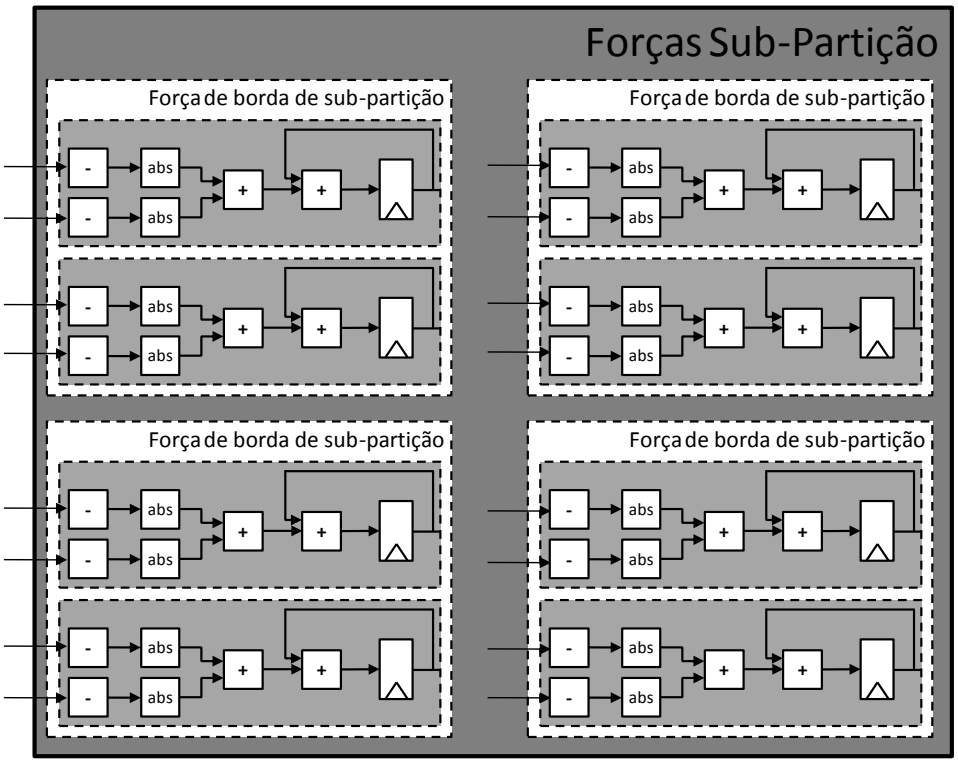

Figura 25 - Arquitetura para o cálculo de força de borda para decisão de tamanho de sub-partição.

Conforme explicado na seção 1.1.10 o cálculo da força de borda para sub-particionamento é realizado em duas partes: a força da metade superior (ou à esquerda) e a força da metade inferior (ou à direita) da sub-partição. São utilizados, portanto, quatro ciclos para computar a força de borda de cada metade da sub-partição, totalizando oito ciclos para o cálculo de força de borda da sub-partição em cada sentido. São necessárias oito instâncias deste módulo que calcula a força de borda para tamanho de sub-partição, sendo cada um deles responsável por metade da borda de cada uma das quatro sub-partições 8x8. Sendo assim, são geradas 16 informações de força de borda, uma para cada metade vertical ou horizontal de cada bloco da partição 8x8 (ver Fig. 17)

1.1.17 Arquiteturas Integradas Para a Decisão de Modo Inter-Quadros

Todas as arquiteturas desenvolvidas para as heurísticas apresentadas na seção 0 foram integradas para gerar o módulo de decisão inter-quadros completo. A Fig. 26 apresenta o diagrama em blocos desta arquitetura onde todos os módulos podem ser visualizados. MB Original e MB Referência se referem ao macrobloco original e de referência, respectivamente.

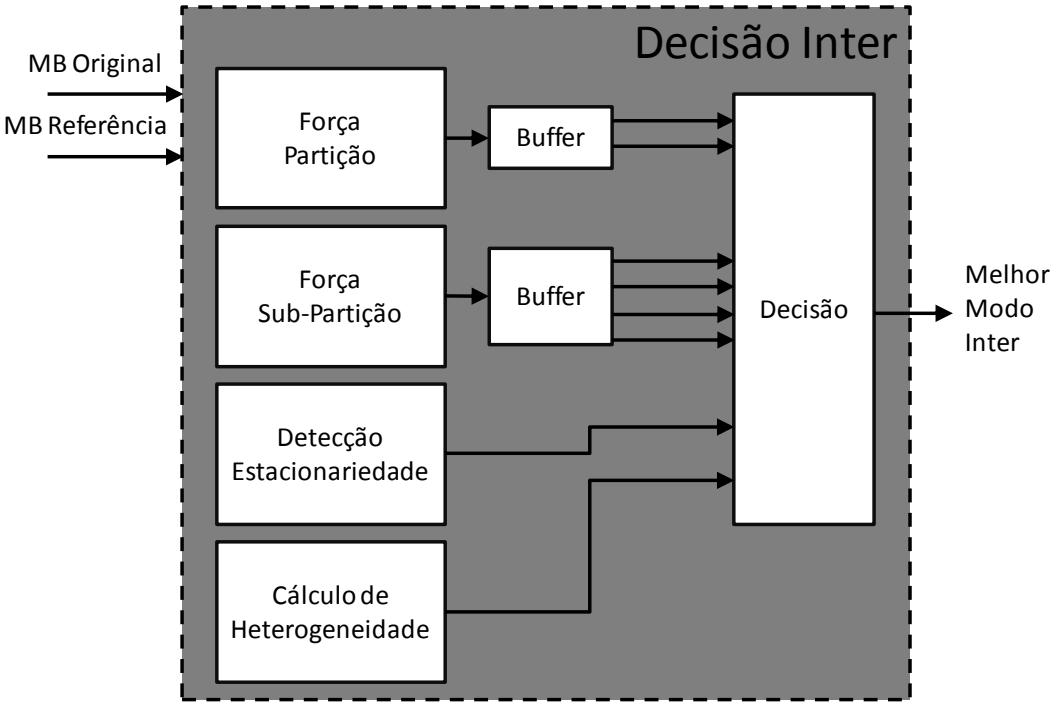

Figura 26 - Diagrama em blocos da arquitetura de decisão de modo inter-quadros completa.

O módulo Força de Partição é o responsável por gerar a força de borda vertical e horizontal para decisão sobre tamanho de partição. O módulo Força de Sub-Partição realiza o cálculo de forças de borda para decisão sobre tamanho de sub-partição. A decisão sobre modo *SKIP* é realizada com base no resultado gerado pelo módulo Detecção de Estacionariedade. Por fim, o módulo que realiza o cálculo do nível de heterogeneidade gera a informação para decidir se haverá ou não sub-particionamento sobre o macrobloco. Todos os cálculos realizados por estes módulos são realizados em paralelo, pois não há dependência de dados entre eles. Após o processamento destes módulos, os resultados gerados são avaliados pelo módulo de decisão, o qual realiza uma avaliação sobre esses dados para gerar a decisão do melhor modo inter-quadros.

A Fig. 27 mostra o diagrama de ciclos da arquitetura de decisão de modo inter-quadros, onde é possível ter uma noção temporal de como os módulos realizam o seu processamento.

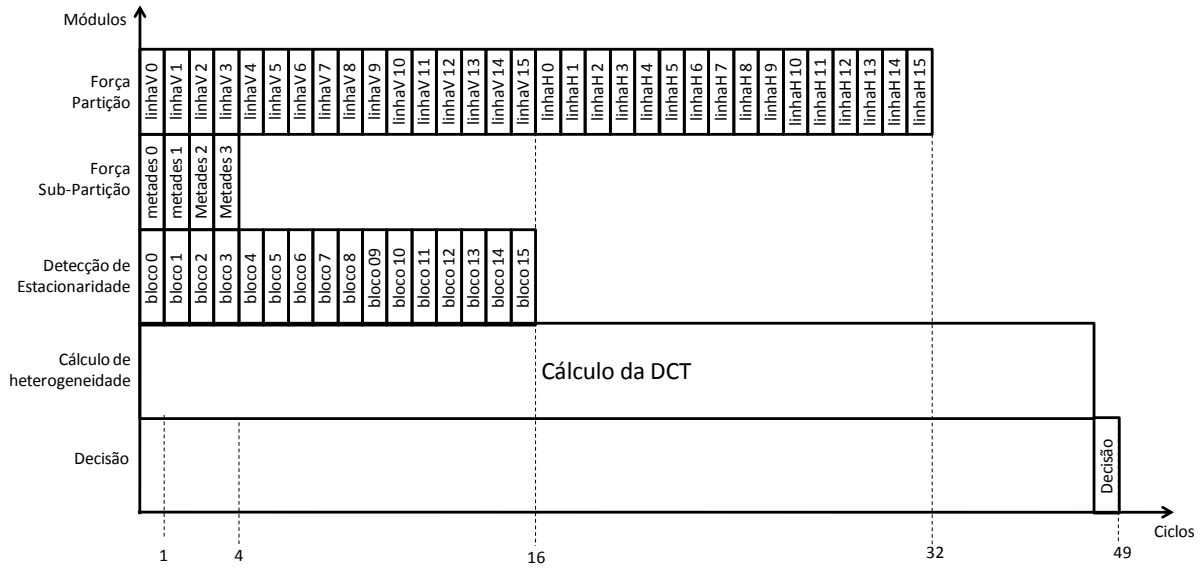

Figura 27 - Diagrama de ciclos para a arquitetura de decisão inter-quadros completa.

O módulo Força de Partição é capaz de processar uma linha por ciclo, ou seja, são necessários 16 ciclos para cada borda (vertical e horizontal), totalizando 32 ciclos para gerar as forças para as duas bordas. O cálculo de Força de Sub-Partição gera quatro forças de metade da borda por ciclo, ou seja, são necessários apenas quatro ciclos para gerar as 16 informações de força de borda mencionadas anteriormente. A Detecção de Estacionariedade é capaz de realizar o SAD entre o bloco

original e o bloco de referência sobre um bloco de tamanho 4x4 por ciclo. Os resultados são acumulados durante 16 ciclos para gerar o nível de estacionariedade do bloco 16x16. Relembrando a Fig. 21, o cálculo de heterogeneidade (*buffer* DCT e DCT 16x16 parcial) necessita de 48 ciclos para terminar o seu processamento. Sendo assim, a decisão sobre um macrobloco inter-quadros é gerada em 49 ciclos (48 ciclos + 1 ciclo para o módulo de decisão final).

A Tab. 17 apresenta os resultados de síntese da arquitetura de decisão inter-quadros completa.

Tabela 17 - Resultados de síntese da arquitetura de decisão inter-quadros completa.

FPGA Stratix II			Standard Cell	
ALUTs	DLRs	Freqüência (MHz)	Área (Número de *Gates*)	Freqüência (MHz)
5.796	2.859	146,56	85.877	142,4

Os recursos de hardware utilizados chegaram a 5796 ALUTs e 2859 DLRs para a síntese em FPGA. Em uma implementação com *standard cells*, foram utilizados 85877 *gates*. A frequência máxima de operação atingida chegou a 142,4 MHz para a síntese em *standard cells*.

Considerando os resultados de síntese apresentados na Tab. 17 e o número de ciclos necessários para processar um macrobloco (apresentado na Fig. 27) a arquitetura desenvolvida para a decisão de modo inter-quadros é capaz de atingir uma taxa de processamento de 366 quadros de resolução HD 1080p (1920x1080 pixels) por segundo. Além disso, para atingir uma taxa de processamento de 30 quadros HD 1080p por segundo, a arquitetura pode operar à uma frequência de 12,0 MHz.

CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho de conclusão de curso apresentou um conjunto de heurísticas para o módulo de decisão de modo do padrão H.264/AVC. O principal objetivo foi o desenvolvimento de heurísticas rápidas para o módulo de decisão de modo, eliminando completamente o uso da técnica RDO para realizar a decisão. As heurísticas foram desenvolvidas e integradas com o software de referência do padrão H.264/AVC com o objetivo de avaliar os resultados. Foram desenvolvidas duas heurísticas para a decisão entre modos intra-quadros e uma para a decisão de modo inter-quadros.

Os resultados, em termos de taxa de bits, qualidade de imagem e complexidade computacional, foram avaliados e comparados com a técnica RDO e com trabalhos relacionados que também tinham como objetivo diminuir a complexidade da decisão de modo. Os resultados mostraram pequenas perdas em taxa de compressão e qualidade da imagem quando comparadas com a técnica RDO. Entretanto, uma elevada redução na complexidade computacional foi atingida pelas heurísticas propostas. Quando comparado com trabalhos relacionados encontrados na literatura, este trabalho apresentou os melhores resultados em termos de diminuição do número de cálculos realizados para a decisão de modo. Após a avaliação dos resultados, as heurísticas foram modeladas e desenvolvidas em hardware, com o objetivo de processar vídeos de alta resolução em tempo real. Foi desenvolvida uma arquitetura para cada um dos dois módulos intra-quadro propostos e outra arquitetura para o módulo de decisão inter-quadros proposto. Os resultados de síntese mostraram que todas as arquiteturas desenvolvidas são capazes de processar vídeos de resolução HD 1080p em tempo real.

Como trabalho futuro pretende-se realizar a verificação das arquiteturas desenvolvidas, bem como a sua integração com os outros módulos do codificador H.264/AVC. Além disso, uma reavaliação sobre a modelagem arquitetural pode gerar arquiteturas que utilizem menos recursos de hardware mantendo o mesmo desempenho obtido neste trabalho.

O desenvolvimento desta monografia gerou resultados satisfatórios que foram submetidos para eventos e revista importantes da área. Como citado anteriormente uma das heurísticas desenvolvidas já foi publicada (CORRÊA, PALOMINO, *et al.*, 2010). Além deste trabalho, o autor desta monografia possui outros trabalhos publicados (ver Apêndice A) que na sua grande maioria, se relacionam com o mesmo assunto desta monografia: arquiteturas para compressão de vídeos.

REFERÊNCIAS

AGOSTINI, L. **Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas a Compressão de Vídeo Segundo o Padrão H.264/AVC**. Universidade Federal do Rio Grande do Sul. Porto Alegre. 2007.

ALTERA. **FPGA, CPLD and ASIC from Altera**. Disponivel em: <http://www.altera.com>. Acesso em: Dezembro 2010.

CHOI, I. Fast Coding Mode Selection With Rate-Distortion Optimization for MPEG-4 part 10 AVC/H.264. **IEEE Transactions on Circuits and Systems for Video Technology**, Dezembro 2006. 1557-1561.

COMPONENTS, A. **TSMC 0.18um 1.8-Volt SAGE-XTM Standard Cell Library Databook.** [S.l.]: [s.n.], 2001.

CORRÊA, G. **Estudo e Desenvolvimento de Heurísticas e Arquiteturas de Hardware para Decisão Rápida do Modo de Codificação de Bloco para o Padrão H.264/AVC**. UFRGS. Porto Alegre, p. 95. 2010.

CORRÊA, G., PALOMINO, D. et al. **Homogeneity and Distortion-Based Intra Mode Decision Architecture for H.264/AVC**. ICECS 2010. Atenas: [s.n.]. 2010.

DINIZ, C. **Arquitetura de Hardware Dedicada para a Predição Intra-Quadro em Codificadores do Padrão H.264/AVC de Compressão de Vídeo**. Universidade Federal do Rio Grande do Sul. Porto Alegre, p. 80. 2009.

FENGQIN, W.; YANGYU, F. **Fast Intra Mode Decision Algorithm in H.264/AVC Using Characteristics of Transformed Coefficients**. 5th International Conference on Visual Information Engineering. [S.l.]: [s.n.]. 2008. p. 245-249.

GHANBARI, M. **Standard Codecs:** Image Compression to Advanced Video Coding. [S.l.]: IEEE Press, 2003.

HANZO, L.; CHERRIMAN, P.; STREIT, J. **Video Compression and Communications**. Second Edition. ed. Southampton: Wiley, 2007.

ISO/IEC. **MPEG Test Model 5, ISO/IEC JTC/SC29/WG11 Document**. [S.l.]. 1993.

ISO/IEC. **MPEG-2 International Standard**. [S.l.]. 1995.

ISO/IEC. **ISO/IEC 14496-2, Coding of audio-visual objects, part 2**. [S.l.]. 1998.

ITU-T. **ITU-T Recommendation H.261, Video Codec for Audiovisual Services at px64 kbit/s**. ITU-T. [S.l.]. 1993.

ITU-T. **ITU-T Recommendation H.263, Video Coding for Low Bit Rate Communication, Draft H.263**. [S.l.]. 1996.

JEON, Y.-I. **Fast Intra Mode Decision Algorithm Using Directional Gradients for H.264/AVC**. International Congress On Image and Signal Processing, CISP. Tianjin: [s.n.]. 2009.

JVT. **Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/AVC or ISO/IEC 14496-10 AVC)**. [S.l.]. 2003.

KIM, H. **Low-Complexity Macroblock Mode Selection For H.264/AVC Encoders**. International Conference on Image Processing, ICIP. Singapura: [s.n.]. 2004.

LEE, Y.-M. **An Improved SATD-Based Intra Mode Decision Algorithm for H.264/AVC**. International Conference on Acoustics, Speech, and Signal Pocessing, ICASSP. Taipei: [s.n.]. 2009. p. 1029-1032.

LIU, P. **A Fast Mode Decision Scheme With Variable Blcok Sizes in H.264/AVC**. IEEE International Conference on Intelligent Systems Design And Applications, ISDA. [S.l.]: [s.n.]. 2008. p. 368-371.

MA, W. **An Efficient Fast Mode Decision Algorithm Based on Motion Cost for H.264/AVC Inter Prediction**. International Symposium on Intelligent Information Technology Application, ITTAW. [S.l.]: [s.n.]. 2008. p. 550-553.

PURI, A.; CHENB, X.; LUTHRA, A. Video Coding Using the H.264/MPEG-4 AVC Compression Standard. **Elsevier Signal Processing Image Communication**, n. 19, 2004. 793-849.

RICHARDSON, I. E. **H.264 and MPEG-4 Video Compression - Video Coding for Next-generation Multimedia**. Aberdeen: Wiley, v. I, 2003.

SBTVD. **ISDTV Standard Draft - Fórum Brasileiro de Televisão Digital**. [S.l.]. 2003.

SHEN, L. Fast Inter Mode Decision Using Spatial Property of Motion Field. **IEEE Transactions on Multimedia**, Outubro 2008. 1208-1214.

SULLIVAN, G.; WIEGAND, T. Rate Distortion Optimization for Video Compression. **IEEE Signal Processing Magazine**, v. 15, p. 74-90, 1998.

SULLIVAN, G.; WIEGAND, T.; STOCKHAMMER, T. Draft H.26L Video Coding Standard for Mobile Applications. **Proceedings of IEEE International Conference on Image Processing**, Thessaloniki, 3, October 2001. 573-576.

SUN, Y. T.; YINYI, L. SATD-Based Intra Mode Decision For H.264/AVC Video Coding. **IEEE International Conference on Multimedia and Expo**, Hannover, 2008. 61-64.

VCEG. **JM Reference Software 16.2**. Disponivel em: <http://iphome.hhi.de/suehring/tml/>. Acesso em: Fevereiro 2010.

WANG, H. **Fast H.264/AVC Inter Mode Decision Based On Inter and Intra Block Conditions**. International Symposium on Circuits and Systems, ISCAS. [S.l.]: [s.n.]. 2007. p. 3647-3650.

WIEGAND, T. et al. Overview of the H.264/AVC Video Coding Standard. **IEEE transactions on Circuits and Systems for Video Technology**, v. 13, p. 560-576, Julho 2003.

WU, D. Fast Inter Mode Decision in H.264/AVC Video Coding. **IEEE Transactions on Circuits and Systems for Video Technology**, 2005. 953-958.

YU, A. **Efficient Block Size Selection Algorithm For Inter-Frame Coding In H.264/AVC/MPEG-4**. International Conference on Acoustics, Speech, and Signal Processing, ICASSP. Montreal: [s.n.]. 2004.

YU, A. Fast Inter Mode Selection in the H.264/AVC Standard Using a Hierarchical Decision Process. **IEEE Transactions on Circuits and Systems for Video Technology**, Fevereiro 2008. 186-195.

APENDICE A – Publicações e prêmios durante a graduação

A.1 Prêmio recebido durante a graduação

CADAthlon Brasil 2010, Universidade Federal do Rio Grande do Sul.

A.2 Trabalhos publicados durante a graduação

A.2.1 Artigos completos publicados em periódicos

SAMPAIO, F., PALOMINO, D. M., DORNELLES, R.S.S., AGOSTINI, L. V. **High Throughput and Low Cost Architecture for the Forward Quantization of the H.264/AVC Video Compression Standard**. CLEI Electronic Journal. , v.13, p.1 - 9, 2010.

DORNELLES, R.S.S., SAMPAIO, F., PALOMINO, D. M., CORRÊA, G., NOBLE, D., AGOSTINI, L. V. **Arquitetura de um Módulo T Dedicado à Predição Intra do Padrão de Compressão de Vídeo H.264/AVC para Uso no Sistema Brasileiro de Televisão Digital**. Hífen (PUCRS. Impresso). , v.82, p.75 - 82, 2008.

A.2.2 Capítulos de livros publicados

PALOMINO, D. M., SAMPAIO, F., CORREA, M., AGOSTINI, L. V. **Investigação e Desenvolvimento de Hardware para as Transformadas do Padrão H.264/AVC de Compressão de Vídeo** In: Desafios e Avanços em Computação - Inovações e Tecnologia, Pelotas: Editora e Gráfica Universitária, 2009, v.1, p. 119-144.

A.2.3 Trabalhos completos publicados em anais de eventos

PALOMINO, D. M., CORRÊA, G., AGOSTINI, L. V. **Desenvolvimento de Heurística Baseada Em Homogeneidade E Distorção Para A Decisão De Modo de Codificação da Predição Intra do Padrão H.264/AVC** In: Congresso de Iniciação Científica, 2010.

CORRÊA, G., PALOMINO, D. M., DINIZ, C., PORTO, R. E. C., AGOSTINI, L. V., BAMPI, Sérgio. **Homogeneity and Distortion-Based Intra Mode Decision Architecture for H.264/AVC** In: IEEE International Conference on Electronics, Circuits, and Systems, 2010, Atenas.

SAMPAIO, F., PALOMINO, D. M., MATTOS, J. C. B., PILLA, M. **Implementação do Escalonador de Tempo Real Rate Monotonic no Sistema Operacional TinyOS In: Southern Conference on Computational Modeling**, 2010, Rio Grande. IV MC Sul, 2010.

PALOMINO, D. M., SAMPAIO, F., AGOSTINI, L. V. **Low Latency and High Throughput Architecture for the H.264/AVC Transforms and Quantization Loop Targeting Intra Prediction** In: South Symposium on Microelectronics, 2010, Porto Alegre, 2010.

DORNELLES, R.S.S., SAMPAIO, F., PALOMINO, D. M., AGOSTINI, L. V. **A Multiplierless Forward Quantization Module Focusing the Intra Prediction Module of the H.264/AVC Standard** In: Simpósio Sul de Microeletrônica, 2009, Pelotas. Editora e Gráfica Universitária, 2009. v.1. p.239 – 242.

SAMPAIO, F., DORNELLES, R.S.S., PALOMINO, D. M., SIEDLER, G. S., CORREA, M., AGOSTINI, L. V. **A Multitransform Architecture for the H.264/AVC Standard And Its Design Space Exploration** In: IEEE International Conference on Electronics Circuits and Systems, 2009, Yasmine Hammamet, 2009. p.711 – 714.

PALOMINO, D. M., SAMPAIO, F., DORNELLES, R.S.S., AGOSTINI, L. V. **An Architecture for a T Module of the H.264/AVC Focusing in the Intra Prediction Restrictions** In: Simpósio Sul de Microeletrônica, 2009, Pelotas. Editora e Gráfica Universitária, 2009. v.1. p.219 – 222.

SAMPAIO, F., PALOMINO, D. M., DORNELLES, R.S.S., AGOSTINI, L. V. **Arquitetura de Alto Desempenho e Baixo Custo para o Módulo de Quantização Direta do Padrão H.264/AVC de Compressão de Vídeo Digita**l In: Conferência Latina Americana de Informática, 2009, Pelotas. Procedings of XXXV CLEI. , 2009.

SAMPAIO, F., PALOMINO, D. M., DORNELLES, R.S.S., AGOSTINI, L. V., BAMPI, Sérgio. **Arquitetura Dedicada para o Loop de Transformadas e Quantização para a Predição Intra-Quadros do Padrão H.264/AVC** In: XV Workshop Iberchip, 2009, Buenos Aires. Ediciones Científicas Americanas, 2009. v.2. p.550 – 555.

SAMPAIO, F., PALOMINO, D. M., DORNELLES, R.S.S., AGOSTINI, L. V. **Dedicated Architecture for the T/Q/IQ/IT Loop Focusing the H.264/AVC Intra Prediction**

In: Simpósio Sul de Microeletrônica, 2009, Pelotas. Editora e Gráfica Universitária, 2009. v.1. p.223 – 226.

PALOMINO, D. M., SAMPAIO, F., DORNELLES, R.S.S., AGOSTINI, L. V. **Low Latency And High Throughput Dedicated Loop Of Transforms And Quantization Focusing** In The H.264/Avc Intra Prediction In: International Conference on Image Processing, 2009, Cairo. New York: IEEE, 2009. v.1. p. 2697 – 2700.

DORNELLES, R.S.S., PALOMINO, D. M., SAMPAIO, F., AGOSTINI, L. V., BAMPI, Sérgio. **Modelos Arquiteturais de Alto Desempenho para as Transformadas 4x4 Diretas e Inversas do Padrão H.264/AVC** In: XV Workshop Iberchip, 2009, Buenos Aires. Buenos Aires: Ediciones Científicas Americanas, 2009. v.1. p.195 – 200.

DORNELLES, R.S.S., SAMPAIO, F., PALOMINO, D. M., AGOSTINI, L. V. **Transforms and Quantization Design Targeting the H.264/AVC Intra Prediction Constraints** In: Symposium On Integrated Circuits And Systems Design, 2009, Natal. Nova Iorque: ACM Digital Library, 2009. p.81 – 86.

SAMPAIO, F., DORNELLES, R.S.S., PALOMINO, D. M., CORRÊA, G., NOBLE, D., AGOSTINI, L. V. **Architectural Templates for the 4X4 Transforms of the H.264/AVC Standard targeting the Intra Prediction** Coder In: SForum 2008 - Student Forum on Microelectronics, 2008, Gramado.

PALOMINO, D. M., CORRÊA, G., DORNELLES, R.S.S., SAMPAIO, F., NOBLE, D., AGOSTINI, L. V. **Implementation and Analysis of Architectures for the 4x4 2-D Forward Hadamard Transform of H.264/AVC** In: CLEI 2008 - XXXIV Conferencia Latinoamericana de Informática, 2008, Santa Fe.

A.2.3 Resumos publicados em anais de eventos

PALOMINO, D. M., SAMPAIO, F., DORNELLES, R.S.S., AGOSTINI, L. V. **Desenvolvimento de Hardware Eficiente para o Loop de Transformadas e Quantização Dedicado à Predição Intra Quadros do Padrão de Compressão de Vídeo H.264/AVC** In: Congresso de Iniciação Cientifica, Pelotas, 2009.

PALOMINO, D. M., SAMPAIO, F., DORNELLES, R.S.S., ROSA JUNIOR, L. S. **Soptmizer: Ferramenta para Otimização de Redes de Transistores Baseada em**

Grafos In: XVIII Congresso de Iniciação Cientifica, Pelotas. Editora Gráfica, 2009.

SAMPAIO, F., PALOMINO, D. M., DORNELLES, R.S.S., AGOSTINI, L. V. **Arquitetura Dedicada para o bloco das Transformadas Diretas do Padrão H.264/AVC de Compressão de Vídeo focando a Predição Intra Quadros** In: Congresso de Iniciação Cientifica, Pelotas, 2008.

SAMPAIO, F., DORNELLES, R.S.S., PALOMINO, D. M., AGOSTINI, L. V. **Exploração do Espaço de Projeto das Transformadas 4x4 Diretas do Padrão H.264/AVC de Compressão de Vídeo** In: Congresso de Iniciação Cientifica, Pelotas, 2008.

PALOMINO, D. M., DORNELLES, R.S.S., SAMPAIO, F., AGOSTINI, L. V. **Implementação e Ánalise de Arquiteturas para as Transformadas Inversas do Padrão H.264/AVC** In: Congresso de Iniciação Cientifica, Pelotas, 2008.

DORNELLES, R.S.S., SAMPAIO, F., PALOMINO, D. M., AGOSTINI, L. V. **Projeto de Hardware para Quantização de Elementos DC com foco na Predição Intra-Quadros do padrão de Compressão de Vídeo do Sistema Brasileiro de Televisão Digitas** In: Congresso de Iniciação Cientifica, Pelotas, 2008.