



COPPE/UFRJ

MARCA D'ÁGUA FRÁGIL E SEMI-FRÁGIL PARA AUTENTICAÇÃO DE
VÍDEO NO PADRÃO H.264

Leandro Pena Moreira de Oliveira

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Eduardo Antônio Barros da
Silva

Rio de Janeiro
Setembro de 2009

MARCA D'ÁGUA FRÁGIL E SEMI-FRÁGIL PARA AUTENTICAÇÃO DE
VÍDEO NO PADRÃO H.264

Leandro Pena Moreira de Oliveira

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
ELÉTRICA.

Aprovada por:

Prof. Eduardo Antônio Barros da Silva, Ph.D.

Prof. Gelson Vieira Mendonça, Ph.D.

Prof. Lisandro Lovisolo, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

SETEMBRO DE 2009

Oliveira, Leandro Pena Moreira de

Marca d'Água Frágil e Semi-Frágil para Autenticação de Vídeo no Padrão H.264/Leandro Pena Moreira de Oliveira.
– Rio de Janeiro: UFRJ/COPPE, 2009.

XIII, 80 p.: il.; 29, 7cm.

Orientador: Eduardo Antônio Barros da Silva

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2009.

Referências Bibliográficas: p. 77 – 80.

1. Marca d'Água.
2. Codificação de Vídeo.
3. H.264. I. Silva, Eduardo Antônio Barros da. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Dedico este trabalho aos meus
pais, Domingos e Dulce, e ao
meu irmão Marcelo.*

Agradecimentos

Gostaria de agradecer a todos que direta ou indireta colaboraram para a realização deste trabalho.

Um agradecimento especial ao meu pai e a minha mãe que me deram todo apoio necessário para que esta obra fosse concluída.

Agradeço também o meu orientador Eduardo pela orientação e confiança ao longo do desenvolvimento desta obra.

Por fim, agradeço a minha namorada, Tatiana, por toda compreensão e apoio.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MARCA D'ÁGUA FRÁGIL E SEMI-FRÁGIL PARA AUTENTICAÇÃO DE VÍDEO NO PADRÃO H.264

Leandro Pena Moreira de Oliveira

Setembro/2009

Orientador: Eduardo Antônio Barros da Silva

Programa: Engenharia Elétrica

Este trabalho propõe dois novos métodos de inserção de marca d'água em dois níveis qualitativos de robustez: Frágil e Semi-Frágil.

Para a proposta de marca d'água frágil é feito um estudo na predição Intra e Inter do padrão H.264 com a finalidade de reduzir o acúmulo de erros (*drift*) provocado pelo método de inserção proposto por Kim, Hong e Won [1]. Os resultados do método de inserção proposto mostraram uma redução no acúmulo de erros à custa de um aumento na complexidade computacional.

Para a marca d'água semi-frágil é proposta a utilização do método de inserção, QIM (*Quantization Index Modulation*), em conjunto com o modelo perceptual de Watson com o objetivo de autenticar tanto imagens quanto vídeos codificados com os padrões de compressão com perda baseados na DCT (JPEG, MPEG-2, H.264). Os resultados do método de inserção proposto mostraram uma boa relação de custo-benefício entre as propriedades: robustez, capacidade e transparência.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

FRAGILE AND SEMI-FRAGILE WATERMARK FOR H.264/AVC VIDEO
AUTHENTICATION

Leandro Pena Moreira de Oliveira

September/2009

Advisor: Eduardo Antônio Barros da Silva

Department: Electrical Engineering

This work proposes two new watermarking methods in two qualitative robustness levels: Fragile and Semi-Fragile.

For the proposal of the fragile watermark, a study in the Intra and Inter prediction of the H.264 standard is made for the purpose of reducing the drift error provoked by the data hiding method published by Kim, Hong e Won [1]. Experimental results for the proposed watermarking method demonstrated a drift reduction to the cost of an increase of the computational complexity.

For semi-fragile watermark, we propose a new QIM(*Quantization Index Modulation*) watermarking algorithm based on Watson perceptual model. This watermarking method aim to authenticate the coded images and videos with the lossy compression standards based on DCT (JPEG, MPEG-2, H.264). Experimental results for the proposed watermarking method demonstrated good trade-off between robustness, capacity and transparency.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xiii
1 Introdução	1
1.1 Organização da Tese	2
2 Marca d'Água Digital	3
2.1 Marca d'Água X Esteganografia	3
2.2 Importância da Marca d'Água Digital	4
2.3 Aplicações	5
2.3.1 Monitoramento de Transmissão	5
2.3.2 Proteção de direitos autorais	6
2.3.3 Rastreamento de Transação	7
2.3.4 Autenticação de Conteúdo	7
2.4 Princípios Básicos de Sistemas de Marca d'Água	7
2.4.1 Propriedades	8
2.4.2 Robustez X Transparência X Capacidade	12
3 O padrão H.264	13
3.1 Introdução	13
3.2 Organização da Camada de Codificação de Vídeo	14
3.3 Visão Geral	15
3.3.1 Codificador	15
3.3.2 Reconstrução	16
3.3.3 Decodificador	17

3.4	Predição	17
3.4.1	Predição Intra	18
3.4.2	Predição Inter	19
3.5	Compensação de Movimento Estruturada em Árvore	20
3.6	Precisão de Um Quarto de Amostra na Compensação de Movimento .	21
3.7	Multiplas Referências Temporais	22
3.8	Filtro Antiblocagem Adaptativo	24
3.9	Estimação de Movimento	24
3.10	Transformada e Quantização	25
3.10.1	Transformada	25
3.10.2	Quantização	25
3.11	Codificação de Entropia	25
3.11.1	CAVLC	26
4	Marca d'Água Frágil no <i>Bitstream</i> do Padrão H.264 no Perfil <i>Baseline</i>	30
4.1	Marca d'Água no <i>bitstream</i> do padrão MPEG-2	31
4.2	Informação oculta no <i>bitstream</i> do H.264 no perfil <i>Baseline</i>	33
4.2.1	Análise do CAVLC no H.264	34
4.2.2	Método de Inserção	34
4.3	Estudo do Acúmulo de Erros na Predição Intra e Inter do Padrão H.264	35
4.3.1	Explorando a Predição Intra	36
4.3.2	Explorando a Predição Inter	37
4.4	Marca d'água frágil para autenticação de vídeo no padrão H.264 no perfil <i>Baseline</i>	39
4.4.1	Método de inserção da marca d'água	40
4.4.2	Criação da marca d'água	41
4.5	Resultados e Discussões	42
4.5.1	Resultados no Quadro Intra do GOP de Vídeo	43
4.5.2	Resultados no Quadro Inter do GOP de Vídeo	45
4.5.3	Resultados nos Quadros Intra e Inter do GOP de Vídeo	45

5	Marca d'Água Semi-Frágil Adaptativa para autenticação dos padrões de compressão com perda baseados na DCT (JPEG, MPEG-2 e H.264)	49
5.1	Modelo perceptual de Watson	50
5.1.1	Modelo perceptual de Watson para blocos 4x4	52
5.2	QIM	53
5.3	Sistema de Marca d'Água Proposto	54
5.3.1	Insersor QIM	55
5.3.2	Seleção dos coeficientes	55
5.3.3	Detector QIM	58
5.3.4	Criação da Marca d'Água	58
5.3.5	Autenticação	59
5.4	Resultados e Discussões	59
5.4.1	Resultados para o codificador JPEG	60
5.4.2	Resultados para o codificador H.264-Intra	65
5.5	Conclusão	67
6	Conclusão	75
	Referências Bibliográficas	77

Lista de Figuras

2.1	Sistema de marca d'água genérico.	8
3.1	Organização hierárquica da camada de codificação de vídeo.	14
3.2	Diagrama do codificador do H.264.	16
3.3	Diagrama do decodificador do H.264.	17
3.4	Modos de predição Intra 4x4.	18
3.5	Modos de predição Intra 16x16.	19
3.6	Possíveis blocos para predição Inter.	21
3.7	Multiplas referências temporais.	23
3.8	Quadros de referência para o quadro B.	23
4.1	Conceito de domínio do <i>bit</i>	32
4.2	Conceito de domínio do coeficiente.	32
4.3	Método de compensação do acúmulo de erros.	33
4.4	Blocos 4x4 que podem utilizar o bloco P para predição.	36
4.5	Mapa de predição de um quadro P.	38
4.6	Etapas do algoritmo de marca d'água frágil proposto.	39
4.7	Etapas do algoritmo de autenticação de vídeo.	41
4.8	Total de blocos 4x4 Intra com T1s.	43
4.9	Total de blocos 4x4 Intra não utilizados para predição Intra.	43
4.10	PSNR - Modificando N blocos Intra com T1s no quadro Intra.	44
4.11	Total de blocos Inter com T1s.	45
4.12	PSNR - Modificando em cada GOP 1024 blocos Inter com T1s.	46
4.13	Total de blocos com T1s no GOP.	46
4.14	PSNR - Modificando 1024 blocos com T1s no GOP.	47
4.15	<i>Quadro P após a inserção da marca d'água.</i>	48

5.1	Sistema de Marca d'Água Proposto	54
5.2	Programa desenvolvido para elaboração do método de inserção pro- posto.	56
5.3	BER do método de inserção com robustez de $QF = 10$ e $TBloco = 0$.	61
5.4	BER do método de inserção com robustez de $QF = 20$ e $TBloco = 0$.	61
5.5	BER do método de inserção com robustez de $QF = 40$ e $TBloco = 0$.	62
5.6	BER do método de inserção com robustez de $QF = 40$ e $TBloco = 0$, 1 e 2.	63
5.7	Imagem original, imagem marcada e imagem codificada pelo JPEG. . .	64
5.8	Blocos utilizados na detecção e a diferença entre a imagem original e a imagem marcada.	65
5.9	BER do sistema de marca d'água com robustez de $QP = 30$ e $Tbloco$ $= 0$	69
5.10	BER do sistema de marca d'água com robustez de $QP = 33$ e $Tbloco$ $= 0$	70
5.11	BER do sistema de marca d'água com robustez de $QP = 36$ e $Tbloco$ $= 0$	71
5.12	Capacidade dos vídeos marcados em diferentes níveis de robustez. . .	72
5.13	PSNR dos vídeos marcados em diferentes níveis de robustez.	73
5.14	Quadro original, quadro marcado e quadro codificado pelo H.264. . .	74
5.15	A diferença entre o quadro original e o quadro marcado mutiplicada por 15.	74

Lista de Tabelas

2.1	<i>Aplicações de marcas d'água em vídeos.</i>	5
3.1	<i>Tipos de MB permitidos em cada fatia</i>	15
3.2	<i>Modos de Predição Intra 16x16 luma</i>	19
3.3	<i>Modos de predição Intra 16x16 croma</i>	20
3.4	<i>Escolha da tabela de consulta para o coeff_token</i>	27
3.5	<i>Limites para incrementar a tabela VLC</i>	28
4.1	<i>Características dos parâmetros de codificação do CAVLC</i>	34
4.2	<i>Modos de predição Intra4x4 que não utilizam o bloco P para predição</i>	37
4.3	<i>Modos de predição Intra16x16 que não utilizam o bloco P para predição</i>	37
5.1	<i>Parâmetros do método de inserção de marca d'água, QIM-Watson.</i>	59
5.2	<i>Capacidade do sistema de marca d'água</i>	63
5.3	<i>PSNR das imagens marcadas em relação à imagem original.</i>	65
5.4	<i>Métrica de Watson das imagens marcadas em relação à imagem original.</i>	65

Capítulo 1

Introdução

Um dos efeitos da disponibilidade cada vez maior de ferramentas de processamento de sinais para modificar o conteúdo multimídia (sem deixar qualquer rastro de modificação perceptível), é a perda de credibilidade da informação digital, visto que tais ferramentas colocam em dúvida a autenticidade de conteúdos multimídia.

Uma promissora solução para superar tal problema é a autenticação dos dados por meio da tecnologia de marca d'água digital. A principal vantagem da marca d'água digital em relação ao método de autenticação por assinatura digital é a inserção da marca d'água diretamente na obra original (imagem, vídeo e áudio), logo não necessitando adicionar uma informação extra num cabeçalho atrelado à obra.

É importante notar que a marca d'água por ser inserida diretamente na obra original sofrerá as mesmas distorções que a obra marcada. Dessa forma, diferentes níveis qualitativos de robustez são requeridos de acordo com a aplicação pretendida.

Normalmente, as aplicações de autenticação utilizam-se de marcas d'água frágeis ou semi-frágeis para verificar a integridade da obra. A marca d'água frágil por ser afetada por qualquer modificação na obra, provê uma autenticação exata. A marca d'água semi-frágil por não ser afetada pelas modificações legítimas, mas destruída pelas modificações ilegítimas, provê um mecanismo de autenticação seletiva.

Assim, este trabalho propõe dois novos métodos de inserção de marca d'água em dois níveis qualitativos de robustez, frágil e semi-frágil, com o intuito de autenticar vídeos codificados no padrão H.264.

1.1 Organização da Tese

Esta seção descreve cada capítulo desta tese.

No capítulo 2 será feita uma revisão sobre os sistemas de marca d'água abordando sua importância, aplicações, princípios básicos e propriedades.

No capítulo 3 serão apresentados alguns conceitos da camada de codificação de vídeo do H.264 relevantes para este trabalho, por exemplo, o codificador de entropia CAVLC, a transformada DCT inteira, os tipos de predição intra e algumas características da predição inter como a compensação de movimento estruturada em árvore com precisão de um quarto de amostra, o filtro antiblocagem adaptativo e a utilização de múltiplas referências temporais.

No capítulo 4 será feito um estudo na predição Intra e Inter do padrão H.264 de modo a elaborar algoritmos para a redução do acúmulo de erros(*drift*) provocado pelo método de inserção de informação oculta proposto em [1].

No capítulo 5 será proposta uma marca d'água semi-frágil adaptativa para autenticação dos padrões de compressão com perda baseados na DCT (JPEG, MPEG-2, H.264). O sistema de marca d'água proposto utiliza o método de inserção, QIM(*Quantization Index Modulation*) [2], em conjunto com o modelo perceptual de Watson [3] com a finalidade de garantir a presença da marca d'água até um determinado nível de compressão com perda.

No capítulo 6 será apresentada as conclusões dos resultados obtidos nos dois sistemas de marca d'água propostos.

Capítulo 2

Marca d'Água Digital

2.1 Marca d'Água X Esteganografia

Na área de ocultação de informação os campos de marca d'água e esteganografia estão tão estreitamente relacionados que acabam se sobrepondo e compartilhando muitas abordagens técnicas. Abaixo seguem suas definições.

Marca d'água é um sinal perceptível ou imperceptível que é inserido numa obra original (imagem, áudio, vídeo e texto), com o intuito de, posteriormente, na detecção do sinal fornecer informações sobre a obra marcada. A marca d'água pode ser categorizada em visível ou invisível.

A esteganografia é a arte de comunicar-se secretamente, ocultando uma mensagem sigilosa dentro de outra informação sem importância, de maneira que não exista forma de detectar que há uma mensagem escondida. [4]

De modo a entender as diferenças filosóficas desses campos será utilizado um exemplo clássico de esteganografia, a história de *Herodotus*, que conta do escravo enviado pelo seu mestre, *Histiæus*, para a cidade de Ionian em Mileto com uma mensagem secreta tatuada em seu corpo cabeludo. O escravo depois de tatuado esperou o seu cabelo crescer, assim ocultando a mensagem e viajou até Mileto. Em Mileto, o escravo teve seu cabelo raspado, revelando a mensagem para o tirano *Aristágoras*. Assim, essa mensagem encorajou *Aristágoras* a iniciar uma revolta contra o rei persa.

Nesse exemplo a existência da mensagem é secreta e não está relacionada ao escravo, levando então a definição de esteganografia. Mas, se imaginarmos que a

mensagem escrita na cabeça do escravo fosse, "Este escravo pertence ao *Histiaeus*", nesse caso a existência da mensagem continua oculta e está relacionada ao escravo levando tanto à definição de marca d'água invisível quanto à de esteganografia.

Além desses dois casos, existe um terceiro caso em que a informação está relacionada à obra, porém não oculta a sua existência, levando a definição de marca d'água visível. Apesar de a marca d'água poder ser categorizada como visível ou invisível, neste trabalho serão apenas consideradas marcas d'água invisíveis tanto em imagens quanto em vídeos.

2.2 Importância da Marca d'Água Digital

Os recentes avanços tecnológicos dos computadores e a difusão da Internet estão tornando a edição, reprodução e distribuição de material multimídia cada vez mais fáceis e rápidos. No entanto, estes avanços colocam em foco questões relacionadas à segurança, como a proteção dos direitos autorais.

Uma tecnologia que pode ser usada na proteção dos direitos dos proprietários de conteúdo é a criptografia. A criptografia pode proteger o conteúdo em trânsito, mas uma vez decifrado, o conteúdo não está mais protegido. Dessa maneira, surge uma necessidade de uma alternativa ou complemento para a criptografia, uma tecnologia que possa proteger o conteúdo mesmo depois de decifrado. Para preencher estas necessidades a tecnologia de marca d'água digital mostra grande potencial, visto que visa-se embutir a informação no conteúdo de modo a não ser possível removê-la através de operações normais (compressão, mudanças de formato do arquivo e outros tipos de processamento).

Apesar da proteção de direitos autorais ser o principal impulsionador para as pesquisas no campo de marca d'água digital, existem outras aplicações onde a marca d'água digital pode ser utilizada ou sugerida. Algumas das aplicações de marcas d'água em vídeos são resumidas na Tabela 2.1 e apresentadas extensivamente na seção 2.3 e em [5].

Tabela 2.1: *Aplicações de marcas d'água em vídeos.*

Aplicação	Proposta da inserção de marca d'água
Proteção de Direitos Autorais	Prova de propriedade
Rastreamento de Transação	Rastreamento de Compradores Piratas
Monitoração de Transmissão	Identificar quando e onde o vídeo está sendo transmitido
Autenticação de Conteúdo	Assegurar que o material original não foi alterado

2.3 Aplicações

A preocupação em relação à proteção dos direitos autorais do conteúdo digital e também o grande potencial comercial da tecnologia de marca d'água digital provocaram um aumento no interesse pela tecnologia na última década. Esta seção é dedicada à apresentação de várias aplicações em que a marca d'água digital pode adicionar valor.

2.3.1 Monitoramento de Transmissão

Há organizações e indivíduos interessados na monitoração de transmissão. Por exemplo, os fornecedores de comerciais querem garantias que o conteúdo publicitário foi transmitido em determinado horário e na íntegra, e as emissoras de televisão querem garantias que o conteúdo transmitido pela estação base está sendo retransmitido pelas estações afiliadas. De modo a atender essas organizações e indivíduos há diferentes métodos de monitoramento de transmissão.

O método de monitoramento de transmissão mais primitivo é ter observadores que assistam o conteúdo transmitido e gravem o material exibido. Este método é caro e propenso a erros. Por isso, há técnicas para automação da monitoração que são divididas em duas categorias:

- Sistema de monitoração passivo : Consiste em um computador que compara os sinais recebidos com um banco de dados com obras conhecidas, tentando diretamente reconhecer o conteúdo transmitido, de forma a simular os observadores.
- Sistema de monitoração ativo : Associa uma informação ao conteúdo trans-

mitido de modo a identificá-lo. Um exemplo de implementação de um sistema ativo é colocar uma informação de identificação numa área separada do sinal transmitido, por exemplo, colocar um código de identificação no intervalo vertical de apagamento (VBI) do sinal de vídeo. No entanto, a inserção de informação no VBI pode ser problemática, pois nos Estados Unidos da América, o distribuidor não tem obrigação legal de manter qualquer outra informação, a não ser a informação do *Closed Caption*, no intervalo vertical de apagamento (VBI) do sinal de vídeo [6]. Conseqüentemente, o código de identificação inserido no VBI não possui garantia que será transmitido.

Os sistemas de marca d'água digital são métodos alternativos para inserção de uma informação de identificação em sistemas de monitoração ativa. A principal vantagem é que a informação está inserida dentro do próprio conteúdo e não se precisa explorar um segmento particular do sinal transmitido. As desvantagens são que o processo de inserção não é trivial e a marca d'água pode degradar a qualidade do conteúdo multimídia.

2.3.2 Proteção de direitos autorais

Proteção de direitos autorais é historicamente o primeiro objetivo da marca d'água digital. A estratégia consiste em inserir uma marca d'água para identificar o proprietário legítimo do conteúdo multimídia. Se uma cópia ilegal for encontrada o proprietário legítimo pode provar através da marca d'água que a obra é de sua autoria e conseqüentemente, o malfeitor pode responder pelo ato ilegal no tribunal.

No entanto, este cenário perfeito é perturbado pelos malfeitores. Por exemplo, se o malfeitor inserir uma segunda marca na obra, ambos os interessados, o legítimo proprietário e o malfeitor podem alegar a propriedade da obra. Para evitar tal problema, o algoritmo de inserção da marca d'água não deve ser invertível. Isto evitaria que o malfeitor conseguisse produzir uma falsa obra marcada.

Além disso, o malfeitor pode utilizar algumas técnicas de processamento de sinais, de modo a remover a marca d'água e manter a obra pouco distorcida. Desse modo os sistemas de marca d'água para aplicações de prova de propriedade devem ser robustos e seguros. A definição de uma marca d'água robusta e segura é descrita na Seção 2.4.1, na descrição de robustez.

2.3.3 Rastreamento de Transação

Para localizar a origem da cópia pirata de uma obra, em cada cópia da obra original é inserida uma marca d'água diferente. Quando uma cópia não autorizada é encontrada, a extração da marca d'água indica qual cópia da obra original gerou a cópia pirata, assim identifica-se quem distribuiu ilegalmente a obra original marcada.

2.3.4 Autenticação de Conteúdo

Autenticação de conteúdo é uma das principais aplicações de marca d'água digital. Em contraste com as outras aplicações, como proteção de direitos autorais e rastreamento de transação onde as aplicações são projetados para suportar diversas formas de distorção, o objetivo da autenticação de conteúdo é verificar a integridade da obra, detectando alguma possível manipulação.

Em sistemas de autenticação é sempre necessária uma referência para verificar a autenticidade da obra marcada. Por exemplo, em aplicações tradicionais de autenticação de mensagem, assinaturas digitais são habitualmente usadas como referências. O receptor reproduz um resumo da obra recebida e compara com o resumo gerado com a obra original. Se os resumos estiverem iguais, a obra será considerada autêntica, caso contrário, haverá um indício de que a obra foi modificada. Na autenticação com marca d'água, a própria serve como referência. A marca d'água é incorporada na obra, submetendo-se às mesmas transformações que a obra sofrerá. Assim, a obra sendo modificada, a marca d'água também será alterada. Portanto, a integridade da obra pode ser verificada comparando a marca d'água extraída com a marca d'água original.

2.4 Princípios Básicos de Sistemas de Marca d'Água

A Figura 2.1 mostra um sistema de marca d'água genérico, representado por um diagrama em blocos. Nele, a entrada do sistema é formada pela obra original, a marca d'água e uma chave opcional que pode ser pública ou secreta. A chave pode ser usada para aumentar a segurança e prevenir que pessoas não autorizadas

decodifiquem a marca d'água e alterem ou destruam a marcação.

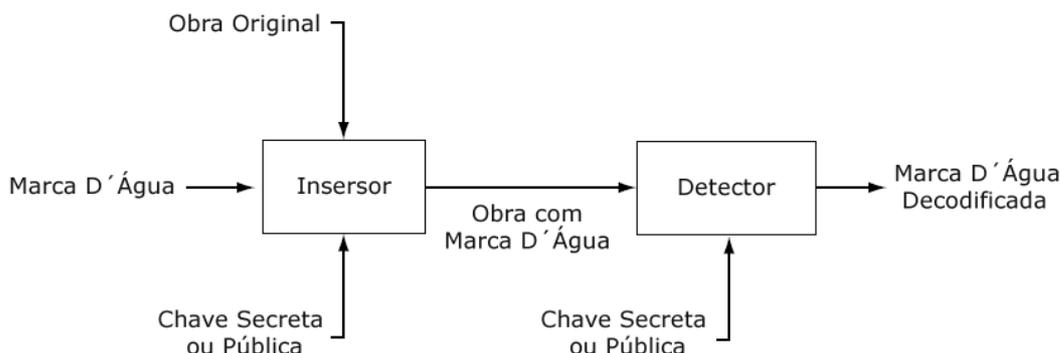


Figura 2.1: Sistema de marca d'água genérico.

Basicamente, os sistemas de marca d'água são formadas por um estágio de inserção e um estágio de detecção. Para inserir uma informação da marca d'água numa obra original, as técnicas de inserção de marca d'água provocam pequenas alterações visualmente imperceptíveis na obra original. Posteriormente, a informação da marca d'água pode ser recuperada a partir da obra marcada através da detecção da presença dessas pequenas alterações.

Os sistemas de marca d'água podem ser caracterizados por algumas propriedades. A relativa importância de cada propriedade é dependente dos requisitos de cada aplicação. Na próxima subsecção comentaremos sobre essas propriedades: transparência, capacidade, robustez, segurança, sistemas públicos e privados e detecção cega e não-cega.

2.4.1 Propriedades

Detecção cega e não-cega

Em algumas aplicações a obra original está disponível durante a detecção, como na aplicação de rastreamento de transações em que o proprietário da obra original é quem utiliza o detector para descobrir quem distribuiu ilegalmente a obra.

Em outras aplicações, a detecção deve ser executada sem o acesso à obra original, como na aplicação de autenticação de conteúdo. Nela, o detector é distribuído para o público de modo que o usuário verifique a integridade do vídeo originário de uma fonte que esteja sendo redistribuído por outras fontes.

O detector que requer acesso à obra original é chamado de detector não-cego. Este termo também se refere aos detectores que requerem alguma informação derivada da obra original. De modo oposto, o detector que não requer qualquer informação relacionada à obra original é chamado de detector cego.

Na prática, sistemas de marca d'água com detecção não-cega conseguem um desempenho melhor em relação aos sistemas de marca d'água com detecção cega. Isso acontece porque a obra marcada pode sofrer alterações não-maliciosas, maliciosas e do próprio processo de inserção da marca d'água que acabam interferindo no processo de detecção. Isto torna a obra original uma melhor referência do que a obra marcada durante a detecção.

Marca d'água pública e privada

A marca d'água é chamada de privada, quando somente usuários autorizados puderem detectar a marca d'água. Em contraste, a marca d'água é chamada de pública, quando qualquer usuário puder detectar a marca d'água.

Transparência

A transparência em sistemas de marca d'água refere-se à similaridade perceptual entre a obra original e a obra marcada. A imperceptibilidade da marca d'água é uma das características mais importantes na tecnologia de marca d'água digital, visto que artefatos introduzidos pelo processo de inserção de marca d'água não são desejáveis e podem reduzir ou destruir o valor comercial da obra marcada. Dessa maneira, alguns métodos de marca d'água chamados adaptativos [7], exploram o sistema visual humano para distribuir a energia da marca d'água de modo que a marca não ultrapasse o limiar de percepção. Para avaliar a transparência do vídeo marcado há alguns conhecidos algoritmos de métricas de qualidade, como PSNR, VQM [8] e SSIM [9].

Capacidade

O termo refere-se à quantidade de dados que a marca d'água insere por obra ou por unidade de tempo. Em aplicações de áudio e vídeo também é conhecido por taxa de dados. O requerimento da taxa de dados varia significativamente dependendo da

aplicação.

Robustez

A robustez é representada pela resistência da marca d'água às manipulações sofridas na obra marcada. Estas manipulações podem ser não-maliciosas ou maliciosas. As não-maliciosas não visam explicitamente eliminar a marca d'água, enquanto, as manipulações maliciosas têm o intuito de danificar a marca d'água.

O nível exato de robustez de uma marca d'água não pode ser especificado sem considerar a sua aplicação. A maioria das situações encontradas na prática englobam-se em quatro níveis qualitativos de robustez, estes descritos em [10] [6], como:

1. **Marca d'Água Segura:** Definida principalmente para aplicações de proteção de direitos autorais e de rastreamento de transação, onde a marca d'água deve sobreviver às manipulações maliciosas e não-maliciosas. Neste caso, a perda da informação oculta somente aconteceria quando ocorresse uma perda significativa na qualidade da obra marcada, de modo a torná-la sem valor. Nas manipulações maliciosas assume-se que o atacante conhece os algoritmos de inserção e detecção da marca d'água, assim possibilitando a elaboração de estratégias específicas para a remoção da marca d'água. As manipulações não-maliciosas incluem uma variedade de operações de processamento de sinais, por exemplo, compressão com perda, filtragem linear e não-linear, conversão D/A e A/D, ruído aditivo e distorções geométricas (rotação, translação, interpolação e decimação). Na literatura há alguns trabalhos com o propósito de obter uma marca d'água segura [11] [12].
2. **Marca d'Água Robusta:** Neste caso requer-se que a marca d'água seja resistente somente contra manipulações não-maliciosas. As aplicações de marca d'água robusta incluem todas as situações em que as manipulações não tenham a intenção de remover a marca d'água. No cenário de suas aplicações, a marca d'água deve resistir ao uso conjunto de inúmeros tipos de manipulações não-maliciosas.
3. **Marca d'Água Semi-Frágil:** Neste caso a marca d'água não é afetada pelas

distorções legítimas, mas destruída pelas distorções ilegítimas. Assim, um mecanismo de autenticação seletiva é desenvolvido. Se a distinção entre legítima e ilegítima estiver baseada pela perceptibilidade, a criação da marca d'água semi-frágil é similar à da marca d'água robusta. Afinal o objetivo da marca d'água robusta é garantir que esta sobreviva às manipulações na obra marcada até o ponto no qual as modificações a tornem sem valor. Para a marca d'água semi-frágil pode se garantir que esta sobreviva as manipulações depois do ponto onde a obra marcada perdeu seu valor, mas também assegura-se a remoção da marca d'água acima de determinado ponto. Isto pode ser conseguido ajustando a robustez da marca d'água de modo que seja destruída se a distorção exceder um determinado nível. Na literatura, há alguns trabalhos utilizando marca d'água semi-frágil para a aplicação de autenticação de conteúdo [13] [14].

4. **Marca d'Água Frágil:** A marca d'água é considerada frágil se qualquer modificação na obra marcada remover a informação oculta. A principal aplicação é na autenticação, onde qualquer alteração da obra marcada evidência a ocorrência de adulteração.

Segurança

A segurança de uma marca d'água refere-se à habilidade de resistir a ataques hostis. Um ataque hostil é qualquer processo com pretensões de destruir a proposta da marca d'água. Os tipos de ataques encaixam-se em três categorias: remoção, detecção e inserção não-autorizada. A relativa importância desses ataques depende da aplicação.

A remoção não autorizada refere-se aos ataques que visam evitar a detecção da marca d'água. Uma interessante forma de remoção não autorizada é conhecida de ataque de colisão. O malfeitor obtém várias cópias de uma obra, cada uma com diferentes marcas e então combina as obras marcadas para produzir uma cópia sem marca d'água. Essa é uma grande preocupação na aplicação de rastreamento de transação, visto que cada cópia da obra original possui uma marca d'água diferente.

A inserção não autorizada refere-se ao ato de inserir marcas d'água ilegítimas em obras que não deveriam contê-las. Por exemplo, na aplicação de autenticação

não existe a preocupação do malfeitor em tornar a obra marcada em não detectável, pois esta obra será considerada adulterada. Entretanto, se o malfeitor tem a habilidade de executar uma inserção não autorizada, pode levar o detector a considerar, erroneamente como autêntica, uma obra adulterada.

A detecção não autorizada ocorre quando o malfeitor consegue detectar e/ou decifrar a mensagem inserida na obra marcada. Em geral, esta forma de ataque é mais preocupante em sistemas de esteganografia do que em sistemas de marca d'água.

2.4.2 Robustez X Transparência X Capacidade

Vale ressaltar que a maioria dos sistemas de marca d'água estabelecem uma relação de custo-benefício entre três propriedades: transparência, robustez e capacidade. Por exemplo, uma marca d'água bastante robusta pode ser obtida, a partir de grandes modificações na obra original. Se a marca d'água tiver que carregar muita informação haverá uma intensa degradação na obra, no caso dela carregar pouca informação, a distorção não será tão intensa. Dessa forma, o nível de robustez da marca d'água limita a quantidade de dados a ser inserida para que não ocorra uma grande distorção na obra.

Capítulo 3

O padrão H.264

3.1 Introdução

O padrão H.264 já foi amplamente discutido na literatura [15], [16], [17]. Muitas publicações ilustram o desempenho do H.264/AVC, comparando-o a outros codificadores de vídeo como o MPEG-2 e o H.263. Todos os resultados apontam para um aumento de desempenho de, pelo menos, o dobro em relação aos padrões MPEG2 e H.263 [18] [19].

O padrão recomendado pelo ITU-T em conjunto com o ISO/MPEG para o H.264, seguindo a tradição dos padrões anteriores, apenas especifica a decodificação de uma seqüência de dados (*bits*). A vantagem é que isso oferece mais liberdade no projeto dos codificadores, cuja única restrição é que gerem seqüências de dados decodificáveis.

Um dos conceitos fundamentais do H.264 é a separação do projeto em duas distintas camadas, chamadas de VLC (Video Coding Layer - Camada de Codificação de Video) e NAL (Network Abstraction Layer). A VLC é responsável pela eficiente representação do conteúdo do vídeo, enquanto a NAL é responsável pelo empacotamento da informação codificada de uma maneira apropriada para transmissão ou armazenamento.

Neste capítulo, serão abordados alguns conceitos da camada de codificação de vídeo relevantes para este trabalho, por exemplo, o codificador de entropia CAVLC, a transformada DCT inteira, os tipos de predição intra e algumas características da predição inter como a compensação de movimento estruturada em árvore com

precisão de um quarto de amostra, o filtro antiblocagem adaptativo e a utilização de múltiplas referências temporais.

3.2 Organização da Camada de Codificação de Vídeo

A estrutura de codificação de vídeo do H.264 é parecida com a dos padrões anteriores, sendo formada por GOP, quadro, fatia, macrobloco e bloco, ilustrados na Figura 3.1.

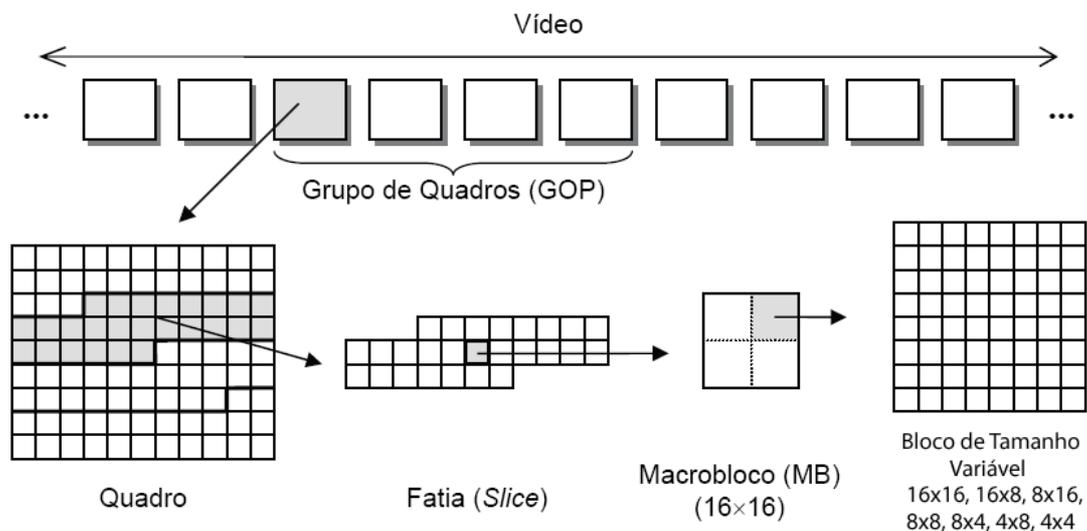


Figura 3.1: Organização hierárquica da camada de codificação de vídeo.

O macrobloco (MB) é a principal unidade do H.264, sendo formado por 16x16 amostras no canal *luma*(Y) e 8x8 amostras no canal *croma*(Cb e Cr) para o formato de vídeo 4:2:0. Estes são classificados de acordo com o modo de codificação ao qual são submetidos, podendo ser do tipo I, P e B.

Macroblocos I são aqueles cuja predição é do modo Intra. Esta ocorre a partir de amostras decodificadas na fatia atual.

Macroblocos P são aqueles cuja predição é do modo Inter. Esta ocorre a partir de um dos N quadros codificados, decodificados e reconstruídos chamados quadros de referência.

Macroblocos B são aqueles cuja predição é do modo Inter. Esta ocorre a partir de um ou dois quadros de referência.

As fatias, estruturas que agrupam os macroblocos, são classificados pelos tipos

I, P, B, SP e SI. Cada tipo de fatia pode conter diferentes modos de codificação de macrobloco, mostrados na Tabela 3.1. Por fim, um quadro codificado consiste de uma ou mais fatias.

Tabela 3.1: *Tipos de MB permitidos em cada fatia*

Tipo de Fatia	Tipos de MB permitidos
I	I
P	I ou P
B	I ou B
SP	P
SI	I

3.3 Visão Geral

O padrão H.264 realiza os processos de predição, transformação, quantização e codificação de entropia para produzir um fluxo de dados comprimido (*bitstream*). Praticamente os mesmos processos dos padrões anteriores são usados, no entanto, importantes mudanças ocorrem nos detalhes da implementação e no funcionamento de cada processo.

O diagrama do codificador do H.264 é mostrado na Figura 3.2. Nas próximas subseções são detalhados os principais passos da codificação e decodificação de um macrobloco no modo Intra e no modo Inter.

3.3.1 Codificador

O codificador processa o quadro de vídeo em unidades que consistem de macroblocos. Cada macrobloco é codificado no modo Intra ou Inter, onde para cada bloco, a predição, definida pela *Pred* na Figura 3.2, é baseada nas amostras dos quadros reconstruídos.

No modo Intra, a *Pred* é formada por amostras da fatia atual que já foram codificadas, decodificadas e reconstruídas. É importante notar que essas amostras uF'_n na Figura 3.2 não utilizam o filtro de antiblocagem para gerar a *Pred*.

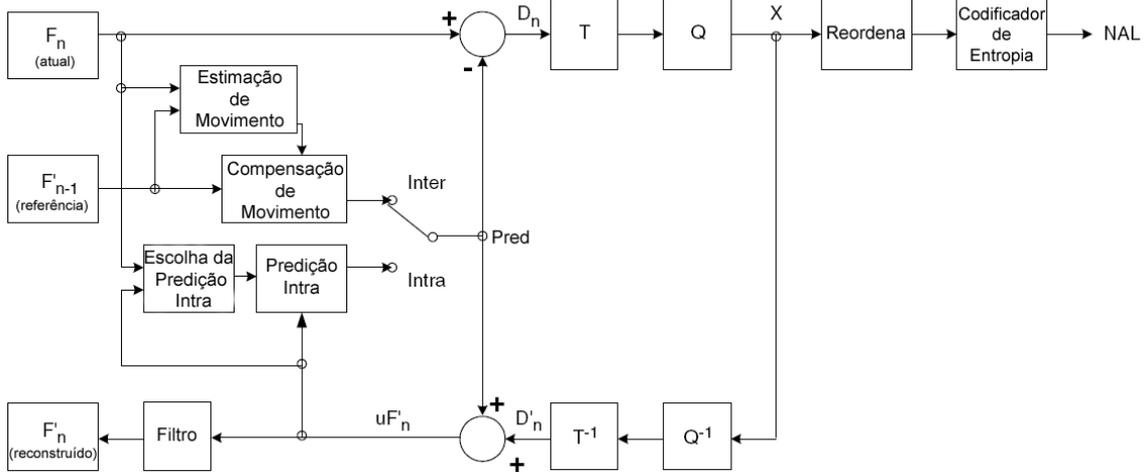


Figura 3.2: Diagrama do Codificador do H.264.

No modo Inter, a $Pred$ é formada pela estimação de movimento de um ou dois quadros de referência selecionados de uma ou duas listas de quadros de referência denominadas lista 0 e lista 1. Para predição no modo Inter o quadro de referência pode ser escolhido de quadros passados ou futuros (segundo a ordem de exibição) que já tenham sido codificados e reconstruídos. Na Figura 3.2, o quadro de referência é mostrado como F'_{n-1} .

Seguindo o fluxo da Figura 3.2, a $Pred$ é subtraída pelo bloco atual para produzir um bloco de resíduo D_n que é transformado e quantizado resultando em X , assim esse conjunto de coeficientes transformados e quantizados são reordenados e em seguida codificados pelo codificador de entropia. Os coeficientes codificados junto com a informação adicional (modos de predição, parâmetros de quantização, vetores de movimento e outros) requerida para decodificar cada bloco formam o fluxo de dados comprimido (*bitstream*).

3.3.2 Reconstrução

Os coeficientes X são escalonados (Q^{-1}) e transformados inversamente (T^{-1}) para produzir um bloco de diferença D'_n . A predição do bloco ($Pred$) é somada a D'_n para criar o bloco reconstruído uF'_n . Em seguida, somente para predição Inter, um filtro é aplicado para reduzir os efeitos de blocagem e o quadro de referência reconstruído é criado de uma série de blocos F'_n .

3.3.3 Decodificador

O decodificador recebe o fluxo de dados comprimido que passa pelo decodificador de entropia para produzir um conjunto de coeficientes quantizados, definidos por X na Figura 3.3. Estes são escalonados e transformados inversamente gerando D'_n . Utilizando a informação do cabeçalho decodificada do fluxo de dados comprimido, o decodificador cria a predição do bloco ($Pred$), que deve ser idêntica à $Pred$ gerada pelo codificador. Em seguida, a $Pred$ é somada a D'_n para produzir amostras uF'_n que são filtradas para criar cada bloco decodificado F'_n .

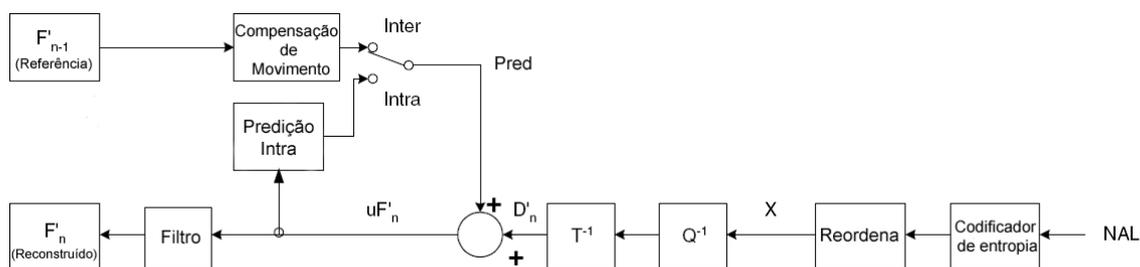


Figura 3.3: Diagrama do Decodificador do H.264.

Até aqui foi descrito o processo de codificação e decodificação no H.264. Na próxima seção são apresentadas algumas das novas técnicas do padrão H.264 relacionadas à predição Intra e Inter.

3.4 Predição

De modo a obter uma codificação de vídeo eficiente, os codificadores utilizam técnicas de predição, com o objetivo de remover a redundância espacial e temporal existente no vídeo.

No H.264, para explorar a redundância espacial, as amostras são preditas a partir das amostras da fatia corrente que já tenham sido codificadas e reconstruídas. Para explorar a redundância temporal, as amostras são preditas de quadros anteriores já codificados e reconstruídos (quadros de referência) presentes em uma ou na duas listas empregáveis.

Vale ressaltar que o codificador baseia suas predições nas amostras codificadas e decodificadas da imagem ao invés das amostras do quadro de vídeo original. Isso visa assegurar que as predições do codificador e do decodificador sejam idênticas.

3.4.1 Predição Intra

Nos quadros Intra dos padrões de compressão de vídeo anteriores, a DCT era aplicada diretamente nos blocos da imagem, assim resultando em quadros Intra codificados com grande quantidade de dados. De modo a aumentar a eficiência no processo de codificação dos quadros Intra, o H.264 passou a explorar a correlação espacial entre os macroblocos vizinhos pertencentes à mesma fatia. Esse procedimento é baseado no fato de que estatisticamente os macroblocos adjacentes tendem a ter amostras similares no canal *luma* e no canal *croma*.

Para amostras do canal *luma*, a predição intra é formada por blocos 4x4 ou 16x16 ou 8x8 (que foi introduzida com as extensões *FRExt* [18]). Existem nove modos de predição para cada bloco 4x4 e quatro modos de predição para cada bloco 16x16.

Modos de Predição Intra 4x4 *luma*

Os nove modos de predição Intra 4x4 são ilustrados na Figura 3.4.

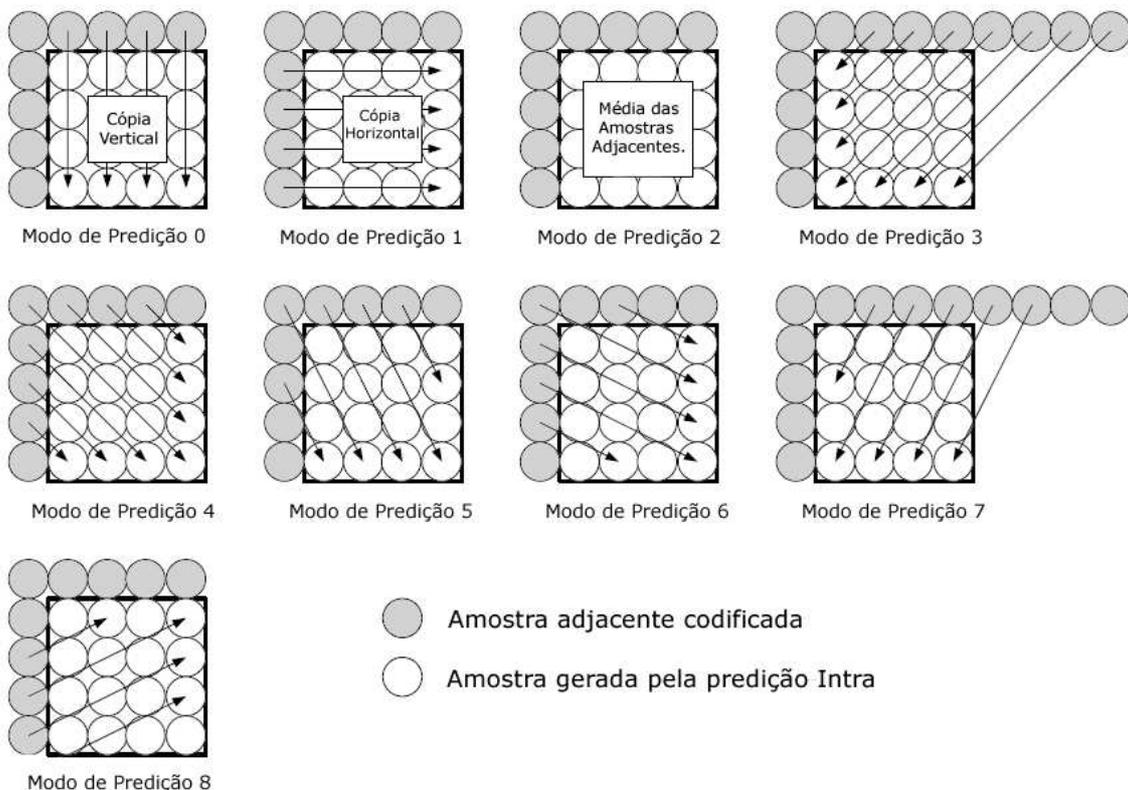


Figura 3.4: Modos de Predição Intra 4x4.

Modos de Predição Intra 16x16 *luma*

Os quatro modos de predição Intra 16x16 são comentados na Tabela 3.2 e ilustrados na Figura 3.5.

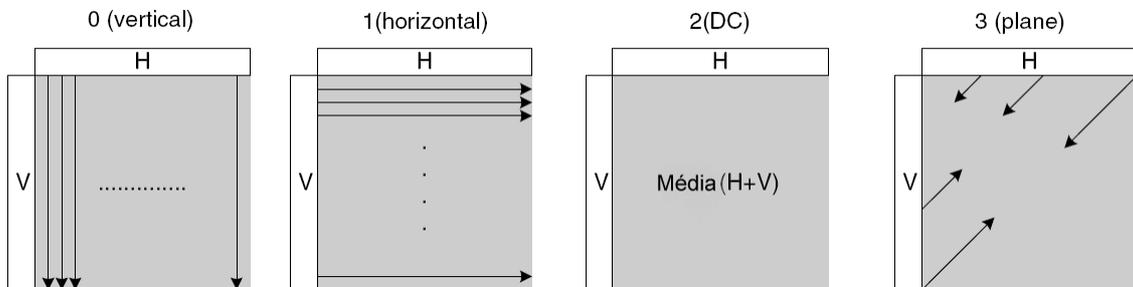


Figura 3.5: Modos de Predição Intra 16x16.

Tabela 3.2: *Modos de Predição Intra 16x16 luma*

Modo 0	Extrapolação a partir das amostras superiores.
Modo 1	Extrapolação a partir das amostras da esquerda.
Modo 2	Média das amostras superiores e da esquerda.
Modo 3	Função linear das amostras superiores e da esquerda.

Para cada bloco, o codificador seleciona o modo de predição que minimiza uma medida de distorção entre a predição e o bloco corrente a ser codificado. Normalmente, nas regiões planas os modos de predição 16x16 são escolhidos, enquanto em regiões de textura os modos de predição 4x4 são selecionados.

Para amostras do canal *chroma*, a predição é formada por blocos com tamanho dependente do formato de vídeo, mas os modos de predição são parecidos com os modos de predição 16x16 *luma* mostrados na Figura 3.5. A exceção é a diferente numeração dos modos de predição, visualizados na Tabela 3.3.

3.4.2 Predição Inter

No domínio temporal, existe uma alta correlação entre os quadros de vídeo vizinhos. O objetivo do processo de predição Inter é reduzir a redundância temporal no vídeo, através da busca de similaridades entre quadros de vídeo vizinhos.

Tabela 3.3: *Modos de predição Intra 16x16 croma*

Modo 0	DC.
Modo 1	HORIZONTAL
Modo 2	VERTICAL
Modo 3	PLANAR

Resumidamente, o processo de predição Inter inicia-se pela criação de uma predição de um MB partir de um ou mais quadros de referência. Então, esta predição é subtraída do macrobloco atual resultando em um MB residual, e quanto menor a energia deste, maior é a precisão do processo de predição.

Como nos padrões de compressão de vídeo anteriores, a base na exploração da redundância temporal no H.264 é feita pelo método de estimação e compensação de movimento baseada em bloco [15]. Isso acontece devido à sua relativa simplicidade na implementação e também a sua fácil adaptação às transformadas de imagem baseadas em blocos.

Nas próximas subseções apresentaremos algumas das novas técnicas adicionadas no padrão H.264, como por exemplo, múltiplas referências temporais, compensação de movimento estruturada em árvore com precisão de um quarto de amostra e o filtro antiblocagem adaptativo.

3.5 Compensação de Movimento Estruturada em Árvore

A energia residual de um bloco é reduzida pela compensação de movimento à medida que se diminui o tamanho do bloco, entretanto um tamanho de bloco cada vez menor conduz a um aumento nas operações de busca e na quantidade de vetores de movimento que precisam ser transmitidos. Então, em blocos menores, a quantidade de vetores de movimento a ser transmitida pode reduzir o benefício fornecido pela menor energia nos blocos residuais.

Dessa forma, uma estratégia eficiente é adaptar o tamanho do bloco às características da imagem. Por exemplo, a escolha de blocos maiores para regiões ho-

mogêneas e a escolha de blocos menores para regiões com detalhes e com movimentos complexos.

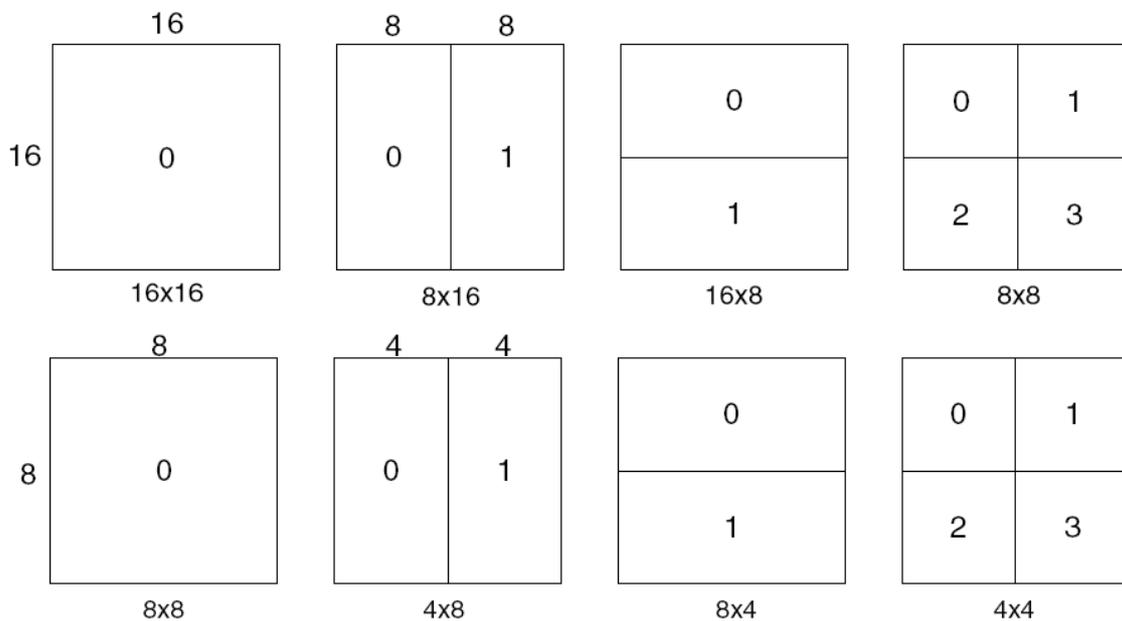


Figura 3.6: Possíveis Blocos para Predição Inter.

O padrão H.264 utiliza um método que provê uma flexibilidade na partição dos macroblocos chamada de compensação de movimento estruturada em árvore. A Figura 3.6 apresenta três possíveis formas em que um bloco 16x16 e um bloco 8x8 podem ser divididos, com intuito de mostrar o grande número de combinações que estas partições possibilitam em um macrobloco.

Vale ressaltar que para cada bloco do macrobloco é requerido um vetor de movimento separado. Assim, é necessário codificar tanto o tipo de bloco escolhido quanto o vetor de movimento.

3.6 Precisão de Um Quarto de Amostra na Compensação de Movimento

Em cenas de um vídeo, raramente o movimento dos objetos encaixa-se perfeitamente num grid fixo $M \times N$, os objetos podem mover-se de frações de amostras entre os quadros, por exemplo, 2.35 amostras ao invés de 2.0 amostras na direção horizontal. Dessa forma, uma melhor predição pode ser formada pela interpolação de regiões do

quadro de referência, antes da busca pelo macrobloco residual com menor energia.

A maioria dos padrões anteriores ao H.264 permitia apenas uma interpolação com precisão de meia amostra, entretanto, o padrão H.264 utiliza um maior refinamento, através da interpolação com precisão de um quarto de amostra. Em geral, um maior refinamento na interpolação provê um melhor desempenho (menor energia residual) na compensação de movimento, às custas de um aumento na complexidade. Além da complexidade, outro ponto desfavorável é o aumento da quantidade de dados (*bits*) requeridos para representar os passos da interpolação, visto que ocorre um aumento no número de posições de busca.

Portanto, quanto mais precisa for a compensação de movimento, mais dados (*bits*) são requeridos para representar o vetor de movimento e menos dados para representar o macrobloco residual. Assim, para compensação de movimento menos acurada, uma menor quantidade de dados é requerida para representar o vetor de movimento e mais dados para representar o macrobloco residual. Com base nessa relação de custo-benefício, o padrão H.264 optou pela interpolação com precisão de um quarto de amostra.

Os valores nas posições de meia-amostra são obtidas pela aplicação de um filtro FIR separável com pesos $(1/32, -5/32, 5/8, 5/8, -5/32, 1/32)$ nas amostras inteiras em cada dimensão [15]. Uma vez que todas as meia-amostras estão disponíveis, as posições de um quarto de amostra são produzidas pela interpolação linear das meia-amostras e das amostras inteiras.

3.7 Múltiplas Referências Temporais

Os padrões mais antigos permitiam utilizar como referência apenas o quadro anterior, no caso do quadro P, e tanto o quadro anterior quanto o posterior, no caso do quadro B. Uma nova arquitetura se estendeu para os padrões H.263++ e H.264, permitindo que o codificador escolha durante a estimação de movimento um quadro de referência dentre uma lista de N quadros previamente codificados. A Figura 3.7 ilustra essa situação.

No caso do quadro B, o codificador e o decodificador gerenciam duas listas contendo N quadros passados e futuros (segundo a ordem de exibição). Cada lista

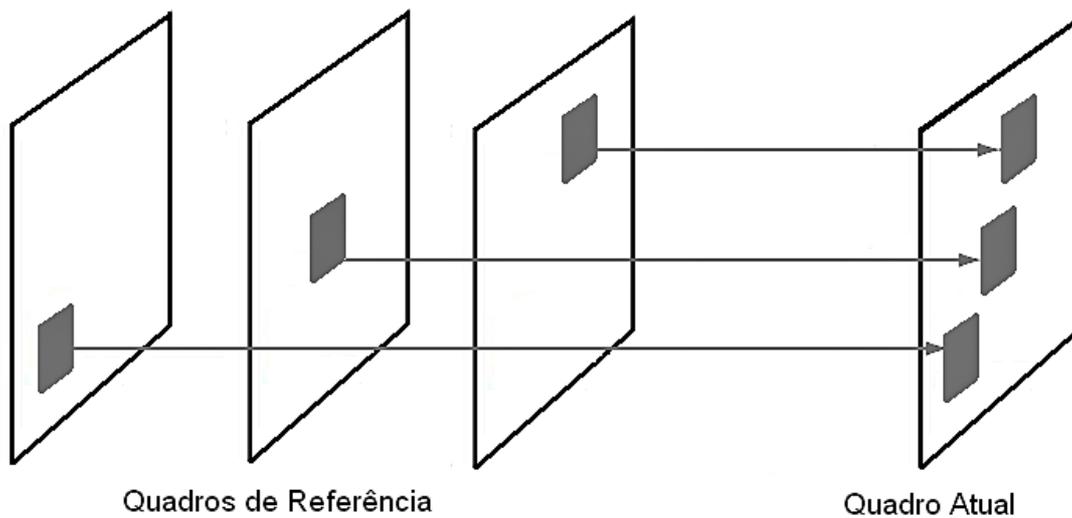


Figura 3.7: Múltiplas Referências Temporais.

fornece um quadro de referência podendo formar um quadro passado e um quadro futuro ou dois quadros futuros ou dois quadros passados, como ilustrado na Figura 3.8.

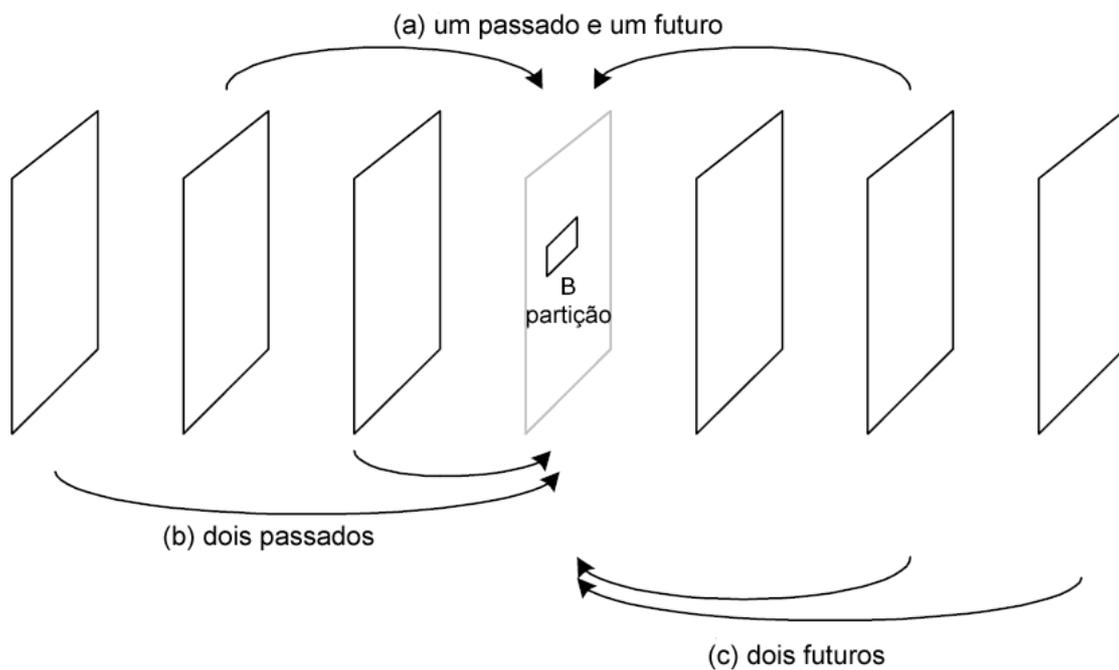


Figura 3.8: Quadros de Referência para o Quadro B.

Esse aumento do espaço de busca temporal possibilita uma melhor predição inter, no entanto, aumenta a complexidade e os requisitos de memória à medida que cresce

a quantidade de quadros de referência.

3.8 Filtro Antiblocagem Adaptativo

Uma característica particular da codificação baseada em blocos é a distorção causada pela visibilidade de estruturas de bloco chamada de blocagem. A blocagem é considerada um dos artefatos mais presentes nos métodos de codificação baseados em blocos. Por essa razão, o H.264 insere um filtro antiblocagem diretamente no processo de reconstrução, com o intuito de melhorar a qualidade dos quadros reconstruídos, conseqüentemente, melhorar a qualidade da predição. A aplicação desse filtro apresenta uma redução significativa na blocagem, conseqüentemente uma melhora na qualidade subjetiva do vídeo [20]. A descrição detalhada do filtro antiblocagem adaptativo pode ser encontrada em [20] [15].

3.9 Estimação de Movimento

Devido ao espaço de busca ser muito grande, na literatura há vários algoritmos de estimação de movimento com diferentes níveis de complexidade computacional [21] [22]. No codificador de referência do H.264 (JM), o algoritmo padrão de estimação de movimento pode ser resumido nas seguintes etapas. Em primeiro lugar, o vetor de movimento predito define uma janela de busca de $M \times M$ amostras nos N quadros de referência. Dentro dessa janela realiza-se buscas pelo vetor de movimento que minimiza a medida de distorção SAD (*Sum of Absolute Differences*) [15] dentre todos os possíveis vetores de movimento com posições inteiras. Em seguida, ocorre uma interpolação com precisão de um quarto de amostra na área em torno do vetor encontrado. Então, nessa área interpolada busca-se o vetor de movimento que minimiza a medida de distorção SATD (*Sum of Absolute Transformed Differences*) [15].

3.10 Transformada e Quantização

3.10.1 Transformada

Dependendo do tipo de resíduo que está sendo codificado, o H.264 utiliza uma transformada diferente. Por exemplo, a transformada Hadamard 4x4 para coeficientes DC no canal *luma*, a transformada Hadamard 2x2 para coeficientes DC no canal *chroma* e a transformada DCT 4x4 para todos os outros resíduos.

A transformada DCT opera em blocos de resíduos 4x4 gerados a partir da predição por compensação de movimento ou da predição Intra. Essa transformada DCT do H.264 apresenta algumas melhorias:

1. A transformada é inteira, ou seja, todas as operações podem ser executadas utilizando aritmética de inteiros, sem perda de precisão na decodificação.
2. A transformada pode ser implementada utilizando somente adições e deslocamentos.

3.10.2 Quantização

A quantização dos coeficientes da transformada é controlada pelo QP (*Quantization Parameter*). Este parâmetro é responsável por controlar o compromisso entre qualidade da imagem reconstruída e a taxa de dados de saída, através de 52 níveis de quantização. O incremento de um nível de QP equivale ao aumento de aproximadamente 12.5% no passo de quantização.

3.11 Codificação de Entropia

H.264 especifica dois métodos alternativos de codificação por entropia: uma técnica de baixa complexidade baseada no CAVLC (*Context Adaptive Variable Length Coding*) [15] e um algoritmo que demanda um maior poder computacional chamado de CABAC (*Context-Based Adaptive Binary Arithmetic Coding*) [23]. Comparando com os outros codificadores de entropia dos padrões de vídeo anteriores, ambos os métodos apresentam melhorias em termos de eficiência na codificação.

3.11.1 CAVLC

Este método é usado para codificar os resíduos, os coeficientes da transformada dos blocos (4x4 ou 2x2) ordenados em *zig-zag*. O CAVLC é projetado para se beneficiar de várias características dos blocos 4x4 quantizados:

- Depois da predição, transformação e quantização, os blocos tipicamente contêm majoritariamente zeros. O CAVLC utiliza a codificação por carreira (*run length coding*) para representar as seqüências de zeros de maneira compacta.
- Os coeficientes não-nulos de maior freqüência são usualmente seqüências de +1 e -1, denominados *trailing ones* ou T1s. O CAVLC codifica o número de T1s de uma maneira compacta.
- Existe uma correlação estatística entre as quantidades de coeficientes não-nulos em blocos vizinhos. Assim, o número de coeficientes não-nulos é codificado escolhendo-se uma tabela VLC dependendo das quantidades de coeficientes não-nulos em blocos vizinhos.
- A magnitude dos coeficientes não-nulos decresce na direção das freqüências mais altas. Então, a escolha da tabela VLC para a magnitude depende das magnitudes dos níveis recém codificados.

O CAVLC codifica os blocos com coeficientes da transformada seguindo o seguinte procedimento:

1. **Codifica o número de coeficientes não-nulos e *trailing ones* (*Coeff_token*)**

O primeiro VLC, *coeff_token*, codifica o número de coeficientes não-nulos e o número de *trailing ones*(T1s), coeficientes com valores iguais a +1 e -1. O número de coeficientes não-nulos pode assumir valores entre 0 e 16. O número de *trailing ones* pode assumir valores entre 0 e 3. No caso de existir mais de 3 T1s, somente os últimos 3 são tratados como *trailing ones*, os outros são codificados como coeficientes normais.

Para codificar o *coeff_token* existem quatro tabelas de consultas: Num-VLC0,

Num-VLC1, Num-VLC2 e Num-FLC.

A escolha da tabela depende do parâmetro N , dado por:

$$N = (N_U + N_L)/2 \quad (3.1)$$

N_U : Número de coeficientes não-nulos no bloco imediatamente superior ao bloco sendo codificado.

N_L : Número de coeficientes não-nulos no bloco imediatamente à esquerda do bloco sendo codificado.

O parâmetro N determina a tabela de consulta segundo a Tabela 3.4.

Tabela 3.4: Escolha da tabela de consulta para o `coeff_token`

N	Tabela para o <code>coeff_token</code>
0, 1	Num-VLC0
2, 3	Num-VLC1
4, 5, 6, 7	Num-VLC2
8 ou mais	FLC

2. Codifica o sinal de cada T1

Para cada T1 (+/-), um único bit codifica o sinal (0 = +, 1 = -). Estes são codificados em ordem reversa, começando com o T1 de maior frequência.

3. Codifica os valores dos coeficientes não-nulos remanescentes (*Level*)

O valor (sinal e magnitude) de cada coeficiente não-nulo remanescente é codificado em ordem reversa. A escolha da tabela VLC para codificar cada valor se adapta dependendo da magnitude de cada valor já codificado. Existem sete tabelas a escolher mostradas na Tabela 3.5.

A escolha da tabela VLC se adapta da seguinte forma:

- A tabela pode ser inicializada com VLC0 ou VLC1. No caso de existir mais que 10 coeficientes não-nulos e menos que 3 T1s inicia-se com VLC1, caso contrário, inicia-se com VLC0.
- O coeficiente não-nulo de mais alta frequência é codificado. Se a magnitude deste coeficiente é maior que um limite pré-definido, a tabela esco-

Tabela 3.5: *Limites para incrementar a tabela VLC*

Tabela VLC atual	Limite para incrementar tabela
VLC0	0
VLC1	3
VLC2	6
VLC3	12
VLC4	24
VLC5	48
VLC6	Não se aplica (Tabela mais alta)

lhida é a próxima tabela VLC. Os limites pré-definidos são mostrados na Tabela 3.5.

4. **Codifica o número total de zeros antes do último coeficiente (*Total-Zeros*)**

A soma de todos os zeros que precedem o mais alto coeficiente não-nulo no vetor reordenado é codificada com um VLC separado.

5. **Codifica cada seqüência de zeros**

O número de zeros precedendo cada coeficiente não-nulo (*run_before*) é codificado na ordem reversa (do coeficiente de maior freqüência para o de menor freqüência). O parâmetro *run_before* é codificado para cada coeficiente não-nulo, começando pelo de maior freqüência. Esta codificação só não é necessária em dois casos:

- Se não existem mais zeros à esquerda para codificar.
- Se o coeficiente não-nulo em questão é o de menor freqüência (último coeficiente não-nulo na ordem reversa).

O VLC para cada seqüência de zeros é escolhido dependendo do par *Zero-Left/run_before*, com *ZerosLeft* consistindo no número total de zeros à esquerda do coeficiente não-nulo e *run_before* consistindo no número de zeros imediatamente à esquerda do coeficiente não nulo.

Em [15] há mais detalhes e exemplos de codificação e decodificação de um bloco utilizando o CAVLC.

Capítulo 4

Marca d'Água Frágil no *Bitstream* do Padrão H.264 no Perfil *Baseline*

Talvez, o H.264 seja o padrão de codificação mais popular nas transmissões sem fio e entre as mídias que circulam na Internet. Logo, há uma necessidade de autenticação desse conteúdo de vídeo, já que o fácil acesso, transmissão e manipulação de vídeos permitem uma rápida distribuição de vídeos adulterados. Assim, este capítulo propõe um sistema de marca d'água frágil para autenticação de vídeos codificados no padrão H.264 no perfil *Baseline*.

O perfil *Baseline* é indicado para sistemas com poucos recursos e direcionado principalmente para aplicações móveis e de videoconferência. Então, sistemas de marca d'água para essas aplicações devem ter como premissa uma baixa complexidade computacional. Esse requisito impossibilita que as etapas completas de decodificação e codificação de vídeo sejam realizadas no processo de inserção de uma marca d'água num vídeo codificado.

Assim, os processos de inserção e detecção de marca d'água devem ser executados diretamente no domínio da compressão. Quando a inserção da marca d'água é no domínio da compressão há dois possíveis cenários. O primeiro cenário é o da inserção da marca d'água no *bitstream* do vídeo codificado, que é abordado neste capítulo, e outro cenário é a inserção da marca d'água durante a codificação do vídeo.

O sistema de marca d'água frágil proposto neste capítulo insere a marca d'água

no *bitstream* do vídeo codificado através de melhorias sugeridas para o método de inserção de informação oculta no bitstream do H.264 proposto por Kim, Hong e Won [1].

Este capítulo apresenta, na Seção 4.1, alguns métodos de inserção de marca d'água no *bitstream* do padrão MPEG-2. Na Seção 4.2 apresenta-se o método de inserção proposto em [1] que embute uma informação oculta no bitstream do padrão H.264 no perfil *Baseline*. Na Seção 4.3 é feito um estudo nas predições Intra e Inter do padrão H.264 para reduzir o acúmulo de erros provocados pelo método de inserção proposto em [1]. A Seção 4.4 sugere um sistema de marca d'água frágil para autenticação de vídeo no padrão H.264 no perfil *Baseline*. Por último na Seção 4.5, compara-se o sistema de marca d'água proposto com o método de inserção proposto em [1].

4.1 Marca d'Água no *bitstream* do padrão MPEG-2

Apesar da inserção de marca d'água proposta neste capítulo ocorrer no *bitstream* do padrão H.264, esta seção tem por objetivo situar o leitor em relação aos problemas e soluções da inserção de marca d'água no *bitstream* de um vídeo codificado no padrão MPEG-2.

Os algoritmos de inserção de marca d'água em *bitstreams* MPEG-2 operam normalmente em dois diferentes níveis de compressão: no domínio do *bit* ou no domínio do coeficiente.

No domínio do *bit*, a marca d'água é aplicada diretamente no *bitstream* do vídeo precisando-se apenas processar os blocos do decodificador de entropia e do inseridor da marca d'água, ilustrados na Figura 4.1. Este procedimento consiste basicamente da decodificação VLC (*Variable Length Coding*) seguida da modificação do VLC. Este método é computacionalmente eficiente e pode ser encontrado em [24] [25]. A técnica proposta modifica o *bit* menos significativo de alguns coeficientes quantizados de modo a não alterar a taxa de *bits* do vídeo.

Já no domínio do coeficiente, ocorre um aumento na complexidade, visto que o processo de inserção da marca d'água precisa utilizar algumas etapas da codificação



Figura 4.1: Conceito de domínio do *bit*.

e decodificação de vídeo para conseguir manipular os coeficientes da DCT. Dentre as etapas de codificação e decodificação de vídeo estão o decodificador e o codificador de entropia, a técnica de *run-level*, o escalonamento e a quantização dos coeficientes da DCT, este processamento é ilustrado na Figura 4.2. Esse conceito de inserção pode ser visto com mais detalhes em [26].



Figura 4.2: Conceito de domínio do coeficiente.

O maior problema de modificar os coeficientes da DCT no domínio do *bit* ou no domínio do coeficiente é o chamado acúmulo de erros (*drift*). No *bitstream* do vídeo, predições feitas a partir de quadros passados são usadas para reconstruir o quadro atual, e este pode servir como referência para futuras predições. Então, um erro causado pela modificação no *bitstream* de um quadro pode alterar os próximos quadros que o utilizam para predição. Logo, se todos os quadros são alterados, os erros dos quadros anteriores e do quadro atual podem se acumular e resultar em artefatos visuais no vídeo.

O artigo [26] propõe uma técnica para compensar o acúmulo de erros, assim evitando a propagação dos erros pelos quadro P e B do padrão MPEG. Entretanto, a complexidade computacional dessa solução limita sua aplicabilidade. A técnica de compensação do acúmulo de erros ocorre pela adição de um sinal representado pela diferença da predição no *bitstream* original e da predição no *bitstream* com a marca d'água.

A desvantagem dessa técnica é o aumento da complexidade computacional cau-

sada pelas operações da DCT e da decodificação completa do MPEG requeridas para o cálculo do sinal de compensação do acúmulo de erros. A Figura 4.3 ilustra esse aumento de complexidade.

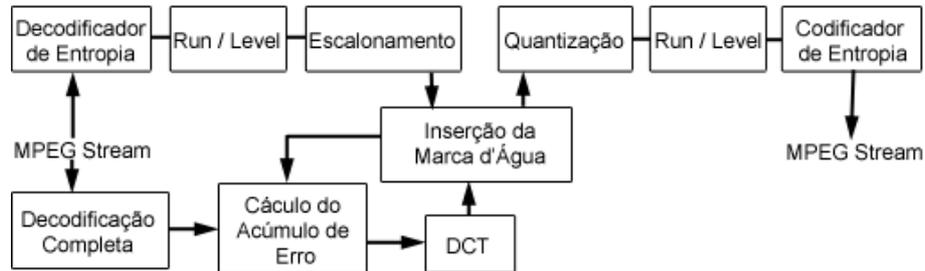


Figura 4.3: Método de compensação do acúmulo de erros (*drift*).

4.2 Informação oculta no *bitstream* do H.264 no perfil *Baseline*

Para obter uma codificação de vídeo mais eficiente, novas técnicas foram adotadas no padrão H.264, como predição espacial direcional, compensação de movimento com tamanho de bloco variável e a codificação de comprimento variável adaptativa baseada em contexto. Essas novas técnicas impossibilitam que métodos convencionais de inserção de marca d'água implementados no MPEG-2 sejam diretamente aplicados no *bitstream* dos vídeos codificados através do H.264.

Dessa forma, novos métodos estão sendo propostos. O artigo [1] propõe a inserção de informações ocultas diretamente no *bitstream* do vídeo codificado pelo H.264 no perfil *Baseline*. Esse método é descrito com detalhes nesta seção, pois forma a base do sistema de marca d'água para autenticação de vídeo proposto neste capítulo.

Os algoritmos de inserção de informação oculta no *bitstream* do vídeo tipicamente buscam uma menor degradação visual e a não alteração da taxa de *bits* do vídeo, visto que uma intensa degradação visual no vídeo inviabilizaria uma inserção efetiva e a alteração da taxa de *bits* do vídeo pode provocar problemas em transmissões sobre canais de taxa de *bits* fixa, como o estouro de memória dos dispositivos decodificadores, ou a possibilidade da perda de sincronização entre áudio e o vídeo [1] [25].

4.2.1 Análise do CAVLC no H.264

Nesta seção o codificador de entropia CAVLC, descrito na seção 3.11.1, é analisado para determinar qual o melhor parâmetro de codificação a ser modificado que satisfaça os requisitos de menor degradação visual e não alteração da taxa de bits do vídeo.

Tipicamente, os coeficientes de mais alta frequência são os menos importantes visualmente, assim, são os que introduzem uma menor de degradação visual ao serem modificados. Já para não alteração da taxa de bits do vídeo é preferível alterar os parâmetros de codificação de tamanho fixo.

A tabela 4.1 mostra as características dos parâmetros de codificação do CAVLC em relação à frequência dos coeficientes e da sua representação em *bits*.

Tabela 4.1: *Características dos parâmetros de codificação do CAVLC*

Parâmetros de codificação	Valores codificados	Representação em bits	Frequência dos coeficientes
Coeff_token	TC e T1	Variável	Todas
Trailing ones	Sinais dos T1s	Fixa	Alta Frequência
Level	Coeficientes não-nulos excluindo os T1s	Variável	Média e Baixa Frequência
TotalZeros	Número de zeros antes do último coeficiente não-nulo	Variável	*
Run_before	Número de zeros precedendo cada coeficiente não-nulo	Variável	*

4.2.2 Método de Inserção

O bit que representa o sinal dos *trailing ones* satisfaz os requisitos de menor distorção visual e taxa de bits do vídeo inalterada, dessa forma sendo o parâmetro de codificação escolhido para inserir uma informação oculta.

Os bits dos sinais de T1 serão definidos como S_{i1} , S_{i2} , S_{i3} , pois três é número máximo de T1s num bloco. O S_{i1} é o bit que representa o sinal do primeiro T1 na ordem inversa do zig-zag do i -ésimo bloco da DCT, ou seja, o sinal da componente com maior frequência entre os coeficientes com valores diferentes de zero. Dessa forma, S_{i1} , por ser o sinal do coeficiente com maior frequência, este é o bit escolhido

para a inserção de um bit definido como b_i , através do seguinte procedimento:

$$S'_{i1} = \begin{cases} \bar{S}_{i1}, & \text{se } ((b_i \neq S_{i1}) \text{ e } (k \geq 1)) \\ S_{i1}, & \text{Nos demais pontos} \end{cases} \quad (4.1)$$

Onde \bar{S}_{i1} representa o complemento do sinal do bit original S_{i1} e a quantidade de T1s é definida por k . Como resultado da inserção, se existir pelo menos um T1, o S'_{i1} será equivalente ao bit b_i .

O algoritmo de detecção é simplesmente expresso como:

$$\hat{b}_i = \hat{S}_{i1} \quad \text{se } (k \geq 1) \quad (4.2)$$

Onde \hat{b}_i representa o bit extraído do i^{th} bloco da DCT e \hat{S}_{i1} é o sinal do primeiro T1 recebido.

Este método de inserção possui pontos favoráveis como: a grande capacidade para inserção de informação oculta, a baixíssima complexidade computacional e a não alteração da taxa de *bits* do vídeo. Um ponto desfavorável é em relação à degradação visual, visto que os blocos com *bits* modificados provocam acúmulo de erros (*drift*) pelo vídeo, principalmente nos quadros Intra onde o acúmulo de erros pode se propagar tanto espacialmente, no quadro Intra que está sendo codificado, quanto temporalmente, nos quadros Inter a serem codificados. Devido a essa desvantagem, a próxima seção apresenta um estudo para a redução do acúmulo de erros a partir das técnicas de predição Intra e Inter empregadas no padrão H.264.

4.3 Estudo do Acúmulo de Erros na Predição Intra e Inter do Padrão H.264

Nos padrões de compressão H.264, MPEG-2 e MPEG, o acúmulo de erros é o maior problema encontrado nas técnicas de inserção de marca d'água no *bitstream*. No H.264 esse problema piora já que uma marca d'água inserida em um quadro Intra pode propagar erros por todo o quadro Intra e conseqüentemente pelos quadros P e B, assim resultando em artefatos indesejáveis.

Dessa forma, esta seção apresenta um estudo para redução do acúmulo de erros provocado pelo método de inserção apresentado na Seção 4.2. Onde propõe-se a

modificação no *bitstream* dos sinais de *trailing ones* com objetivo de inserir uma informação oculta.

Os *trailing ones*(T1s) surgiram pelo fato de a maioria dos blocos 4x4 da DCT possuírem coeficientes com valores de 1. Então, um vídeo codificado com CAVLC contém uma grande quantidade de blocos com T1s. Tal quantidade possibilita a elaboração de um método para selecionar os blocos com T1 que modificados resultem numa menor degradação visual no vídeo. Assim, nas próximas subseções será realizado um estudo na predição Intra e Inter do padrão H.264 de modo a avaliar os blocos com T1 que quando alterados resultem num menor acúmulo de erros nos vídeos codificados usando o CAVLC no padrão H.264.

4.3.1 Explorando a Predição Intra

A nova técnica de predição espacial direcional do H.264 pode provocar um enorme acúmulo de erros quando amostras de um bloco Intra são modificadas, pois as amostras dos blocos Intra são formadas pela soma dos resíduos decodificados com uma predição gerada a partir de amostras adjacentes ao bloco. Uma vez que as amostras adjacentes são modificadas, as predições dos blocos que as utilizam são calculadas erroneamente, assim propagando erros por toda a fatia Intra.

Uma forma de evitar esse acúmulo de erros na fatia Intra é buscar por blocos onde suas amostras não sejam utilizadas pela predição Intra.

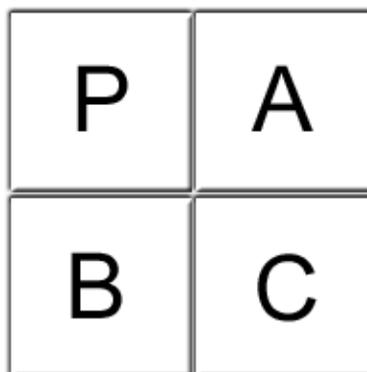


Figura 4.4: Blocos 4x4 que podem utilizar o bloco P para predição.

No caso da predição Intra4x4, o bloco P, ilustrado na Figura 4.4, pode ser utilizado para predição Intra por um ou mais blocos adjacentes, definidos como A, B e C. Os blocos A, B e C possibilitam diferentes combinações dos modos de

predição Intra, onde algumas combinações resultam na não utilização do bloco P para predição Intra. Por exemplo, um caso em que o bloco P não é utilizado para predição Intra acontece quando o bloco A tem o modo de predição vertical, o bloco B tem o modo de predição horizontal e o bloco C tem o modo de predição vertical.

A Tabela 4.2 mostra os modos de predição Intra4x4 dos blocos A, B e C que não utilizam o bloco P para predição. Os modos de predição Intra são descritos em detalhes na seção 3.4.1.

Tabela 4.2: *Modos de predição Intra4x4 que não utilizam o bloco P para predição*

Bloco 4x4	Modos de predição Intra4x4				
Bloco A	Modo 0	Modo 3	Modo 7	*	*
Bloco B	Modo 1	Modo 8	*	*	*
Bloco C	Modo 0	Modo 1	Modo 3	Modo 7	Modo 8

Para o caso da ocorrência de uma predição Intra16x16 abaixo ou a direita do bloco P, a Tabela 4.3 mostra os modos de predição Intra16x16 que não utilizam o bloco P para predição.

Baseado nos modos de predição das Tabelas 4.2 e 4.3, uma máquina de estados pode ser desenvolvida com objetivo de selecionar os blocos com predição Intra4x4 e T1 que não são utilizados na predição Intra de outros blocos.

Tabela 4.3: *Modos de predição Intra16x16 que não utilizam o bloco P para predição*

Bloco 16x16	Modos de predição Intra16x16
Bloco abaixo do bloco P	Modo 1
Bloco a direita do bloco P	Modo 0

4.3.2 Explorando a Predição Inter

O problema do acúmulo de erros gerado pela modificação do *bitstream* nos quadros Inter ocorre na maioria dos padrões de compressão. O H.264 diferentemente dos outros padrões de compressão utiliza múltiplos quadros de referência para predição Inter com a opção de quadros de longo prazo. Assim, um erro no primeiro quadro pode se propagar até o último quadro.

Dessa forma, para redução do acúmulo de erros, um mapa de predição é proposto para representar a quantidade de vezes que cada amostra é utilizada pela predição Inter, pois os blocos poucas vezes usados pela predição ao serem alterados provocam um menor acúmulo de erros do que os blocos muitas vezes utilizados. Para criação do mapa de predição, informações do decodificador, como o vetor de movimento e o tamanho do bloco são utilizados.

É importante notar que o mapa de predição tem o intuito de determinar o nível de acúmulo de erros provocado pela alteração de cada bloco. Assim, para cada bloco, o nível de acúmulo de erros é definido pela soma de vezes que suas amostras foram utilizadas pela predição Inter.

A Figura 4.5 mostra um exemplo de um mapa de predição em um quadro P, onde as regiões mais claras representam as amostras mais utilizadas pela predição Inter.

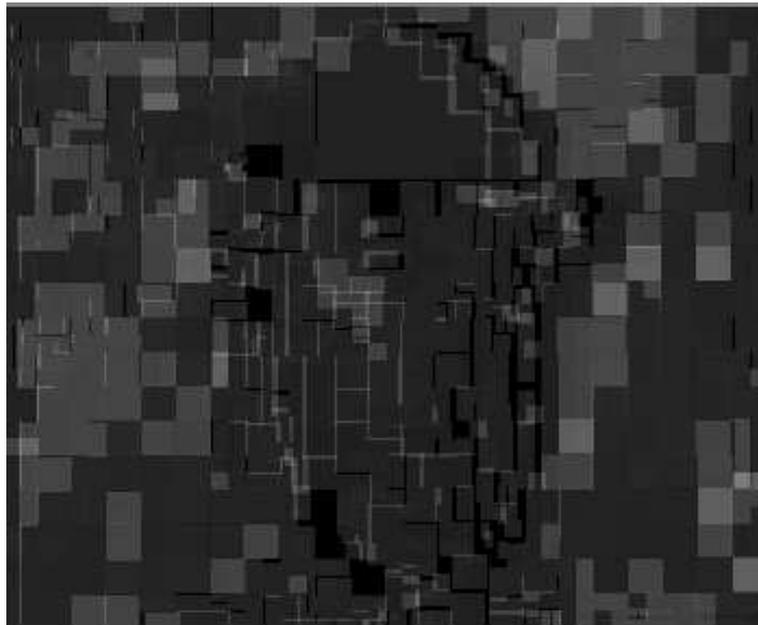


Figura 4.5: Mapa de predição de um quadro P.

Uma desvantagem na implementação de um mapa de predição é a sua propriedade não-causal que introduz no sistema uma janela de atraso definida pela quantidade de quadros utilizados para determinar o nível de acúmulo de erros.

4.4 Marca d'água frágil para autenticação de vídeo no padrão H.264 no perfil *Baseline*

A proposta desta seção é desenvolver um sistema de marca d'água de modo a autenticar cada GOP de vídeo no padrão H.264 no perfil *Baseline*. Para isso, utiliza-se o método de inserção de informação oculta apresentado na Seção 4.2.2 em conjunto com as estratégias para redução do acúmulo de erro descritas na Seção 4.3.

O sistema de marca d'água proposto, para assegurar que nenhuma adulteração transcorreu no vídeo, insere nos *bits* dos sinais de *trailing ones* uma marca d'água com informações de cada GOP de vídeo. Nele, somente as informações do canal de luma são consideradas na autenticação do vídeo, já que o sistema visual humano é mais sensível ao canal de luma do que ao canal de croma [15]. A Figura 4.6 mostra as etapas do algoritmo de marca d'água frágil proposto para autenticação de vídeo.

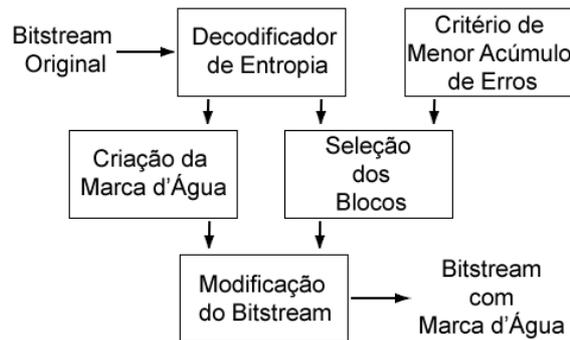


Figura 4.6: Etapas do algoritmo de marca d'água frágil proposto.

O decodificador de entropia fornece informações para a criação da marca d'água e para a seleção dos blocos com T1s. Para a criação da marca d'água utiliza-se os coeficientes quantizados do bloco residual, descrito com detalhes na Seção 4.4.2. Já a seleção dos blocos com T1s ocorre a partir das informações necessárias para a realização das estratégias de redução do acúmulo de erro, descritas na Seção 4.3, como: os modos de predição Intra, os vetores de movimento e o tamanho de cada bloco.

A modificação do *bitstream* é realizada pelo método de inserção de informação oculta apresentado na Seção 4.2.2. As próximas subseções descrevem o método de inserção da marca d'água e a criação de uma marca d'água.

4.4.1 Método de inserção da marca d'água

O método de inserção de informação oculta em [1], e apresentado na Seção 4.2.2, não define um critério na escolha dos blocos com T1 que serão usados para inserção da informação oculta. Dessa forma, propõe-se uma estratégia para selecionar os blocos com T1s que provocam menor degradação visual no vídeo, através de uma ordenação pelo critério de menor acúmulo de erros. O critério de menor acúmulo de erros define que os blocos poucas vezes usados pela predição Intra e Inter ao serem alterados provocam um menor acúmulo de erros do que os blocos muitas vezes utilizados. Logo, as estratégias apresentadas no estudo do acúmulo de erro na predição Inter e Intra são utilizados para selecionar e ordenar os blocos.

Os blocos selecionados para ordenação são formados de blocos 4x4 no canal luma contendo T1s, onde na fatia Intra são considerados somente os blocos não utilizados para predição Intra e na fatia Inter são considerados apenas os blocos Inter.

Após a seleção dos blocos, a ordenação dos blocos selecionados ocorre no final de cada GOP pelo critério de menor acúmulo de erros, onde os blocos poucas vezes usados pela predição são ordenados de modo a serem os primeiros utilizados na inserção da marca d'água. Outro ponto interessante é que, permitido um aumento da complexidade computacional na ordenação dos blocos, alguns critérios perceptuais priorizando os coeficientes de alta frequência podem ser adotados. Por exemplo:

1. Os blocos Inter e Intra com T1 não utilizados para predição podem ser ordenados priorizando os blocos com *tralling ones* de maior frequência.
2. No caso de igualdade no nível do acúmulo de erros, o bloco com T1 que possuir a maior frequência é priorizado.

Vale ressaltar que os blocos Intra utilizados pela predição Intra e os blocos Intra na fatia Inter não são usados na seleção dos blocos.

Finalmente, depois da seleção e ordenação dos blocos, o método de inserção apresentado na Seção 4.2.2 insere uma marca d'água nos bits dos sinais de *tralling ones* selecionados. As etapas do método de inserção em conjunto com uma sugestão para criação da marca d'água são ilustradas na Figura 4.7.

A detecção da marca d'água realiza as mesmas etapas da inserção, com a diferença de que ao invés de modificar os *bits* dos *tralling ones* selecionados, estes bits

são extraídos e comparados com a marca d'água gerada.

É importante notar que a marca d'água é considerada frágil, pois qualquer modificação em uma das informações fornecidas pelo decodificador de entropia podem destruir a marca d'água. Dentre estas informações estão os modos de predição Intra, os vetores de movimento e o tamanho de cada bloco. Além disso, ainda há as informações dos coeficientes quantizados do bloco residual para criação da marca d'água.

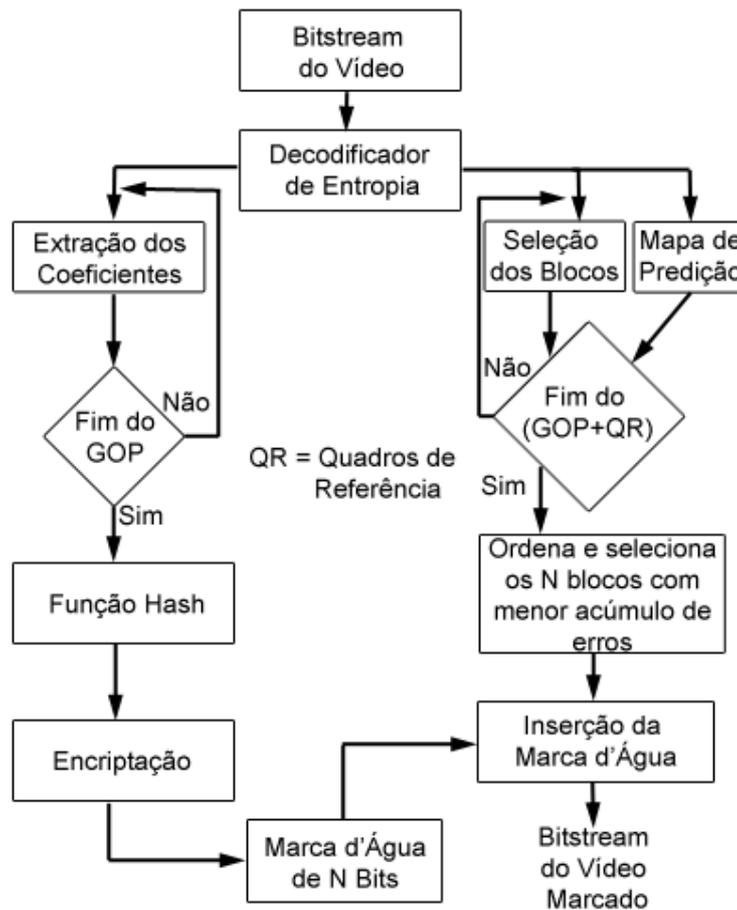


Figura 4.7: Etapas do algoritmo de autenticação de vídeo.

4.4.2 Criação da marca d'água

O conteúdo de uma marca d'água inserida no vídeo é essencial para segurança de um sistema de autenticação, pois existem conhecidos ataques de inserção que procuram autenticar os vídeos adulterados. Por exemplo, o ataque de cópia tem por objetivo estimar uma marca d'água de um vídeo marcado para copiar essa marca para outros

vídeos. Nos casos em que a marca d'água é fixa, ou seja, não se adapta às características do vídeo, a simples cópia da marca d'água permite que vídeos adulterados sejam autenticados.

Outro importante ataque que compromete os sistemas de marca d'água públicos é o chamado ataque de busca. Nele, o malfeitor, de posse do detector, efetua buscas pela chave de segurança do sistema, pois uma vez encontrada a chave de segurança outros vídeo podem ser indevidamente autenticados. Portanto, sistemas de marca d'água devem ser projetados para que o tamanho do espaço de busca torne o ataque impraticável.

Dessa forma, para segurança do sistema de marca d'água proposto, uma marca d'água é sugerida com base na assinatura digital proposta em [27]. A única diferença em relação ao artigo em [27] é a não utilização dos coeficientes de valores iguais a +1 e -1 na assinatura digital, visto que os *tralling ones* são modificados durante a inserção da marca d'água.

Então, a marca d'água é formada pelos valores dos coeficientes mais significativos dos blocos residuais que são o coeficiente DC e os dois primeiros coeficientes AC do bloco com valores diferentes de +1 ou -1. No final de cada GOP, esses valores passam por uma função *hash* e em seguida por um algoritmo de criptografia assimétrica ou de chave pública. A intenção da função *hash* é reduzir a informação, formada pelos coeficientes mais significativos dos blocos, a uma determinada quantidade de bits. Esse procedimento embutido no processo de inserção da marca d'água é ilustrado na Figura 4.7.

4.5 Resultados e Discussões

Para verificar as propriedades de capacidade e transparência do sistema de marca d'água proposto e compará-lo ao método de inserção em [1], alguns testes foram realizados nos quadros Intra, Inter e nos GOPs do vídeo.

Os algoritmos propostos para o sistema de marca d'água foram implementados no decodificador do *software* de referência do H.264 (JM 14.2) e as simulações utilizam o vídeo *Foreman* de 300 quadros na resolução CIF (352x288) codificado no perfil Baseline do padrão H.264 com GOP de 10 quadros, 4 quadros de referência, QP=27

e a fatia é igual a um quadro.

4.5.1 Resultados no Quadro Intra do GOP de Vídeo

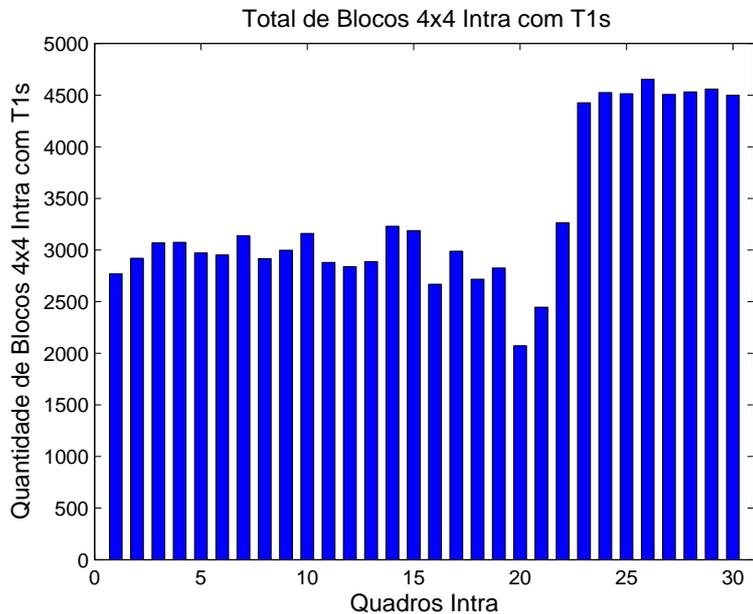


Figura 4.8: Total de Blocos 4x4 Intra com T1s.

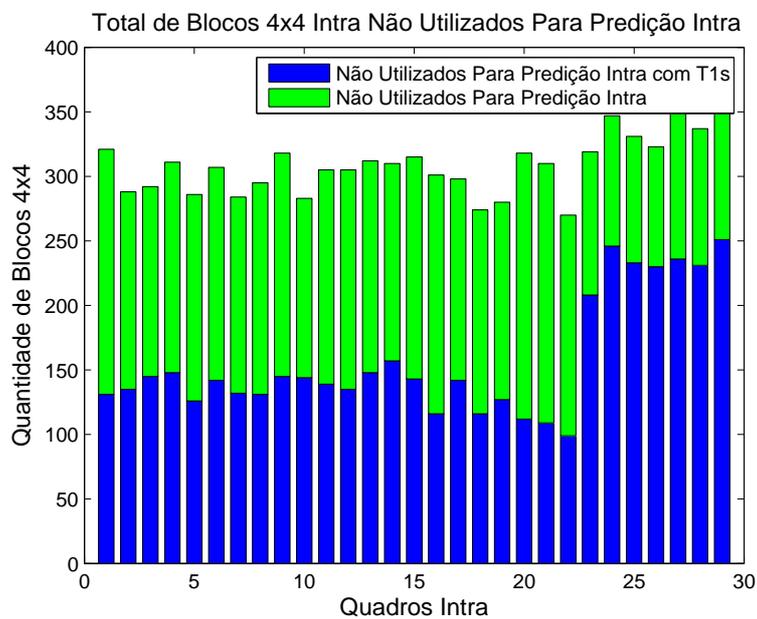


Figura 4.9: Total de blocos 4x4 Intra não utilizados para predição Intra.

Para as simulações nos quadros Intra, o gráfico em barras da Figura 4.8 ilustra uma grande quantidade de blocos com T1 nos quadros Intra do vídeo codificado.

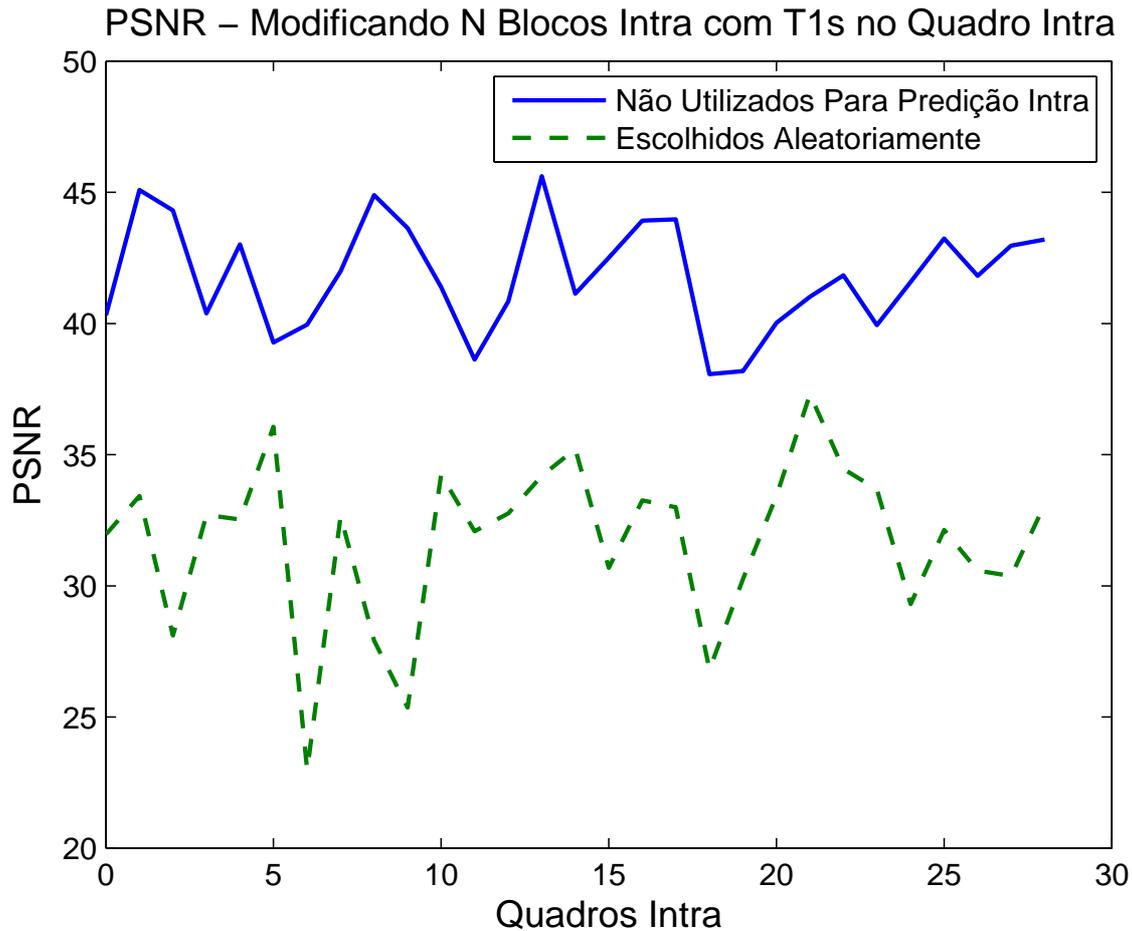


Figura 4.10: PSNR - Modificando N blocos Intra com T1s no quadro Intra.

A Figura 4.9 mostra uma quantidade razoável de blocos com T1 não utilizados para predição Intra. Esses resultados da Figura 4.9 foram obtidos a partir da estratégia apresentada na Seção 4.3.1, que seleciona os blocos que não são utilizados para predição Intra.

Para comparar os resultados obtidos nos quadros Intra pelo método de inserção proposto neste capítulo com os do método de inserção proposto em [1] é realizada uma comparação entre a modificação de N blocos com T1 escolhidos aleatoriamente e N blocos com T1 que não são utilizados para predição Intra. Esta comparação é ilustrada na Figura 4.10.

Observa-se na Figura 4.10 que o método proposto apresenta um PSNR maior e conseqüentemente, uma redução do acúmulo de erros no quadro Intra em relação ao método [1].

4.5.2 Resultados no Quadro Inter do GOP de Vídeo

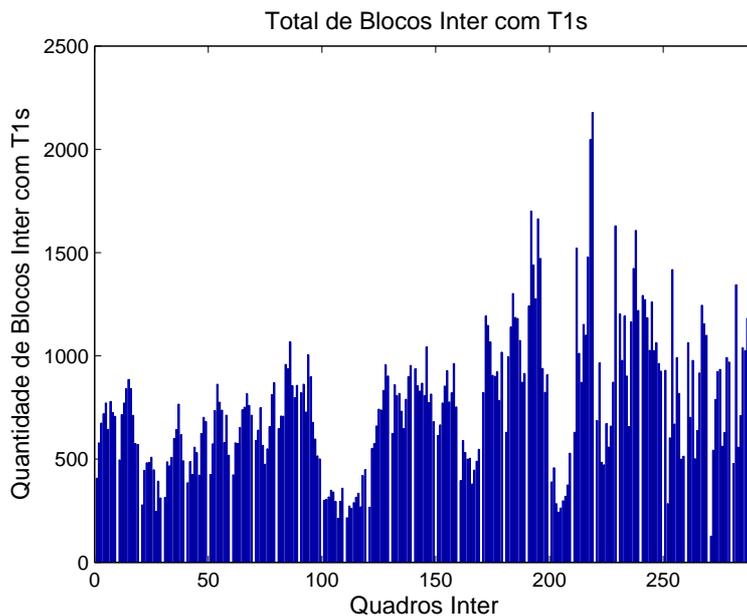


Figura 4.11: Total de blocos Inter com T1s.

Da mesma forma que nos quadros Intra, os quadros Inter apresentam uma grande quantidade de blocos com T1, vista na Figura 4.11.

Para comparar os resultados obtidos nos quadros Inter pelo método de inserção proposto neste capítulo e pelo método de inserção proposto em [1], a Figura 4.12 mostra a comparação entre a modificação de 1024 blocos Inter com T1 em cada GOP escolhidos aleatoriamente ou escolhidos pelo critério de menor acúmulo de erro.

É importante notar na Figura 4.12 que em intervalos de tempos constantes há a formação de picos e vales no PSNR, onde, no início do GOP temos os picos e no fim do GOP os vales. Assim, evidencia-se a ocorrência do acúmulo de erros em ambos os métodos de inserção. No entanto, o método proposto apresenta uma redução no acúmulo de erros em relação ao método em [1].

4.5.3 Resultados nos Quadros Intra e Inter do GOP de Vídeo

Como os quadros Intra e Inter apresentaram uma grande quantidade de blocos com T1, o GOP do vídeo também apresenta, podendo ser visto na Figura 4.13. Além disso, a Figura 4.13 mostra os blocos com T1s utilizados e não utilizados para

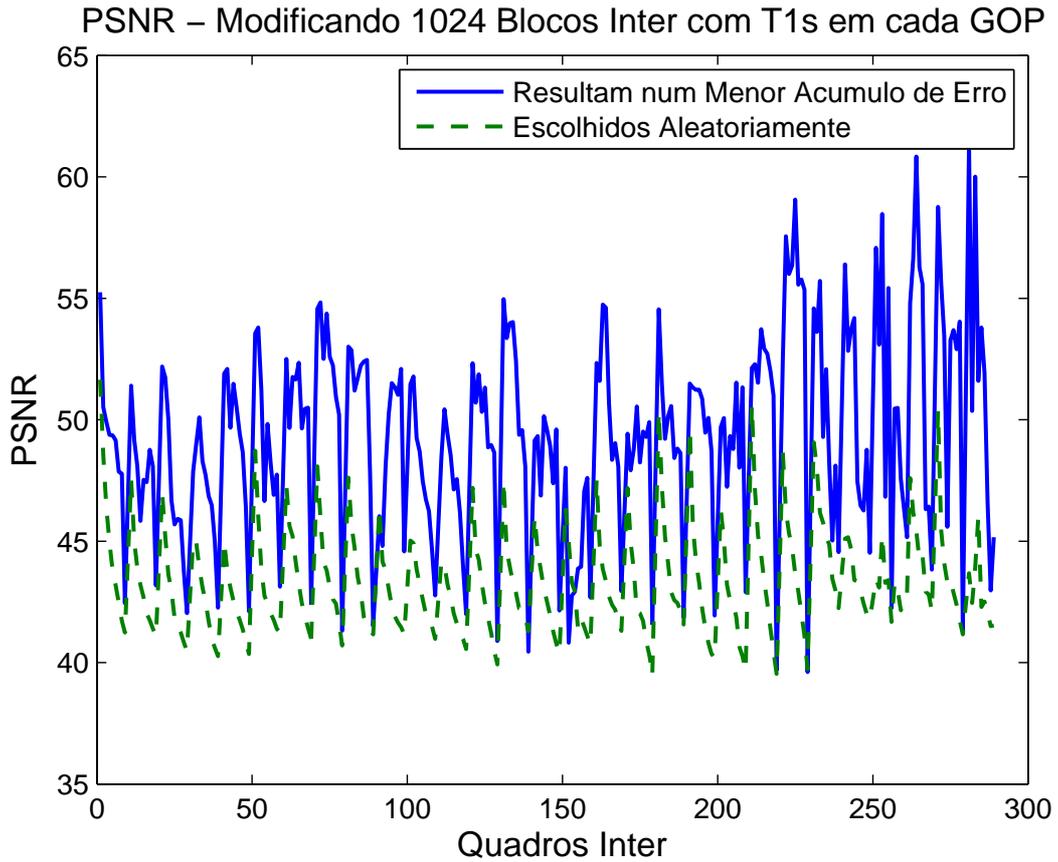


Figura 4.12: PSNR - Modificando em cada GOP 1024 blocos Inter com T1s.

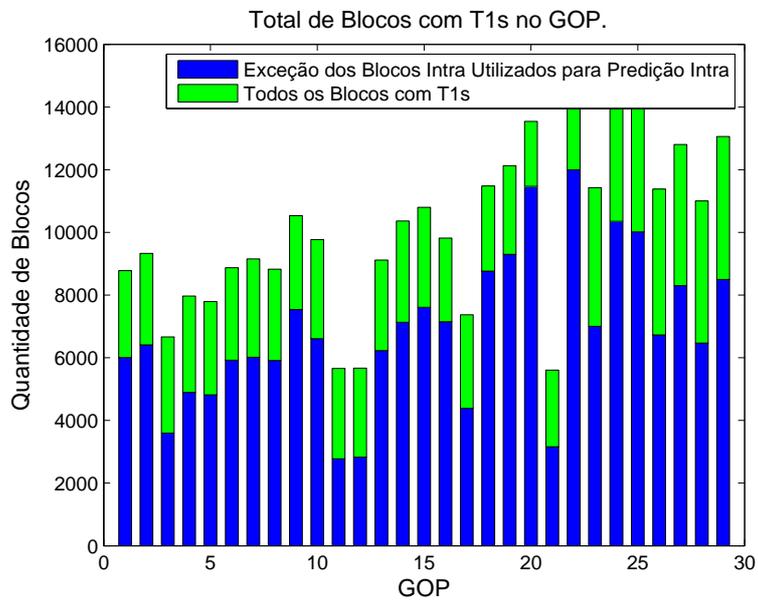


Figura 4.13: Total de blocos com T1s no GOP.

ordenação e seleção dos N blocos com menor acúmulo de erros.

Por fim, para comparar os dois métodos comentados neste capítulo, a Figura 4.14

apresenta o resultado do PSNR para a modificação de 1024 blocos com T1 em cada GOP do vídeo utilizando o critério de menor acumulo de erro e 1024 blocos com T1 em cada GOP do vídeo escolhidos aleatoriamente.

Observa-se que o método de inserção em [1] pode provocar uma enorme distorção no vídeo, visto pelo PSNR da Figura 4.14 e pelo resultado subjetivo de um quadro P marcado que é ilustrado na Figura 4.15. Enquanto isso, o método apresentado neste capítulo, apesar de aumentar a complexidade computacional do método em [1], mostra um resultado aceitável tanto pelo PSNR quanto pelo resultado subjetivo de um quadro P marcado que é visto na Figura 4.15.

Vale ressaltar na Figura 4.14 que os valores do PSNR iguais a 70 são de quadros não modificados pelo processo de inserção da marca d'água, visto que o valor do PSNR muito alto dificultaria a visualização do resultado.

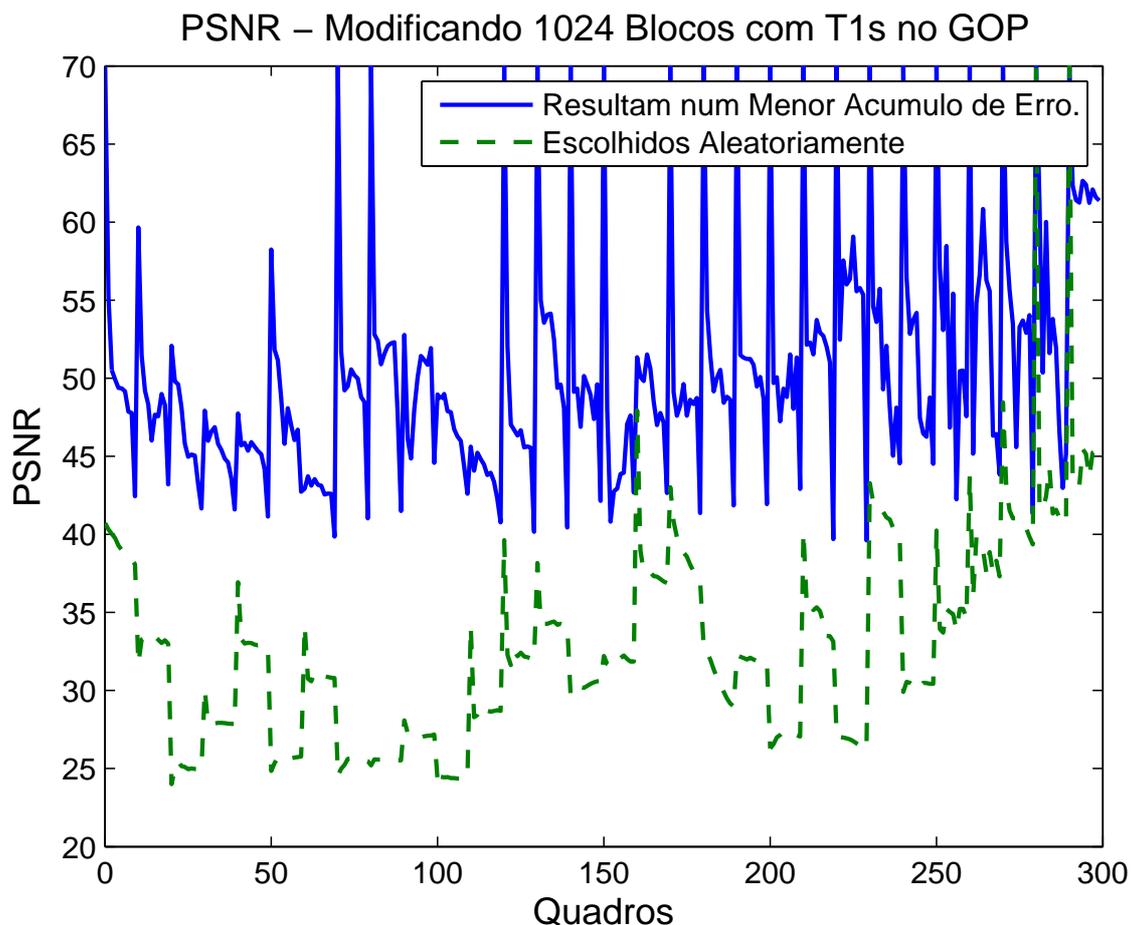


Figura 4.14: PSNR - Modificando 1024 blocos com T1s no GOP.



(a) Quadro P do método de inserção proposto.



(b) Quadro P do método de inserção [1].

Figura 4.15: *Quadro P* após a inserção da marca d'água.

Capítulo 5

Marca d'Água Semi-Frágil Adaptativa para autenticação dos padrões de compressão com perda baseados na DCT (JPEG, MPEG-2 e H.264)

À medida que surge um novo padrão de compressão mais eficiente, diminui-se o espaço para inserção de uma marca d'água imperceptível. Enquanto os métodos de compressão com perda exploram a irrelevância do vídeo para reduzir a quantidade de bits para armazenamento, os sistemas de marca d'água utilizam as áreas irrelevantes do vídeo para inserir uma informação oculta. Esse conflito entre compressão com perda e marcação d'água tem motivado o desenvolvimento de diferentes técnicas de inserção com a finalidade de incorporar conhecimento perceptual nos sistemas de marca d'água. O objetivo dessas técnicas é obter um melhor balanceamento entre as conflitantes propriedades: transparência, capacidade e robustez.

A proposta deste capítulo é elaborar um sistema de marca d'água adaptativo que garanta a presença da marca d'água até um determinado nível de compressão com perda. Para isso, propõe-se um método de inserção de marca d'água baseado no modelo perceptual de Watson [3] integrado ao QIM (*Quantization Index Modulation*), de modo a obter uma melhor relação de custo-benefício entre as propriedades:

transparência, capacidade e robustez.

Na literatura, Wolfgang em [7] e Maneli em [12] propõem sistemas de marca d'água adaptativos utilizando o método de inserção de marca d'água por espalhamento espectral em conjunto com modelo perceptual de Watson [3]. O método de Maneli se diferencia do Wolfgang pela escolha dos coeficientes da DCT a serem modificados e pela derivação do modelo perceptual de Watson de blocos 8x8 para blocos 4x4 que é visto na seção 5.1.1.

Este capítulo apresenta na Seção 5.1 o modelo perceptual de Watson em blocos 8x8 [3] e em blocos 4x4 [12]. A Seção 5.2 apresenta o método de inserção QIM. Na Seção 5.3 é discutido o sistema de marca d'água proposto. Por último, os codificadores JPEG e H.264 são utilizados para gerarem resultados em relação à robustez, capacidade e transparência do sistema de marca d'água proposto.

5.1 Modelo perceptual de Watson

Para determinar a quantidade máxima de ruído de quantização que cada coeficiente da DCT tolera sem afetar a qualidade visual da imagem, Watson propõe em [3], um modelo perceptual baseado em blocos 8x8 da DCT. Esse modelo foi criado para que na codificação JPEG os passos da quantização se adaptassem às características da imagem.

Com esse propósito o modelo perceptual de Watson estima, em JND¹, a perceptibilidade da modificação de cada coeficiente de um bloco 8x8 da DCT. Para mensurar essa perceptibilidade, Watson considera duas características do sistema visual humano: sensibilidade e mascaramento.

Sensibilidade

Sensibilidade refere-se à resposta do olho humano a um estímulo direto. Nos experimentos com o objetivo de mensurar a sensibilidade, os observadores recebem estímulos isolados de modo que as percepções desses estímulos sejam testadas.

Watson define uma tabela com a sensibilidade das frequências, onde cada entrada

¹O nível de distorção que é percebido em 50% dos experimentos e não é percebido nos outros 50% é referenciado como 1 JND (Just Noticeable Difference).

da tabela, $t_{i,j}$, definida em [3], representa a quantidade de modificação de um coeficiente necessária para produzir 1 JND, ou seja, a menor magnitude de modificação de um coeficiente para torná-lo perceptível na ausência de qualquer mascaramento.

Mascaramento

Mascaramento refere-se ao fato de que o contexto afeta a percepção, ou seja, a presença de um sinal pode esconder ou mascarar a presença de outro sinal. Por exemplo, uma textura que é fácil de enxergar isoladamente, diminui sua visibilidade quando adicionada numa imagem com textura. Assim, o mascaramento é uma medida de resposta para um estímulo quando um segundo estímulo está presente. Watson utiliza dois tipos de mascaramento: Mascaramento de luminância e de contraste.

- Mascaramento de Luminância: A adaptação à luminância refere-se ao fato de que a modificação de um coeficiente por uma determinada quantidade é menos perceptível, quanto maior for a média da luminância do bloco 8x8. Assim, Watson ajusta a tabela de sensibilidade de acordo com o coeficiente DC do bloco.

$$t_{i,j,k} = t_{i,j} \left(\frac{X_{0,0,k}}{\bar{X}_{0,0}} \right) a^T \quad (5.1)$$

Os coeficientes de cada bloco k são representados por $X_{i,j,k}$. Onde a^T é uma constante de valor 0.649 sugerida em [3], $X_{0,0,k}$ é o coeficiente DC do k-ésimo bloco e $\bar{X}_{0,0}$ é a média dos coeficientes DC da imagem.

- Mascaramento de Contraste: Refere-se à redução da visibilidade de uma modificação numa frequência por causa da energia presente nessa frequência.

$$s_{i,j,k} = \max \left\{ t_{i,j,k}, |X_{i,j,k}|^{w_{i,j}} t_{i,j,k}^{1-w_{i,j}} \right\} \quad (5.2)$$

Watson em [3], define o valor de $w_{i,j} = 0.7$ para todos i e j. O valor de $s_{i,j,k}$ estima a quantidade de modificação em cada coeficiente da DCT para produzir 1 JND.

Pooling

Os modelos de sensibilidade e mascaramento podem ser usados para estimar a perceptibilidade da modificação de uma única frequência. Entretanto, se múltiplas

freqüências são modificadas é preciso saber como combinar as informações de sensibilidade e mascaramento para cada freqüência.

Watson emprega duas formas de *pooling*. Um pela combinação das modificações dos blocos e outro pela combinação das modificações nas diferentes freqüências de cada bloco. Ambos os *pooling* são calculados pela métrica de Minkowski, onde Watson sugere o valor do expoente p igual a 4.

$$D(x_o, x_w) = \left(\sum |d_{i,j,k}|^p \right)^{\frac{1}{p}} \quad (5.3)$$

Para obter a medida de distorção, $d_{i,j,k}$, entre a imagem original, x_o , e a imagem modificada, x_w , Watson calcula a diferença entre os coeficientes da DCT das duas imagens, $e_{i,j,k} = x_{w_{i,j,k}} - x_{o_{i,j,k}}$; e faz

$$d_{i,j,k} = \frac{e_{i,j,k}}{s_{i,j,k}} \quad (5.4)$$

onde $s_{i,j,k}$ é definido na Equação 5.2.

5.1.1 Modelo perceptual de Watson para blocos 4x4

Maneli em [12], propõe o modelo perceptual de Watson para blocos 4x4. Expandindo o bloco em termos das funções base da DCT4x4, onde os coeficientes são definidos por $X_{i,j}$ e as amostras do bloco 4x4 por $x(n_1, n_2)$ temos

$$x(n_1, n_2) = X_{i,j} c_i c_j \cos\left(\frac{\pi(2n_1+1)i_4}{2 \times 4}\right) \cos\left(\frac{\pi(2n_2+1)j_4}{2 \times 4}\right), \quad (5.5)$$

onde,

$$0 \leq i_4, j_4 \leq 3 \quad e \quad c_i = \begin{cases} \sqrt{1/4} & i = 0 \\ \sqrt{2/4} & i > 0 \end{cases}$$

Para expansão do bloco em termos das funções base da DCT8x8 temos

$$x(n_1, n_2) = X_{i,j} c_i c_j \cos\left(\frac{\pi(2n_1+1)i_8}{2 \times 8}\right) \cos\left(\frac{\pi(2n_2+1)j_8}{2 \times 8}\right), \quad (5.6)$$

onde,

$$0 \leq i_8, j_8 \leq 3 \quad e \quad c_i = \begin{cases} \sqrt{1/8} & i = 0 \\ \sqrt{2/8} & i > 0 \end{cases}$$

Comparando as Equações 5.5 e 5.6 observa-se que as funções base da DCT4x4 têm as mesmas frequências das funções bases da DCT8x8 com índices pares.

$$i_8 = 2 \times i_4 \quad (5.7)$$

$$j_8 = 2 \times j_4$$

Assim, Maneli sugere que os limiares de visibilidade, $s_{i,j,k}$, das funções base da DCT4x4 sejam calculados como os $s_{i,j,k}$ das funções bases com índices pares da DCT8x8 são calculados.

5.2 QIM

Em [2], Chen e Wornell propõem métodos de inserção de marca d'água baseados na quantização chamados de QIM (Quantization Index Modulation). A idéia básica da Modulação por Índice Quantizado é a quantização da amostra do sinal, x , de acordo com a informação a ser inserida, m , através da escolha de um quantizador, $Q(\cdot)$, a partir de um conjunto de quantizadores. Um dos métodos de inserção apresentados por Chen e Wornell é uma realização de baixa complexidade chamada de DM (*Dither Modulation*), onde as amostras do sinal marcado, y são dadas por:

$$y = Q(x, \Delta, m, \delta), \quad m \in 0, 1 \quad (5.8)$$

Δ é um passo de quantização fixo e δ é uma forma de ruído pseudo-aleatória para tornar aleatório o erro de quantização. O quantizador $Q(\cdot)$ é definido como:

$$Q(x, \Delta, m, \delta) = \Delta \cdot \text{Round} \left(\frac{x - \delta - m \frac{\Delta}{2}}{\Delta} \right) + \delta + m \frac{\Delta}{2}, \quad (5.9)$$

onde o operador *Round* tem a função de arredondar para o inteiro mais próximo.

No detector, a amostra do sinal recebido, z , é novamente quantizada usando um conjunto de quantizadores. O quantizador que apresentar a menor distância em relação à amostra do sinal recebido, z , identifica a informação embutida, \hat{m} .

$$\hat{m} = \arg \min_{b \in \{0,1\}} |z - Q(z, \Delta, b, \delta)| \quad (5.10)$$

Importante notar que δ tem por objetivo melhorar a qualidade perceptual do sinal marcado e também fornecer uma maior segurança ao método de inserção, já que o ruído pseudo-aleatório pode ser considerado uma chave de segurança, onde sem ela não é possível a detecção.

5.3 Sistema de Marca d'Água Proposto

Baseado no modelo perceptual de Watson, esta seção elabora um método de inserção de marca d'água, QIM-Watson, que seleciona os coeficientes da DCT com maior tolerância perceptual à distorção, para inserir uma marca d'água, m , através da quantização dos coeficientes da DCT selecionados, X .

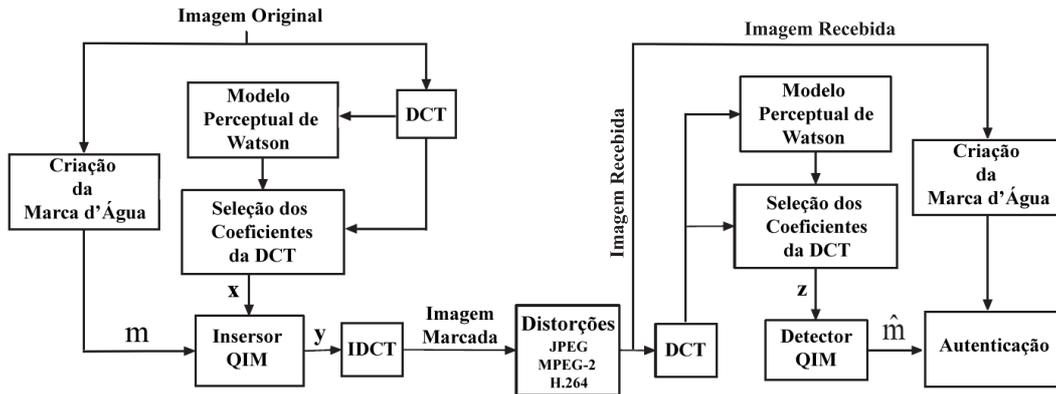


Figura 5.1: Sistema de Marca d'Água Proposto.

A proposta do sistema de marca d'água, ilustrado na Figura 5.1, é garantir a presença da marca d'água até um determinado nível de compressão com perdas, com intuito de desenvolver aplicações na autenticação ou no monitoramento de transmissão. Para estas aplicações, o estágio de detecção da marca d'água deve ser necessariamente cega.

Dessa forma, a detecção da marca d'água proposta não conhece as posições dos coeficientes selecionados durante a inserção. Assim, estas posições dos coeficientes selecionados são obtidas a partir da imagem marcada recebida através do modelo perceptual de Watson, como ilustrado na Figura 5.1.

Para associar cada bit da marca d'água a um coeficiente da imagem, na criação da marca d'água propõe-se que a quantidade de bits da marca d'água corresponda a quantidade de amostras da imagem. Logo, nem todos os bits da marca d'água serão inseridos na imagem, somente os bits correspondentes aos coeficientes selecionados.

Vale ressaltar que a inserção da marca d'água é no domínio da frequência. Assim, o tamanho do bloco usado pelo modelo perceptual de Watson e pela DCT é definido pelo padrão de compressão com perda ao qual as imagens ou os vídeos serão submetidos. Por exemplo, para autenticar os vídeos codificados pelo H.264 utiliza-

se o modelo perceptual de Watson e a DCT para blocos 4x4, e para autenticar as imagens codificadas pelo JPEG utiliza-se o modelo perceptual de Watson e a DCT para blocos 8x8.

5.3.1 Inersor QIM

Para este trabalho, considera-se o método mais simples do QIM, onde os coeficientes selecionados, x , são quantizados usando um quantizador $Q(\cdot)$ que é escolhido de um conjunto de quantizadores de acordo com a informação, m , a ser inserida.

Os coeficientes selecionados marcados, y são dados por:

$$y = Q(x, \Delta, m), \quad m \in 0, 1 \quad (5.11)$$

$$\Delta = 2 \times \Delta_{i,j} \quad (5.12)$$

$\Delta_{i,j}$ é o passo de quantização que representa a robustez do sistema e o quantizador $Q(\cdot)$ é definido como:

$$Q(x, \Delta, m) = \Delta \cdot \text{Round} \left(\frac{x - m \frac{\Delta}{2}}{\Delta} \right) + m \frac{\Delta}{2} \quad (5.13)$$

onde o operador *Round* tem a função de arredondar para o inteiro mais próximo.

O erro máximo de quantização é igual a $\Delta_{i,j}$ e $\frac{\Delta_{i,j}}{2}$ é o maior valor que um coeficiente selecionado pode ser alterado para que a informação inserida seja mantida.

5.3.2 Seleção dos coeficientes

Para avaliar o modelo perceptual de Watson e elaborar um método de inserção que melhor relacione as propriedades de transparência, capacidade e robustez desenvolveu-se um programa, ilustrado na Figura 5.2, contendo informações dos coeficientes da DCT, mostrados na parte inferior da Figura 5.2, e informações dos limiares de visibilidade do modelo perceptual de Watson, mostrados na parte superior da Figura 5.2. Além de inserir uma marca d'água na imagem, o programa permite avaliar a degradação visual causada pela modificação de cada coeficiente.

O método de inserção proposto diferentemente da maioria das técnicas apresentadas na literatura, [14] [28] [29] [30], não altera todos os blocos da imagem, visto

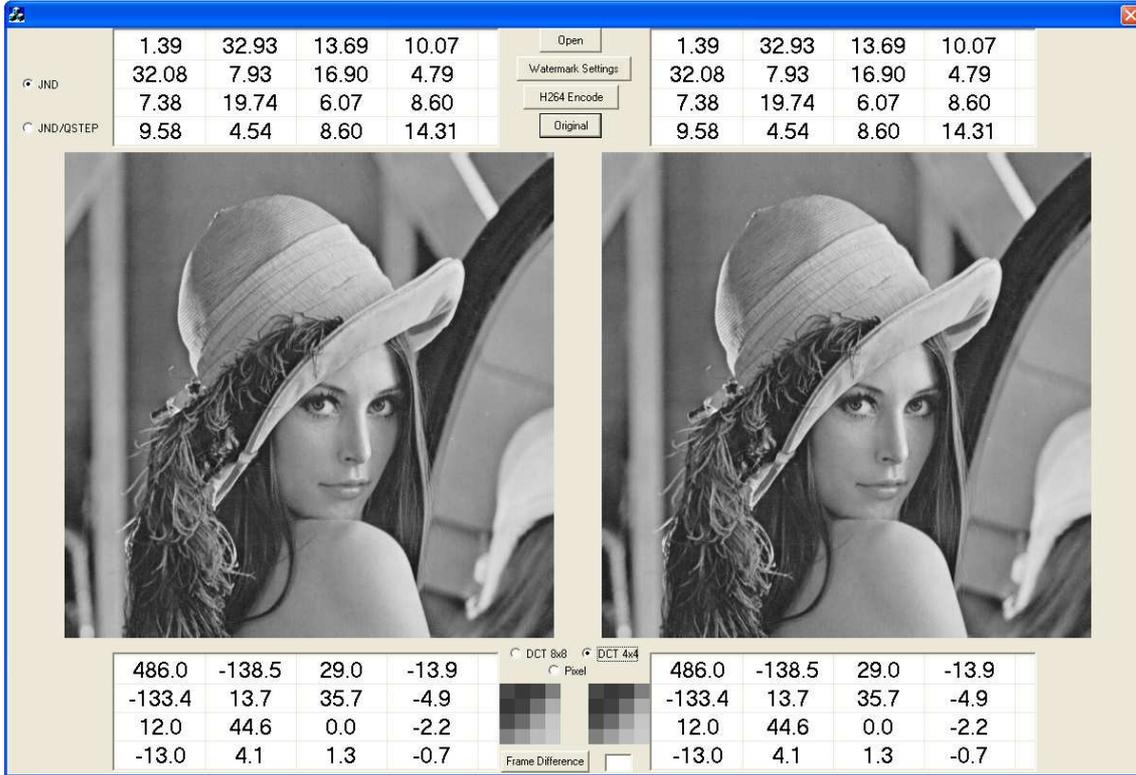


Figura 5.2: Programa desenvolvido para elaboração do método de inserção proposto.

que os blocos com área uniforme ao serem alterados para resistirem à um padrão de compressão causam artefatos visuais na imagem.

Dessa forma, o sistema de marca d'água proposto altera os coeficientes de acordo com as características de cada imagem, assim tornando a sua capacidade variável. Para controlar e melhor balancear a capacidade e a transparência do método de inserção, define-se o parâmetro T_{cap} , que representa a fronteira entre os coeficientes selecionados e não selecionados, ou seja, os coeficiente que possuem uma capacidade visual, CV , maior que T_{cap} são selecionados para a inserção da marca d'água.

$$CV = \frac{s_{i,j,k}}{\Delta_{i,j}} \quad (5.14)$$

Onde $\Delta_{i,j}$ é o passo de quantização que representa a robustez do sistema e $s_{i,j,k}$ é o limiar de visibilidade definido pelo modelo perceptual de Watson.

É importante notar que não há ocorrência de coeficientes com distorções maiores que 1 JND, quando o valor de CV é maior ou igual a 1.0, já que o erro máximo de quantização do método de inserção é igual a $\Delta_{i,j}$.

Na literatura, Wolfgang [7] e Maneli [12] propõem diferentes algoritmos para a seleção dos coeficientes da DCT. Wolfgang propõe alterar somente os coeficientes,

$X_{i,j,k}$, com valores maiores que seus respectivos limiares de visibilidade, $s_{i,j,k}$, enquanto Maneli define dois limiares, T_{coef} e T_{block} , para selecionar os coeficientes a serem modificados. O algoritmo de inserção de Maneli somente altera os coeficientes com valores maiores que T_{coef} e para diminuir o erro de *pooling* limita a quantidade de coeficientes alterados por bloco pelo limiar de T_{block} .

Das propostas de Wolfgang e Maneli optou-se pela de Wolfgang, onde apenas seleciona-se os coeficientes com os valores maiores que os seus respectivos limiares de visibilidade. Isto porque, em aplicações com detecção cega, há uma dificuldade em limitar a quantidade de coeficientes por bloco, visto que coeficientes não selecionados na inserção podem ser escolhidos durante a detecção, assim podendo reduzir o desempenho do sistema de marca d'água.

Então, para obter um melhor desempenho na detecção cega define-se o parâmetro T_{dis} , com a finalidade de evitar que os coeficientes da DCT não selecionados na inserção sejam escolhidos durante a detecção. Assim, somente os coeficientes que possuem uma capacidade visual, CV, maior que $(T_{cap} - T_{dis})$ são escolhidos na inserção da marca d'água, enquanto na detecção da marca d'água, somente os coeficientes que possuem uma capacidade visual, CV, maior que T_{cap} são selecionados.

Em relação ao erro de *pooling* nos blocos, o método de inserção proposto possui um interessante balanceamento entre as propriedades de transparência, capacidade e robustez, pois somente ocorre a seleção de muitos coeficientes por bloco quando a robustez do sistema, o passo de quantização, provoca pequenas distorções em cada coeficiente e quando ocorre um aumento no passo de quantização há uma diminuição na quantidade de coeficientes selecionados por bloco.

O algoritmo de inserção proposto segue os seguintes passos:

1. Dependendo da aplicação, calcula-se a DCT 8x8 ou a DCT aproximada 4x4.
2. Calcula-se o limiar de mascaramento, $s_{i,j,k}$, para cada coeficiente, $X_{i,j}$, de cada bloco k.
3. Depois, define-se o valor da capacidade visual (CV) de cada coeficiente, dividindo o valor de $s_{i,j,k}$ pelo passo de quantização $\Delta_{i,j}$.

$$CV = \frac{s_{i,j,k}}{\Delta_{i,j}} \quad (5.15)$$

4. Em cada bloco, se o coeficiente tiver um $CV \geq (T_{cap} - T_{dis})$ e $X_{i,j,k} > s_{i,j,k}$, insere-se um bit da marca d'água, b_i , pelo quantizador $Q(\cdot)$ definido como:

$$Q(x, \Delta, m) = \Delta \cdot \text{Round} \left(\frac{x - m \frac{\Delta}{2}}{\Delta} \right) + m \frac{\Delta}{2} \quad (5.16)$$

5.3.3 Detector QIM

No detector, os coeficientes que tiverem um $CV \geq T_{cap}$ e $X_{i,j,k} > s_{i,j,k}$ são selecionados. Estes coeficientes selecionados, z , são novamente quantizados por um conjunto de quantizadores. O quantizador que apresentar a menor distância em relação aos coeficientes selecionados, z , identifica a informação embutida, \hat{m} .

$$\hat{m} = \arg \min_{b \in \{0,1\}} |z - Q(z, \Delta, b)| \quad (5.17)$$

Nota-se que as Equações 5.16 e 5.17 assumem que cada coeficiente selecionado contém um bit. Para melhorar a robustez, sugere-se a inserção do mesmo bit, m , nos coeficientes selecionados dos blocos $N \times N$ da imagem, onde N é definido pelo parâmetro T_{bloco} e pelo tamanho do bloco $M \times M$ da DCT.

$$N = T_{bloco} \times M \quad (5.18)$$

Então, para cada bloco $N \times N$ da imagem, a Equação de detecção é definida como:

$$\hat{m} = \arg \min_{b \in \{0,1\}} \sum |z - Q(z, \Delta, b)| \quad (5.19)$$

Vale ressaltar que o valor de T_{bloco} igual a zero significa que cada coeficiente possui um bit.

5.3.4 Criação da Marca d'Água

Para avaliar o método de inserção proposto neste capítulo utiliza-se uma marca d'água fixa do tamanho da imagem a ser marcada, onde cada bit está associado a uma posição na imagem. Em trabalhos futuros para obter uma maior segurança e um melhor desempenho na autenticação é necessário que a marca d'água se adapte as características da imagem. Além disso, essas características extraídas da imagem devem ser robustas aos padrões de compressão. Alguns artigos já abordam esse tema [31] [32].

5.3.5 Autenticação

Depois da detecção, a marca d'água gerada, \hat{m} , é comparada com a marca d'água original, m , para informar se a imagem é autêntica ou adulterada. A autenticidade da imagem é verificada pela comparação da porcentagem de bits detectados corretamente com um limiar previamente definido T_{aut} .

5.4 Resultados e Discussões

Para autenticação dos padrões de compressão é necessário que a marca d'água inserida no domínio sem compressão resista pelo menos as distorções provocadas pelos codificadores. Assim, nessa seção são feitas simulações com os codificadores JPEG e H.264, para avaliar o desempenho do método de inserção, QIM-Watson, em relação à robustez, à transparência e à capacidade. Essas propriedades conflitantes são balanceadas pelo QIM-Watson através dos parâmetros, $\Delta_{i,j}$, T_{cap} , T_{dis} e T_{bloco} , mostrados na Tabela 5.1.

Tabela 5.1: *Parâmetros do método de inserção de marca d'água, QIM-Watson.*

Parâmetro	Função do parâmetro
$\Delta_{i,j}$	Define a robustez do método.
T_{cap}	Controla a transparência e a capacidade do método. Quanto menor o valor de T_{cap} , maior será a capacidade do método à custa de uma maior distorção.
T_{dis}	É um percentual da capacidade visual, CV, com o objetivo de evitar a seleção de coeficientes na detecção que não foram selecionados durante a inserção da marca d'água.
T_{bloco}	Define o tamanho do bloco no qual 1 bit é inserido. Onde um maior tamanho do bloco aumenta a robustez e diminui a capacidade do método.

Nas simulações define-se diferentes níveis de robustez, $\Delta_{i,j}$, e uma qualidade perceptual fixa, $T_{cap} = 1.1$ e $T_{dis} = 0.1$, onde nenhum dos coeficientes marcados possui distorções maiores que 1 JND.

5.4.1 Resultados para o codificador JPEG

Nas simulações com codificador JPEG são usadas três conhecidas imagens de teste: Lena, Mandrill e Peppers. O método de inserção, QIM-Watson, utiliza a DCT e o modelo perceptual de Watson para blocos 8x8. A robustez, $\Delta_{i,j}$, equivale a matriz de quantização definida pelo valor de QF_w (Fator de Qualidade).

$$q = \begin{cases} \frac{50.0}{QF_w}, & \text{se } QF_w < 50 \\ 2.0 - 0.02 \times QF_w, & \text{Nos demais pontos} \end{cases} \quad (5.20)$$

$$\Delta_{i,j} = q \times m_{i,j} \quad (5.21)$$

$$m_{i,j} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 98 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 48 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \quad (5.22)$$

Onde $m_{i,j}$ é a matriz de quantização utilizada no codificador JPEG.

Robustez

As Figuras 5.3, 5.4 e 5.5 ilustram o BER (*Bit Error Rate*) do método de inserção para três níveis de robustez, $QF_w = 10, 20$ e 40 . Na simulação, a imagem marcada é codificada e decodificada por todos valores de QF do JPEG.

Observa-se pelas Figuras 5.3, 5.4 e 5.5 que somente a partir dos valores de QF menores que o nível de robustez determinado, QF_w , ocorre o aumento na quantidade de bits detectados incorretamente. A razão disso está na propriedade de quantização uniforme descrita a seguir.

Seja α um valor real e q_1 e q_2 valores inteiros de passos de quantização, com $q_2 \leq q_1$. Logo, se

$$\alpha \diamond q = q \left[\frac{\alpha}{q} + 0.5 \right], \quad (5.23)$$

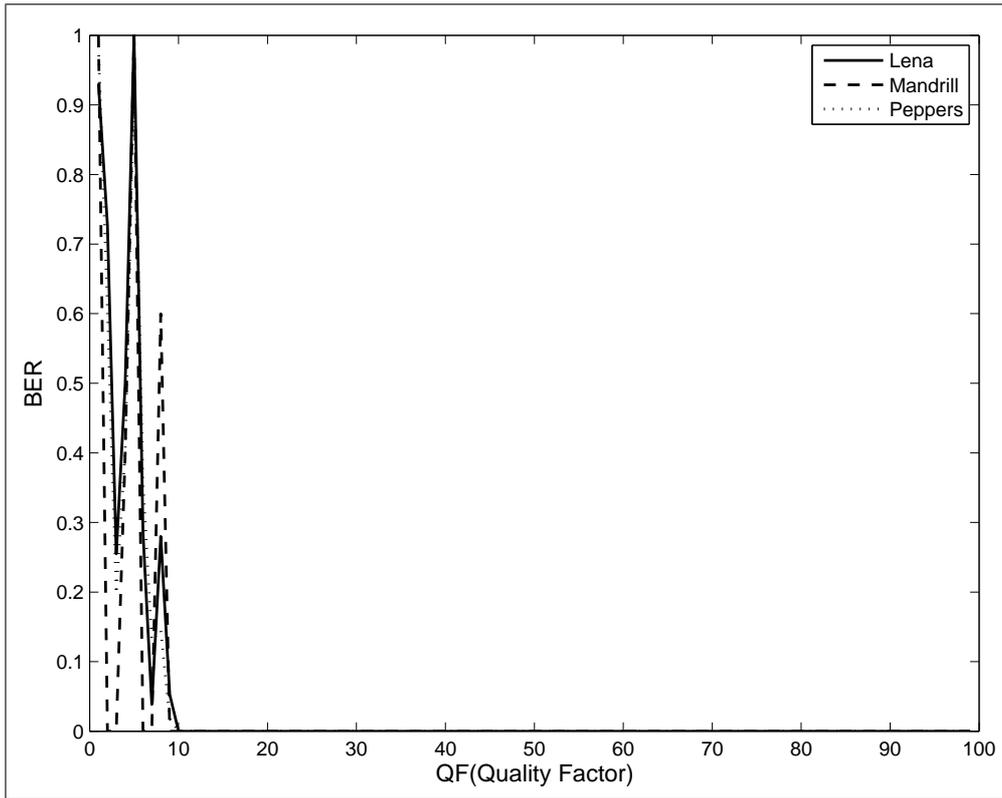


Figura 5.3: BER do método de inserção com robustez de $QF = 10$ e $T_{bloco} = 0$.

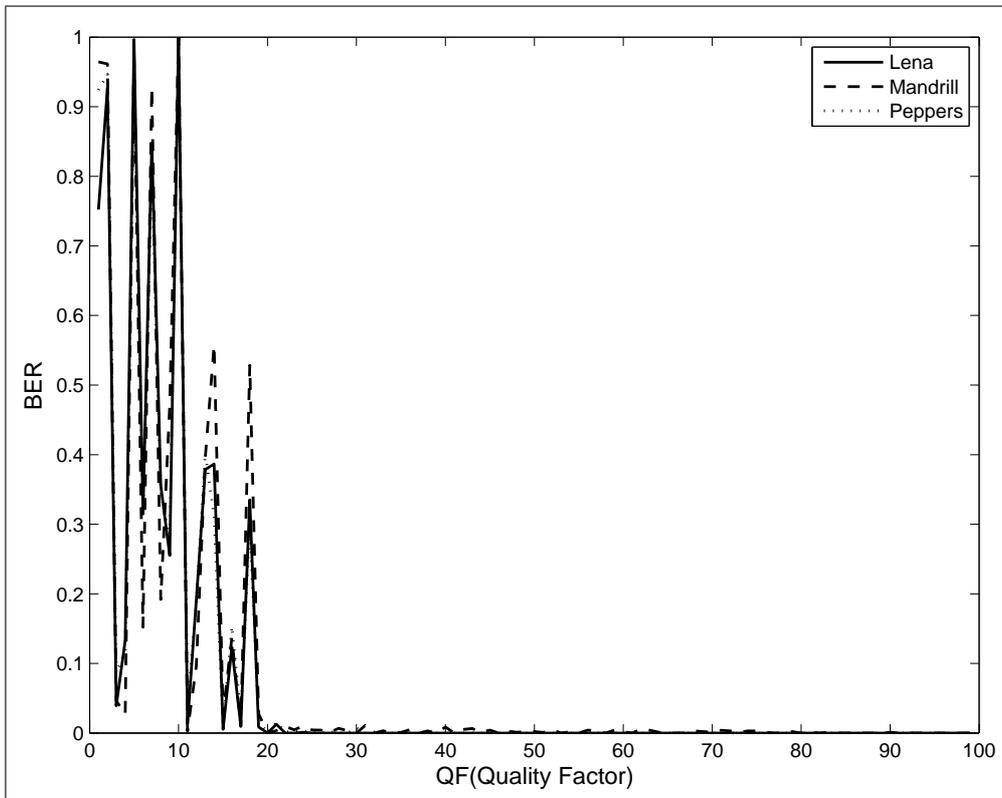


Figura 5.4: BER do método de inserção com robustez de $QF = 20$ e $T_{bloco} = 0$.

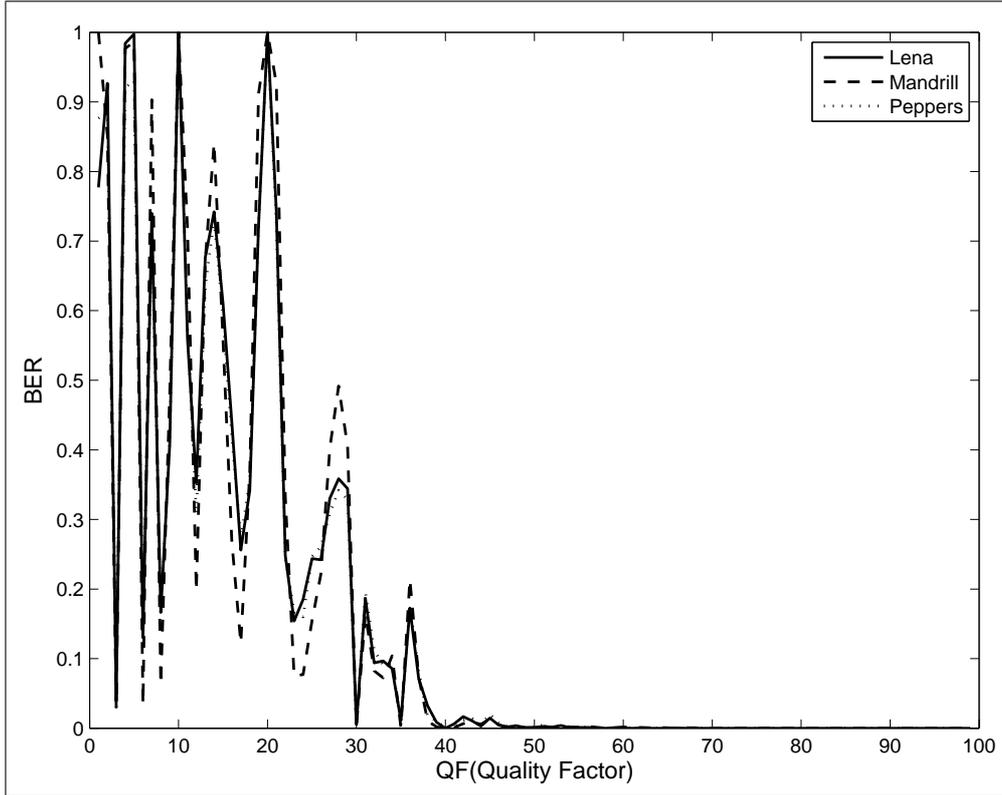


Figura 5.5: BER do método de inserção com robustez de $QF = 40$ e $T_{bloco} = 0$.

então

$$((\alpha \diamond q_1) \diamond q_2) \diamond q_1 = \alpha \diamond q_1 \quad (5.24)$$

Portanto, se α é quantizado por q_1 e em seguida quantizado por q_2 , enquanto $q_2 \leq q_1$, o efeito da segunda quantização é cancelado, quando se quantiza novamente por q_1 . Então, se durante a codificação os passos de quantização forem menores que $\Delta_{i,j}$, a marca d'água do método de inserção proposto resistirá, visto que o erro máximo de quantização uniforme será sempre menor que $\pm \frac{\Delta_{i,j}}{2}$.

A Figura 5.6 mostra a robustez da marca d'água na imagem de teste, *Mandrill*, para diferentes valores de T_{bloco} . Nota-se que nos valores de QF abaixo do nível de robustez, QF_w , há uma maior robustez nos maiores valores de T_{bloco} .

Capacidade

A Tabela 5.2 mostra a quantidade de bits em cada imagem de teste para os três níveis de robustez. Como a qualidade perceptual é fixa, $T_{cap} = 1.1$ e $T_{dis} = 0.1$, observa-se que quanto maior a robustez, menor a quantidade de bits. Para obter

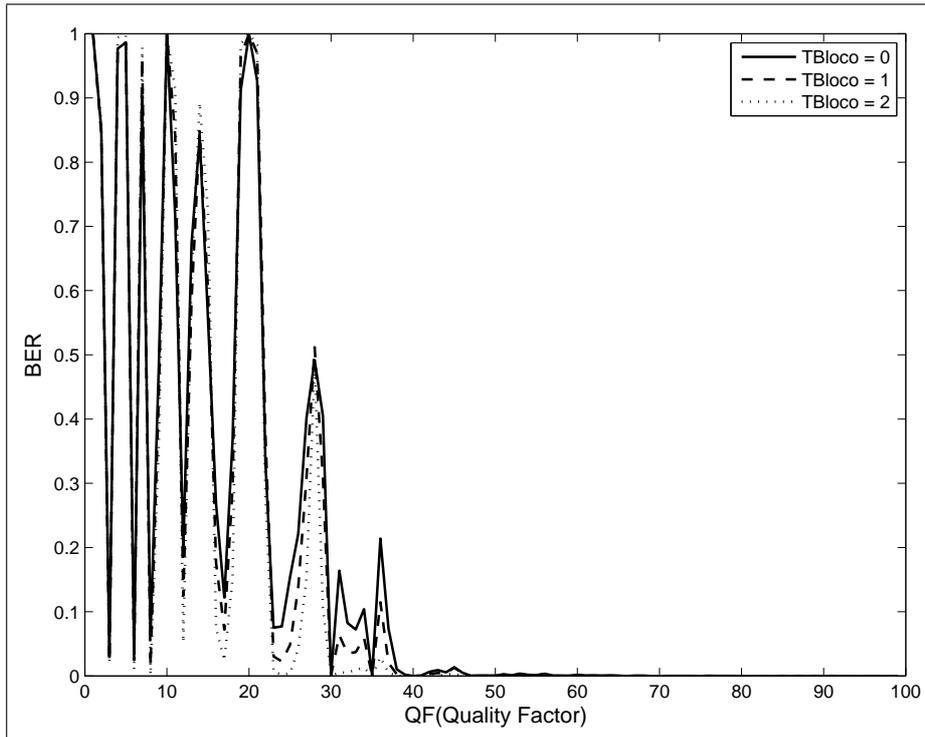


Figura 5.6: BER do método de inserção com robustez de $QF = 40$ e $T_{bloco} = 0, 1$ e 2 .

uma maior quantidade de bits por imagem é preciso diminuir o valor do parâmetro T_{cap} , que provocaria um aumento na distorção da imagem em benefício de uma maior capacidade.

Tabela 5.2: *Capacidade do sistema de marca d'água*

QF	LENA	PEPPERS	MANDRILL
10	42	67	3
20	862	844	1029
40	3114	2986	7341

Transparência

A transparência da imagem marcada em relação à imagem original é avaliada pelo PSNR e pela métrica de Watson para diferentes níveis de robustez, nas Tabelas 5.3 e 5.4. Para uma avaliação subjetiva, a Figura 5.7 ilustra a imagem original, a imagem marcada com uma robustez de $QF_w = 20$ e a codificação de ambas as imagens por um $QF = 20$ do JPEG.

Observa-se na maioria dos resultados das Tabelas 5.3 e 5.4 que quanto maior a robustez, menor a distorção da imagem. Isto acontece porque para uma qualidade perceptual fixa, uma menor quantidade de coeficientes são modificados à medida que ocorre um aumento na robustez da marca d'água.



(a) Imagem original.



(b) Imagem marcada com robustez de $QF = 20$.



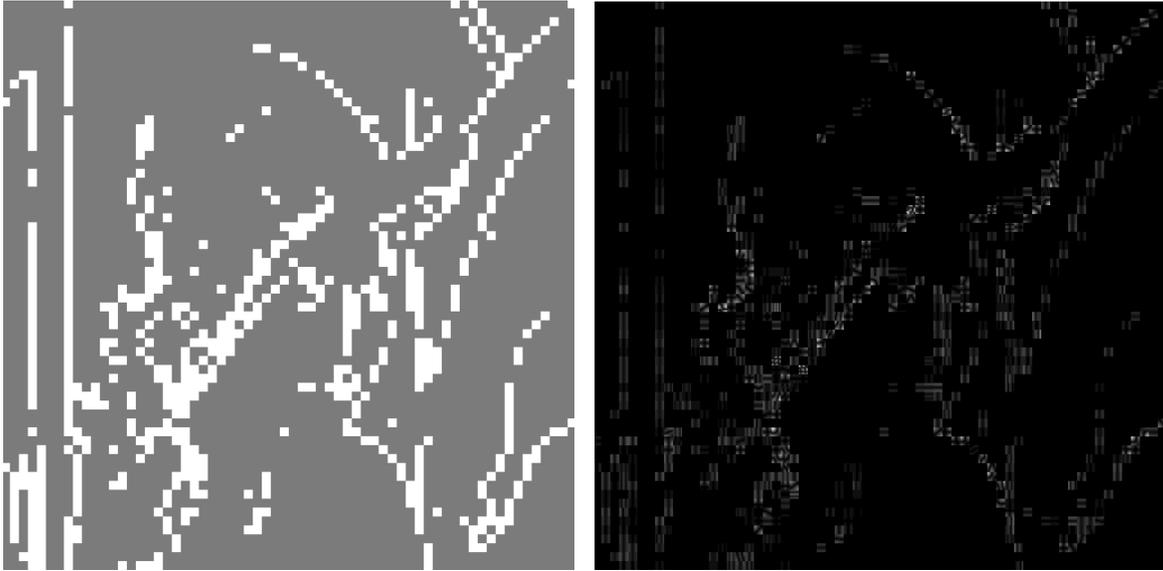
(c) Imagem original codificada pelo JPEG com $QF = 20$.



(d) Imagem marcada e codificada pelo JPEG com $QF = 20$.

Figura 5.7: Imagem original, imagem marcada e imagem codificada pelo JPEG.

A Figura 5.8 mostra os blocos utilizados na detecção e também a diferença entre a imagem original e a imagem marcada multiplicada por 15.



(a) Os blocos na cor branca são utilizados na detecção. (b) Diferença entre a imagem original e a imagem marcada mutiplicada por 15.

Figura 5.8: Blocos utilizados na detecção e a diferença entre a imagem original e a imagem marcada.

Tabela 5.3: *PSNR das imagens marcadas em relação à imagem original.*

QF	LENA	PEPPERS	MANDRILL
10	52.934410	51.742584	64.846748
20	48.177425	47.628521	45.652763
40	47.267593	47.930130	42.052120

Tabela 5.4: *Métrica de Watson das imagens marcadas em relação à imagem original.*

QF	LENA	PEPPERS	MANDRILL
10	1.955946	2.443965	0.944098
20	3.550950	4.000110	7.966116
40	5.027602	5.403074	5.698861

5.4.2 Resultados para o codificador H.264-Intra

Pelos testes realizados no codificador JPEG, nota-se que o conhecimento do processo de quantização do codificador JPEG é utilizado para obter um melhor desempenho do método de inserção proposto, visto que passos de quantização, $\Delta_{i,j}$, são escolhidos a partir da matriz de quantização, $m_{i,j}$ na Equação 5.22. No entanto, as

normas dos padrões de compressão com perda definem apenas o processo de decodificação, permitindo os projetos de codificadores terem diferentes implementações de quantização. Dessa forma, a distorção introduzida pelo processo de quantização é dependente da implementação de cada codificador.

Para as simulações com o codificador H.264 utiliza-se o programa de referência do H.264 (JM 14.2) e os vídeos de teste *Foreman*, *Hall* e *Mobile* de 300 quadros na resolução CIF (352x288). Estes vídeos são codificados somente no modo Intra porque o modelo perceptual do método de inserção proposto aborda apenas o domínio espacial, enquanto a abordagem de um modelo perceptual incluindo o domínio temporal é considerada para trabalhos futuros.

Como a transformada do H.264 utiliza blocos 4x4, o método de inserção, QIM-Watson, para autenticar os vídeos codificados pelo H.264-Intra utiliza a DCT aproximada e o modelo perceptual de Watson para blocos 4x4.

Robustez

Para avaliar a robustez do método de inserção, os vídeos de teste são marcados em três níveis de robustez, $QP_w = 30, 33$ e 36 , e depois codificados e decodificados no modo Intra por três valores de QP. Nos testes, a robustez $\Delta_{i,j}$, equivale ao passo de quantização definido pelo valor de QP (Parâmetro de Quantização), visto na seção 3.10.2.

A Figura 5.9 mostra o BER dos vídeos de teste marcados com robustez de $QP_w = 30$ que foram codificados e decodificados pelo H.264 no modo Intra para três valores de QP = 28, 30 e 32.

A Figura 5.10 mostra o BER dos vídeos de teste marcados com robustez de $QP_w = 33$ que foram codificados e decodificados pelo H.264 no modo Intra para três valores de QP = 31, 33 e 35.

A Figura 5.11 mostra o BER dos vídeos de teste marcados com robustez de $QP_w = 36$ que foram codificados e decodificados pelo H.264 no modo Intra para três valores de QP = 34, 36 e 38.

Pelas Figuras 5.9, 5.10, 5.11, observa-se que os vídeos marcados com robustez, QP_w , somente apresentam o BER menor que 10% quando o valor de QP do codificador é de pelo menos dois níveis de quantização abaixo de QP_w . Isso acontece porque

a quantização uniforme utilizada no programa de referência do H.264 (JM 14.2), mostrada na Equação 5.25, possui um parâmetro de controle de arredondamento, f , que dependendo do seu valor permite a ocorrência de erros de quantização maiores que $\frac{\Delta_{i,j}}{2}$. Na Equação 5.25, a função $sign()$ retorna o sinal do coeficiente, $X_{i,j,k}$.

$$X'_{i,j,k} = \Delta_{i,j} \left[\frac{|X_{i,j,k}|}{\Delta_{i,j}} + f \right] sign(X_{i,j,k}) \quad (5.25)$$

No programa de referência do H.264 (JM 14.2), o parâmetro de controle de arredondamento, f , tem o valor de $\frac{1}{3}$ para os quadros Intra e $\frac{1}{6}$ para os quadros Inter.

Outros fatores que podem influenciar na detecção incorreta dos bits são os limites da faixa dinâmica da imagem e a modificação dos coeficientes da DCT quando se transforma do domínio da frequência para domínio espacial.

Capacidade

A Figura 5.12 mostra a quantidade de bits nos vídeos de teste marcados com qualidade perceptual fixa, $T_{cap} = 1.1$ e $T_{dis} = 0.1$, e em três níveis de robustez. Nota-se que para o mesmo nível de robustez, os vídeos de teste marcados apresentam em cada quadro diferentes quantidades de bits.

Transparência

Na Figura 5.13, observa-se o PSNR dos vídeos marcados em diferentes níveis de robustez. É importante notar que o PSNR dos vídeos de teste apresentam altos valores por causa dos parâmetros fixados, $T_{cap} = 1.1$ e $T_{dis} = 0.1$.

Para uma avaliação subjetiva, a Figura 5.14 ilustra o quadro original do *Hall*, o quadro marcado com uma robustez de $QP_w = 33$ e a codificação de ambos os quadros por um $QP = 31$ no H.264. A Figura 5.15 mostra a diferença entre o quadro original e o quadro marcado multiplicada por 15.

5.5 Conclusão

Nas simulações com os codificadores JPEG e H.264, os parâmetros do sistema de marca d'água proposto foram definidos de modo a avaliar três níveis de robustez,

$\Delta_{i,j}$, e uma qualidade perceptual fixa, $T_{cap} = 1.1$ e $T_{dis} = 0.1$. Os resultados dessas simulações mostraram um interessante balanceamento entre a robustez, a transparência e a capacidade do sistema de marca d'água proposto.

Pelas simulações observou-se que uma ótima relação de custo-benefício das propriedades conflitantes é obtida quando as informações da quantização do codificador são conhecidas, mas mesmo sem essas informações, o método de inserção, QIM-Watson, apresentou resultados satisfatórios para aplicações de autenticação que exigem imagens ou vídeos de alta qualidade.

A contribuição deste capítulo foi a elaboração de um sistema de marca d'água adaptativo com a finalidade de autenticar tanto imagens quanto vídeos codificados com os padrões de compressão com perda baseados na DCT (JPEG, MPEG-2 e H.264).

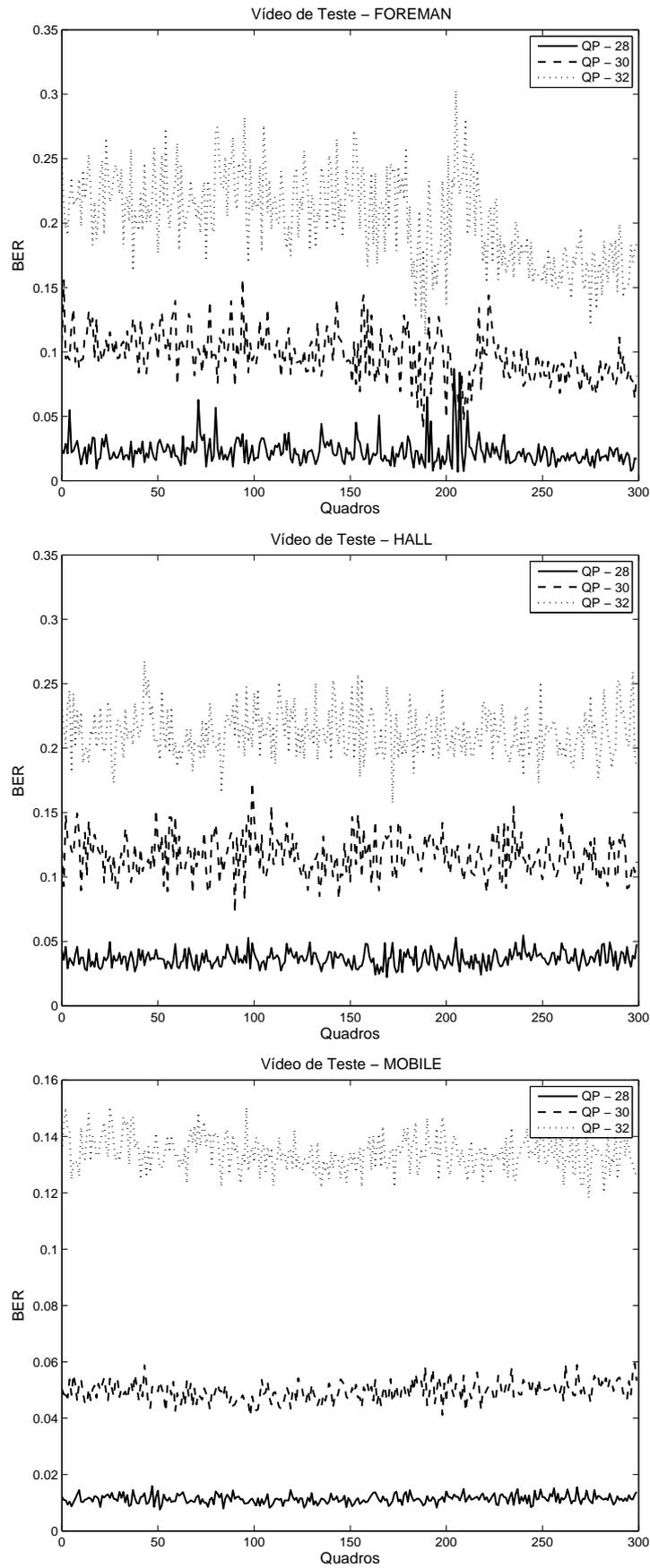


Figura 5.9: BER do sistema de marca d'água com robustez de $QP = 30$ e $T_{\text{bloco}} = 0$.

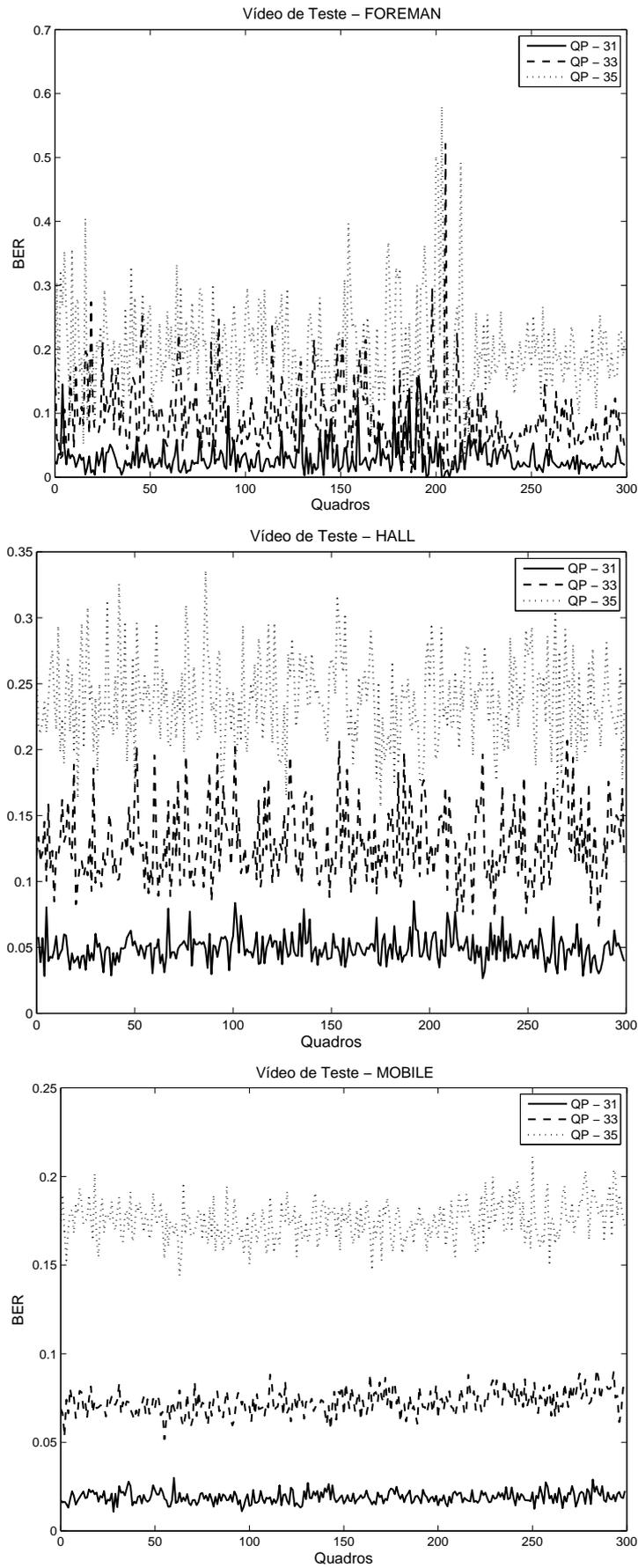


Figura 5.10: BER do sistema de marca d'água com robustez de QP = 33 e $T_{\text{bloco}} = 0$.

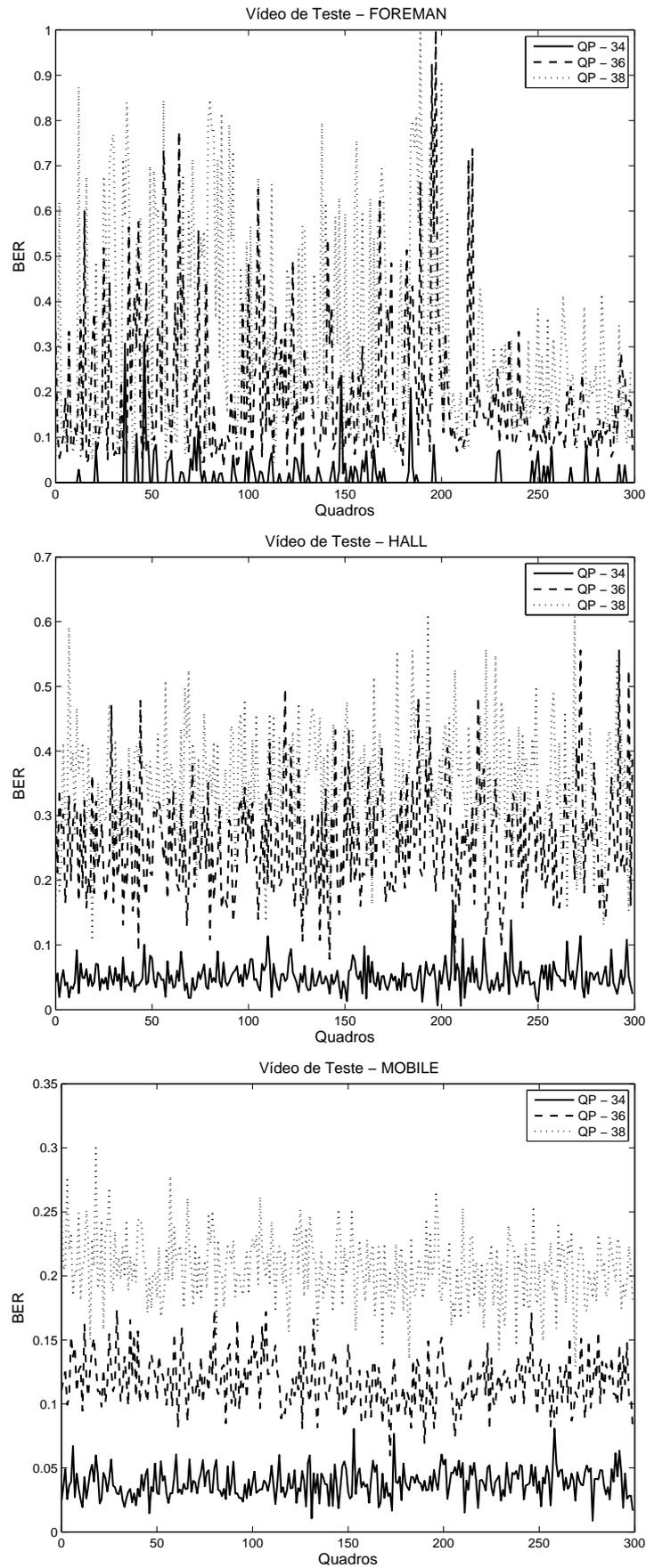


Figura 5.11: BER do sistema de marca d'água com robustez de QP = 36 e $T_{\text{bloco}} = 0$.

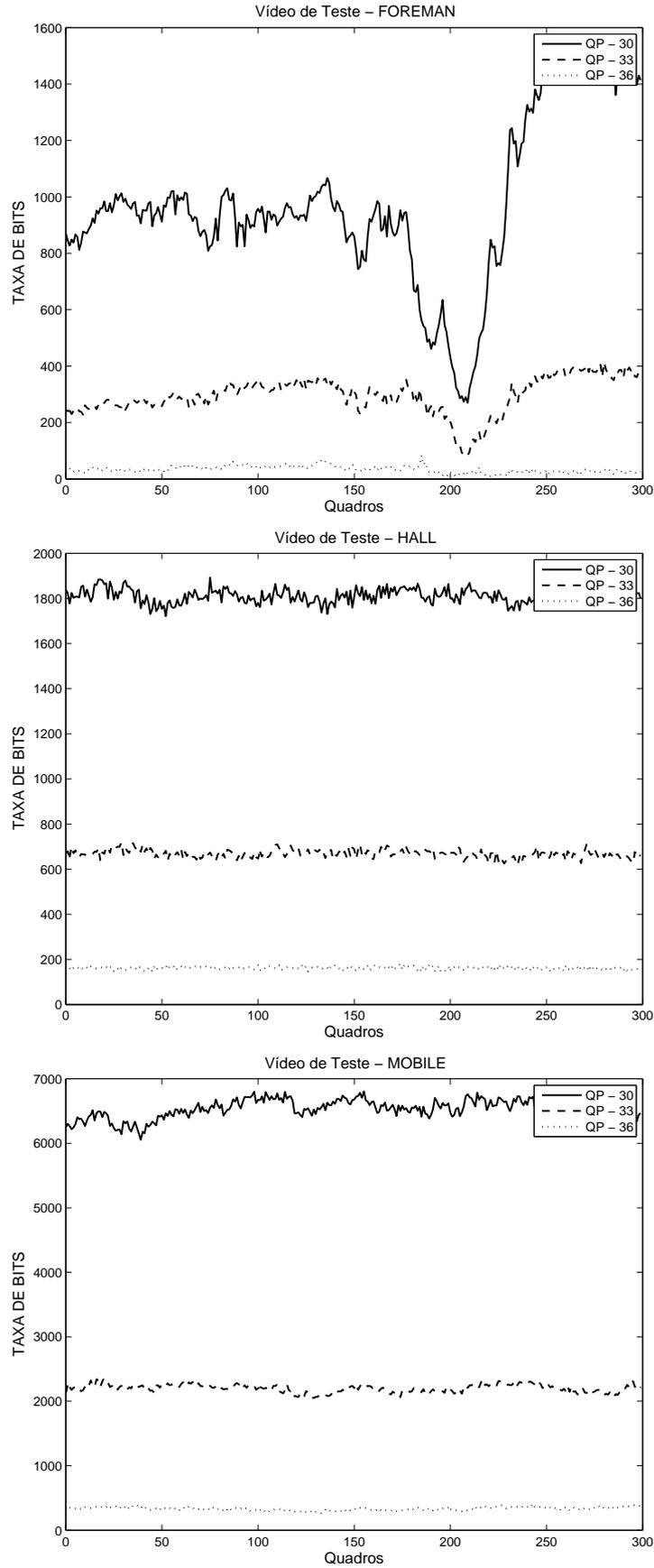


Figura 5.12: Capacidade dos vídeos marcados em diferentes níveis de robustez.

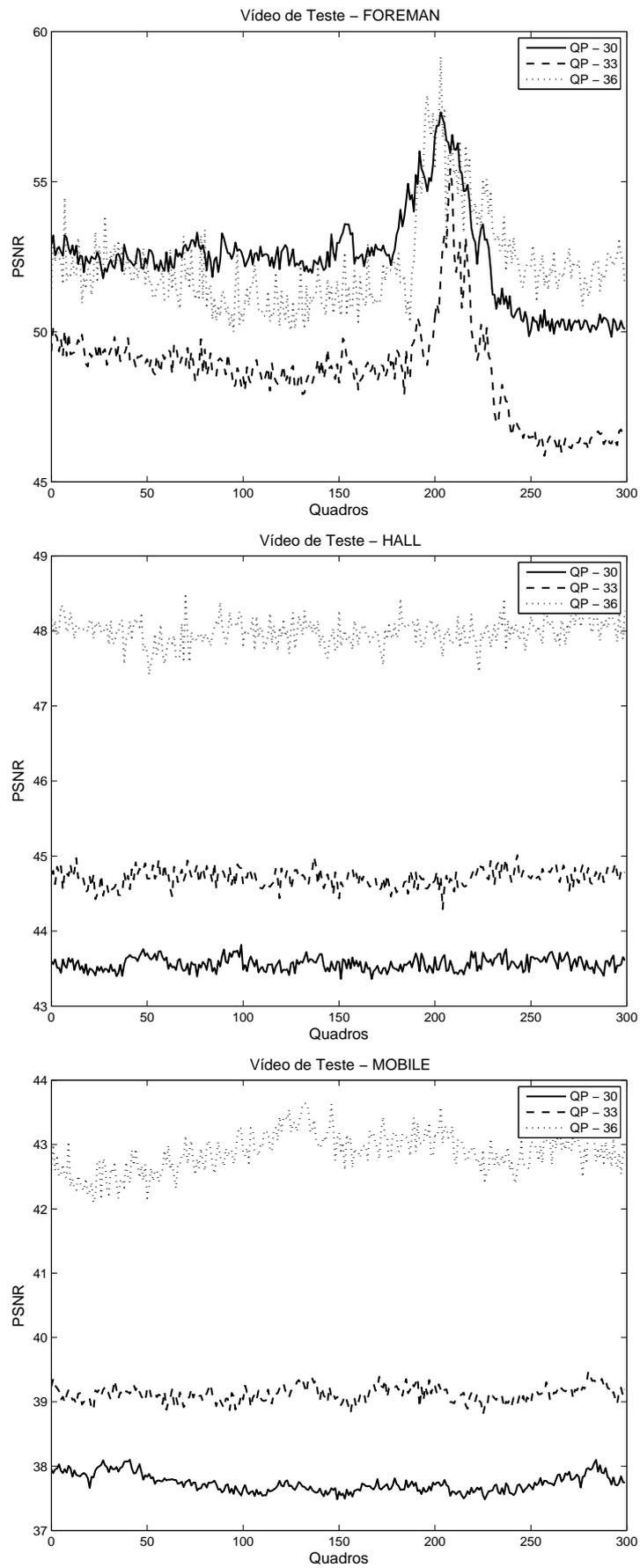


Figura 5.13: PSNR dos vídeos marcados em diferentes níveis de robustez.



(a) Quadro original.



(b) Quadro marcado com robustez de $QP = 33$.



(c) Quadro original codificado pelo H.264 com $QP = 31$.



(d) Quadro marcado e codificado pelo H.264 com $QP = 31$.

Figura 5.14: Quadro original, quadro marcado e quadro codificado pelo H.264.



Figura 5.15: A diferença entre o quadro original e o quadro marcado multiplicada por 15.

Capítulo 6

Conclusão

Este trabalho apresentou dois sistemas de marca d'água para autenticação de vídeo no padrão H.264. O primeiro sistema de marca d'água apresentado no capítulo 4 tem o intuito de obter uma autenticação exata através da inserção de uma marca d'água frágil no *bitstream* do padrão H.264 no perfil *Baseline*. Enquanto o segundo sistema de marca d'água apresentado no capítulo 5 tem a finalidade de inserir uma marca d'água semi-frágil no domínio da frequência que resista às distorções provocadas pelos padrões de compressão com perda baseados na DCT. As simulações desses sistemas de marca d'água foram feitas com o objetivo de avaliar as conflitantes propriedades: robustez, transparência e capacidade.

Nas simulações do Capítulo 4 observou-se que método de inserção proposto em [1] provoca uma intensa degradação visual no vídeo e que a proposta para redução do acúmulo de erros resultou numa maior transparência da marca d'água tanto pela avaliação objetiva, PSNR, quanto pela avaliação subjetiva. Um ponto desfavorável da aplicação de tal técnica de redução do acúmulo de erros no método de inserção proposto em [1] é o aumento da complexidade computacional, visto que na seleção dos blocos é utilizado um algoritmo de ordenação.

Os resultados das simulações com o sistema de marca d'água semi-frágil adaptativo do Capítulo 5 mostraram um interessante balanceamento das propriedades: robustez, transparência e capacidade. Isso aconteceu porque foi inserido um conhecimento perceptual ao método de inserção, QIM, com a finalidade de selecionar os coeficientes da DCT que possuíssem uma maior tolerância à distorção. Além disso, observou-se que o método de inserção proposto tem a possibilidade de se adaptar

às informações da quantização dos diferentes codificadores baseados na DCT, assim podendo obter uma ótima relação de custo-benefício entre as propriedades conflitantes.

Uma proposta para trabalho futuro seria a substituição do modelo perceptual de Watson do sistema de marca d'água semi-frágil adaptativo por um modelo perceptual que inclua tanto o domínio espacial quanto temporal. Outra idéia de trabalho futuro seria a criação de uma marca d'água que se adaptasse as características da imagem, pois se obteria uma maior segurança e um melhor desempenho na autenticação. Vale ressaltar que essas características extraídas da imagem para criação da marca d'água deverão ser robustas aos padrões de compressão baseados na DCT.

Referências Bibliográficas

- [1] KIM, S.M., K. S. H. Y. W. C., “Data hiding on. H.264/AVC compressed video”.
In: *Proceedings of ICIAR 2007*, 2007.
- [2] CHEN, B., W. G. W., “Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding”,
IEEE Transactions On Information Theory, v. 47, pp. 1423–1443, Maio 2001.
- [3] WATSON, A., “DCT quantization matrices optimized for individual images”.
In: *Proceedings of SPIE*, 1993.
- [4] JOHNSON, N. F., *Steganography*, Tech. rep., 1995.
- [5] DOERR, G., D. J. L., “A guide tour of video watermarking”, *Signal Processing: Image Communication*, v. 18, n. 4, pp. 263–282, Abril 2003.
- [6] COX, I. J., M. M. L. B. J. A., *Digital Watermarking*. 1st ed. Academic Press, 2002.
- [7] WOLFGANG, R. B., P. C. I. D. E. J., “Perceptual watermarks for digital images and video”. In: *Proceedings of the IEEE*, 1999.
- [8] XIAO, F., *DCT-based Video Quality Evaluation*, Projeto final, Stanford University, 2000.
- [9] WANG, Z., B. A. C. S. H. R. S. E. P., “Image Quality Assessment: From Error Visibility to Structural Similarity”, *IEEE Transactions on Image Processing*, v. 13, n. 4, Abril 2004.
- [10] BARNI, M., B. F., *Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications*. 1st ed. Marcel Dekker, Inc, 2004.

- [11] COX, I. J., K. J. L. T. S. T., “Secure spread spectrum watermarking for multimedia”, *IEEE Transactions on Image Processing*, v. 6, n. 12, pp. 1673–1687, Dezembro 1997.
- [12] NOORKAMI, M., *Secure and Robust Compressed-Domain Video Watermarking For H.264*, Tese de doutorado, Georgia Institute of Technology, Agosto 2007.
- [13] SCHLAUWEG, M., P. D. P. T. M. E., “Quantization-based semi-fragile public-key watermarking for secure image authentication”. In: *Proceedings of SPIE*, 2005.
- [14] LIN, C. H., S. T. S. H. W. S., “Semi-Fragile Watermarking Scheme for Authentication of JPEG Images”, *Tamkang Journal of Science and Engineering*, v. 10, n. 1, pp. 57–66, 2007.
- [15] RICHARDSON, I., *H.264 and MPEG-4 Video Compression*. 1st ed. Wiley, 2003.
- [16] HANZO, L., C. P. J. S. J., *Video Compression and Communications From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. 2nd ed. Wiley, 2007.
- [17] WIEGAND, T., S. G. J. B. G. L. A., “Overview of the H.264/AVC Video Coding Standard”, *IEEE Transactions On Circuits and Systems For Video Technology*, v. 13, n. 7, pp. 560–576, Julho 2003.
- [18] SULLIVAN, G. J., T. P. L. A., “The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions”. 2004.
- [19] KAMACI, N., A. Y., “Performance Comparison of the Emerging H.264 Video Coding Standard with the Existing Standards”, *IEEE Transactions On Circuits and Systems For Video Technology*, v. 13, n. 7, pp. 560–576, Julho 2003.

- [20] RAJA, G., M. M. J., “In-loop Deblocking Filter for H.264/AVC Video”. In: *Second International Symposium on Communications, Control and Signal Processing*, 2006.
- [21] TOIVONEN, T., H. J., “Fast full search block motion estimation for H.264/AVC with multilevel successive elimination algorithm”. 2004.
- [22] ZHU, C., L. X. C. L. P., “Hexagon-based search pattern for fast block motion estimation”, *IEEE Transactions On Circuits and Systems For Video Technology*, v. 12, pp. 349–355, Maio 2002.
- [23] MARPE, D., S. H. W. T., “Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard”, *IEEE Transactions On Circuits and Systems For Video Technology*, v. 13, n. 7, pp. 620–636, Julho 2003.
- [24] LANGELAAR, G. C., L. R. L. B. J., “Real-time Labeling of MPEG-2 Compressed Video”, *Journal of Visual Communication and Image Representation*, v. 9, n. 4, pp. 256–270, Dezembro 1998.
- [25] LANGELAAR, G. C., *Real-time Watermarking Techniques for Compressed Video Data*, Tese de doutorado, Delft University of Technology, Países Baixos, Janeiro 2000.
- [26] HARTUNG, F., G. B., “Watermarking of Uncompressed and Compressed Video”, *Signal Processing*, v. 66, n. 3, pp. 283–301, Maio 1998.
- [27] RAMASWAMY, N., R. K. R., “Video authentication for H.264/AVC using digital signature standard and secure hash algorithm”. In: *Proceedings of the 2006 international workshop on Network and operating systems support for digital audio and video*, 2006.
- [28] KIM, J., B. Y. C. J., “Semi-fragile Watermarking Based on Dither Modulation”, *Advances in Multimedia Information Processing*, v. 3333, pp. 721–730, Outubro 2004.

- [29] LI, Q., C. I. J., “Using Perceptual Models to Improve Fidelity and Provide Resistance to Volumetric Scaling for Quantization Index Modulation Watermarking”, *IEEE Transactions On Information Forensics And Security*, v. 2, pp. 127–139, Junho 2007.
- [30] LI, Q., C. I. J. D. G., “Spread Transform Dither Modulation using a Perceptual Model”, *IEEE - Multimedia Signal Processing*, pp. 98–102, Outubro 2006.
- [31] YU, L., S. S., “Image Robust Hashing based on DCT Sign”, *Intelligent Information Hiding and Multimedia Signal Processing*, pp. 131–134, Dezembro 2006.
- [32] SCHNEIDER, M., C. S. F., “A robust content based digital signature for image authentication”, *IEEE Int. Conf. Image Processing*, v. 3, pp. 227–230, Maio 1996.
- [33] SU, Y., X. J. V. A. S. H., “Efficient MPEG-2 to H.264/AVC intra transcoding in transform-domain”, *IEEE International Symposium on Circuits and Systems (ISCAS)*, v. 2, pp. 1234–1237, Maio 2005.