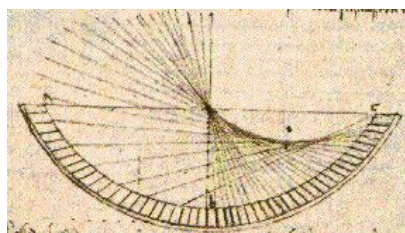


UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO



Assinatura Digital de Imagens

Baseada em Espalhamento de Espectro

Tomás Gomes da Silva Serpa Brandão
(Licenciado)

Dissertação para obtenção do Grau de Mestre
em Engenharia Electrotécnica e de Computadores

Orientadora: Prof^a. Doutora Maria Paula dos Santos Queluz Rodrigues

Presidente: Prof. Doutor Mário Alexandre Teles de Figueiredo

Vogais: Prof^a. Doutora Maria Paula dos Santos Queluz Rodrigues
Prof. Doutor Francisco António Bucho Cercas

Maio de 2002

Resumo

A assinatura digital de produtos multimédia (imagens, vídeo, som ou texto) com recurso a marcas-de-água, consiste na inserção de informação adicional de forma imperceptível, sobre esses produtos. Esta tecnologia tem sido proposta como forma de assegurar novas funcionalidades relacionadas com a protecção da informação digital, tais como a gestão de direitos de autor, a autenticação de conteúdos e o controlo de cópias. Poderá também constituir um mecanismo para inserção de informação descritiva ou de referência (e.g., *meta-data*) num dado sinal.

Muitos algoritmos de marcas-de-água para imagens fixas e vídeo recorrem aos princípios do espalhamento de espectro: um sinal de banda estreita (marca) é transmitido num canal de banda larga (imagem ou vídeo), sujeito a ruído e distorção. Seguindo esta abordagem, é possível estabelecer uma analogia entre um sistema de assinatura de imagens e um sistema de comunicações, o que sugere a utilização de conceitos e resultados da teoria da comunicação digital. O principal assunto em análise nesta tese é a avaliação do impacto que a utilização de *modulação multinível*, *codificação de canal* e *combinação de sinal*, tem no desempenho dos algoritmos de marcas-de-água baseados em espalhamento de espectro.

Palavras-chave:

Marcas-de-água; Espalhamento de espectro; Modulação multinível; Codificação de canal; Detecção óptima; Combinação de sinal.

Abstract

Watermarking has been proposed as a technology providing security systems with new features, such as copyright management, content-based authentication and playback/copy control. It may also provide a mechanism for embedding descriptive or reference information (e.g., meta-data) in a given signal. The concept of digital watermarking is to embed a digital mark into the information (audio, image or video) in a statistical and perceptually undetectable way.

Most watermarking methods use a so-called *spread spectrum* approach: a narrowband signal (the watermark information) has to be transmitted via a wideband channel that is subject to noise and distortion (the multimedia host data, e.g., still images and video). Under this approach, digital watermarking can be treated as a communication problem, and concepts and techniques from communication theory can be used. The application of some of those techniques, namely *multilevel modulation*, *channel coding* and *signal combining*, to spread spectrum based digital watermarking, is the main subject under analysis in this thesis.

Keywords:

Watermarking; Spread spectrum; Multilevel modulation, Channel coding; Optimal detection; Signal combining.

Agradecimentos

Durante o período de desenvolvimento desta tese, beneficiei da colaboração de diversas pessoas que me ajudaram a ultrapassar dificuldades e obstáculos. Pretendo deste modo demonstrar a minha gratidão para com essas pessoas.

Em primeiro lugar, gostaria de agradecer à minha orientadora científica, a Professora Maria Paula Queluz, que contribuiu activamente para a realização deste trabalho, apoiando-me incondicionalmente, dando-me diversos conselhos úteis e mostrando sempre grande disponibilidade para discussões científicas (muitas das quais deram os seus frutos).

Gostaria também de agradecer ao Professor António Rodrigues o tempo que dispensou para apoio à realização deste trabalho, nomeadamente para esclarecimento de dúvidas e para resolução de “problemas de logística”.

Um grande obrigado a todo o Grupo de Imagem do IT-Lisboa pela ajuda prestada. Agradeço em particular aos Engenheiros João Ascenso, Luís Ducla, Paulo Nunes e Paulo Correia, que tantas vezes me ajudaram a resolver problemas de ordem diversa. O bom ambiente e o espírito de camaradagem existentes neste grupo possibilitaram também que o desenvolvimento do trabalho tenha sido realizado de uma forma bastante agradável.

Um grande agradecimento para o pessoal da sala de bolseiros do 10º piso do IT-Lisboa, em especial para a Eng^a. Gabriela Marques, o Eng. Paulo Marques e o Doutor Fernando Velez, pelas pequenas (mas preciosas) ajudas e pelo apoio moral que me deram.

Para finalizar, gostaria de agradecer à Márcia toda a sua paciência e compreensão pelo tempo que sacrificámos.

Índice

RESUMO	III
ABSTRACT	III
AGRADECIMENTOS	III
ÍNDICE	III
ÍNDICE DE FIGURAS	III
ÍNDICE DE TABELAS	III
CAPÍTULO 1 INTRODUÇÃO	3
CAPÍTULO 2 ASSINATURA DE IMAGENS COM MARCAS-DE-ÁGUA	3
2.1 INTRODUÇÃO.....	3
2.2 BREVE HISTORIAL DAS MARCAS-DE-ÁGUA	3
2.3 PRINCÍPIOS DE FUNCIONAMENTO E CONCEITOS RELACIONADOS	3
2.3.1 Esquema geral de um sistema de assinatura de imagens	3
2.3.2 Classes de marcas-de-água	3
2.3.3 Qualidade visual	3
2.3.4 Fiabilidade na extracção	3
2.3.5 Fiabilidade na detecção	3
2.3.6 <i>Payload</i>	3
2.3.7 Complexidade	3
2.3.8 Formas de ataque e sua caracterização	3
2.3.9 Segurança.....	3
2.4 ÁREAS DE APLICAÇÃO DAS MARCAS-DE-ÁGUA.....	3
2.4.1 Prova do proprietário	3
2.4.2 Identificação.....	3
2.4.3 Autenticação	3
2.4.4 Monitorização da difusão/distribuição.....	3

2.4.5 Avaliação da integridade.....	3
2.4.6 Impressões digitais (<i>fingerprinting</i>).....	3
2.4.7 Controlo de utilização	3
2.4.8 Transporte de informação adicional.....	3
2.4.9 Síntese de requisitos.....	3
2.5 “AMBIGUIDADES” NAS MARCAS-DE-ÁGUA.....	3
CAPÍTULO 3 ESPALHAMENTO DE ESPECTRO EM MARCAS-DE-ÁGUA	3
3.1 INTRODUÇÃO	3
3.2 ESPALHAMENTO DE ESPECTRO: PRINCÍPIOS BÁSICOS E APLICAÇÃO À ASSINATURA DE IMAGENS	3
3.3 INSERÇÃO DA MARCA-DE-ÁGUA.....	3
3.3.1 Esquema geral de inserção.....	3
3.3.2 Modulador multinível	3
3.3.3 Mapeamento bidimensional	3
3.3.4 Modelo perceptual.....	3
3.4 EXTRACÇÃO DA MARCA-DE-ÁGUA	3
3.5 MARCAS-DE-ÁGUA NO DOMÍNIO ESPACIAL	3
3.5.1 Inserção e extracção da marca-de-água no domínio espacial.....	3
3.5.2 Análise do desempenho	3
3.5.3 Resultados	3
3.6 MARCAS-DE-ÁGUA NO DOMÍNIO DA FREQUÊNCIA	3
3.6.1 Inserção e extracção da marca-de-água no domínio da frequência.....	3
3.6.2 Modelo perceptual para o domínio da frequência.....	3
3.6.3 Modelo estatístico dos coeficientes DCT	3
3.6.4 Estrutura do desmodulador	3
3.6.5 Análise do desempenho	3
3.6.6 Resultados	3
3.7 COMPARAÇÃO ENTRE RESULTADOS REFERENTES AOS DOIS DOMÍNIOS	3
3.8 CONSIDERAÇÕES FINAIS.....	3
CAPÍTULO 4 CODIFICAÇÃO PARA CORRECÇÃO DE ERROS	3
4.1 INTRODUÇÃO	3

4.2 PRINCIPAIS CLASSES DE CODIFICAÇÃO DE CANAL.....	3
4.2.1 Taxonomia	3
4.2.2 Códigos de bloco	3
4.2.3 Códigos convolucionais binários	3
4.3 APLICAÇÃO DE CODIFICAÇÃO DE CANAL A MARCAS-DE-ÁGUA.....	3
4.3.1 Códigos de bloco binários.....	3
4.3.2 Códigos de bloco não binários	3
4.3.3 Códigos convolucionais binários	3
4.4 COMPARAÇÃO DO DESEMPENHO	3
4.5 RESULTADOS EM PRESENÇA DE COMPRESSÃO JPEG	3
4.6 RESULTADOS EM PRESENÇA DE RUÍDO BRANCO GAUSSIANO E CORTES.....	3
4.7 RESULTADOS NO DOMÍNIO DA FREQUÊNCIA	3
4.8 CONSIDERAÇÕES FINAIS	3
 CAPÍTULO 5 TÉCNICAS DE COMBINAÇÃO DE SINAL	 3
5.1 INTRODUÇÃO.....	3
5.2 COMBINAÇÃO COM LÓGICA DE MAIORIA	3
5.3 COMBINAÇÃO LINEAR	3
5.3.1 Combinação <i>ótima</i>	3
5.3.2 Combinação <i>quase ótima</i>	3
5.3.3 Combinação com pesos iguais e constantes.....	3
5.3.4 Comparação das várias estratégias de combinação linear	3
5.3.5 Estimador de máxima verosimilhança (ML)	3
5.4 RESULTADOS.....	3
5.4.1 Resultados com inserção no domínio espacial.....	3
5.4.2 Resultados com inserção no domínio da frequência	3
5.5 CONSIDERAÇÕES FINAIS	3
 CAPÍTULO 6 CONCLUSÕES	 3
 BIBLIOGRAFIA	 3
 ANEXO A LIMITE DE SHANNON	 3

ANEXO B FREQUÊNCIAS ESPACIAIS EM CICLOS POR GRAU	3
ANEXO C ESTRUTURA DO DESMODULADOR	3
ANEXO D CLASSIFICAÇÃO DE TRAMAS NA NORMA MPEG-2 VÍDEO	3

Índice de figuras

Figura 2.1 – Assinatura digital de uma imagem.....	3
Figura 2.2 – Extracção de uma marca-de-água digital.....	3
Figura 2.3 – Detecção de uma marca-de-água digital.....	3
Figura 3.1 – A tecnologia das marcas-de-água numa perspectiva de comunicação digital.	3
Figura 3.2 – Espalhamento do espectro de um sinal.	3
Figura 3.3 – Comunicação digital com modulação baseada em espalhamento de espectro.	3
Figura 3.4 – Exemplo de modulação (a) e de desmodulação (b) por espalhamento de espectro. ...	3
Figura 3.5 – Assinatura de imagens baseada em espalhamento de espectro.....	3
Figura 3.6 – Esquema geral de inserção da marca-de-água.	3
Figura 3.7 – Exemplo ilustrativo do mapeamento bidimensional: a) Sequências a mapear; b) Tabela de atribuições; c) Sequências mapeadas.....	3
Figura 3.8 – Esquema geral de extracção da marca-de-água.	3
Figura 3.9 – Esquema geral do desmodulador multinível.....	3
Figura 3.10 – Esquemas de inserção e extracção de marcas-de-água no domínio espacial: a) Inserção; b) Extracção.....	3
Figura 3.11 – Esquema do desmodulador multinível para o domínio espacial.....	3
Figura 3.12 – Imagens utilizadas nas simulações.	3
Figura 3.13 – Resultados teóricos – P_b vs. N° de Pixels / Bit de informação.	3
Figura 3.14 – Resultados experimentais – P_b vs. N° de Pixels / Bit de informação.....	3
Figura 3.15 – Resultados experimentais – P_b em presença de compressão JPEG.	3
Figura 3.16 – Ruído gaussiano na imagem <i>Lena</i> : a) Imagem marcada; b) Imagem marcada corrompida por ruído gaussiano com $\sigma_r = 10$; c) Módulo da diferença entre (a) e (b) multiplicada por 8.	3
Figura 3.17 – Resultados experimentais – P_b em presença de ruído branco gaussiano.	3
Figura 3.18 – Corte sobre a imagem <i>Lena</i> (marcada) com $L_{Crop} = 320$	3
Figura 3.19 – Resultados experimentais – P_b em presença de cortes.....	3
Figura 3.20 – Esquemas de inserção e extracção de marcas-de-água no domínio da frequência: a) Inserção; b) Extracção.....	3

Figura 3.21 – Coeficientes DCT utilizados para inserção da marca.	3
Figura 3.22 – Coeficientes DCT para os quais foi realizado o teste χ^2	3
Figura 3.23 – Distribuição estatística de alguns coeficientes DCT – imagem <i>Lena</i>	3
Figura 3.24 – Estrutura do desmodulador para inserção no domínio da frequência.	3
Figura 3.25 – Resultados experimentais para $M=2$, com e sem contabilização de $B[m]$, considerando que os coeficientes DCT seguem uma distribuição de Laplace ($c[m]=1$).	3
Figura 3.26 – Resultados teóricos e experimentais – P_b vs. N° de pontos / Bit de informação.....	3
Figura 3.27 – Resultados experimentais – P_b em presença de compressão JPEG.	3
Figura 3.28 – PSNR resultante após inserção da marca – imagem <i>Lena</i>	3
Figura 3.29 – Comparação de resultados nos dois domínios de inserção – imagem <i>Lena</i>	3
Figura 3.30 – PSNR resultante após inserção da marca – imagem <i>02</i>	3
Figura 3.31 – Comparação de resultados nos dois domínios – imagem <i>02</i>	3
Figura 3.32 – PSNR resultante após inserção da marca e compressão JPEG.	3
Figura 3.33 – Comparação de resultados nos dois domínios quando em presença de compressão.3	
Figura 4.1 – Taxonomia das técnicas de codificação de canal usuais em telecomunicações.....	3
Figura 4.2 – Estrutura de palavra de um código de bloco linear.	3
Figura 4.3 – Codificador convolucional simples.....	3
Figura 4.4 – <i>Árvore</i> do código convolucional referente à figura 4.3.....	3
Figura 4.5 – Esquema geral de inserção da marca-de-água (com codificador).	3
Figura 4.6 – Esquema geral de extracção da marca-de-água (com decodificador).	3
Figura 4.7 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos BCH (Imagem <i>Lena</i>).	3
Figura 4.8 – Resultados teóricos e experimentais de P_b vs. N° de pixels / Bit de informação para o código BCH(127,64).	3
Figura 4.9 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos RS utilizando 16 níveis de sinalização (Imagem <i>Lena</i>).	3
Figura 4.10 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos RS utilizando 256 níveis de sinalização (<i>Lena</i>).	3
Figura 4.11 – Resultados experimentais de P_b vs. N° de pixels / Bit de informação para o código RS(14,8) utilizando 16 e 256 níveis de sinalização.....	3
Figura 4.12 – Resultados teóricos e experimentais de P_b vs. N° de pixels / Bit de informação para o código convolucional com $c_r=1/2$ com decisões dura e suave.	3

Figura 4.13 – Resultados experimentais – P_b vs. N° de Pixels / Bit de informação.....	3
Figura 4.14 – Resultados experimentais em presença de compressão JPEG.....	3
Figura 4.16 – Resultados experimentais em presença de cortes.	3
Figura 4.17 – Resultados teóricos e experimentais (domínio da frequência).	3
Figura 4.18 – Resultados experimentais em presença de compressão JPEG (domínio da frequência).	3
Figura 5.1 – Esquema de combinação de sinal.	3
Figura 5.2 – Esquema de extracção da marca com saídas utilizadas na combinação de sinal.	3
Figura 5.3 – Exemplo de aplicação de lógica de maioria.....	3
Figura 5.4 – Diagrama de blocos do sistema de combinação linear de sinal (para $M=2$).	3
Figura 5.5 – a) Valores de G_1 , G_2 e G_3 para algumas combinações de α, β ; b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=1$	3
Figura 5.6 – a) Valores de G_1 , G_2 e G_3 para algumas combinações de α, β ; b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=1.5$	3
Figura 5.7 – a) Valores de G_1 , G_2 e G_3 , para algumas combinações de α, β ; b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=2$	3
Figura 5.8 – Sequências de vídeo CCIR-601 com 300 tramas cada: a) <i>Stefan</i> ; b) <i>Mobile & Calendar</i> ; c) <i>Table-Tennis</i>	3
Figura 5.9 – Evolução de a e σ em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 2 Mbit/s.	3
Figura 5.10 – Evolução de a e σ em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 4 Mbit/s.	3
Figura 5.11 – Evolução de a e σ em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 6 Mbit/s.	3
Figura 5.12 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 2 Mbit/s, utilizando o método 1.....	3
Figura 5.13 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 4 Mbit/s, utilizando o método 1.....	3
Figura 5.14 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 6 Mbit/s, utilizando o método 1.....	3
Figura 5.15 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 2 Mbit/s, utilizando o método 2.....	3
Figura 5.16 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 4 Mbit/s, utilizando o método 2.....	3
Figura 5.17 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 6 Mbit/s, utilizando o método 2.....	3

Figura 5.18 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 2 Mbit/s, utilizando conhecimento <i>a priori</i> dos símbolos inseridos no cálculo de a_i e σ_i .	3
Figura 5.19 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 4 Mbit/s, utilizando conhecimento <i>a priori</i> dos símbolos inseridos no cálculo de a_i e σ_i .	3
Figura 5.20 – Evolução de c_i em 3 GOPs da sequência <i>Stefan</i> MPEG-2 @ 6 Mbit/s, utilizando conhecimento <i>a priori</i> dos símbolos inseridos no cálculo de a_i e σ_i .	3
Figura A.1 – Limite de <i>Shannon</i> – imagem <i>Lena</i> : a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.	3
Figura A.2 – Limite de <i>Shannon</i> – imagem <i>Mandrill</i> : a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.	3
Figura A.3 – Limite de <i>Shannon</i> – imagem <i>02</i> : a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.	3
Figura B.1 – Ângulo de observação.	3
Figura D.1 – Estrutura de tramas <i>GOP</i> 3, com passo de predição 1.	3
Figura D.2 – Estrutura de tramas <i>GOP</i> 6, com passo de predição 3.	3

Índice de tabelas

Tabela 2.1 – Síntese de requisitos comuns.....	3
Tabela 3.1 – Parâmetros utilizados no modelo perceptual do domínio da frequência.	3
Tabela 3.2 – Resultados do teste chi-quadrado (χ^2).	3
Tabela 4.1 – Códigos de bloco binários BCH analisados.	3
Tabela 4.2 – Códigos de bloco não-binários (RS) analisados.....	3
Tabela 5.3 – Taxa de sucesso na extração com compressão MPEG-2 @ 2Mbit/s – $\beta = 0.2$	3
Tabela 5.4 – Taxa de sucesso na extração com compressão MPEG-2 @ 2Mbit/s – $\beta = 0.3$	3
Tabela 5.5 – Taxa de sucesso na extração com compressão MPEG-2 @ 4Mbit/s – $\beta = 0.2$	3
Tabela 5.6 – Taxa de sucesso na extração com compressão MPEG-2 @ 4Mbit/s – $\beta = 0.3$	3
Tabela 5.7 – Taxa de sucesso na extração com compressão MPEG-2 @ 6Mbit/s – $\beta = 0.2$	3
Tabela 5.8 – Taxa de sucesso na extração com compressão MPEG-2 @ 6Mbit/s – $\beta = 0.3$	3
Tabela 5.9 – Taxa de sucesso na extração com compressão MPEG-2 @ 2 Mbit/s – $\beta = 1.5$	3
Tabela 5.10 – Taxa de sucesso na extração com compressão MPEG-2 @ 4 Mbit/s – $\beta = 1.5$	3
Tabela 5.11 – Taxa de sucesso na extração com compressão MPEG-2 @ 6 Mbit/s – $\beta = 1.5$	3
Tabela B.1 – Frequências espaciais (em ciclos/bloco) correspondentes aos coeficientes DCT.....	3
Tabela B.2 – Frequências espaciais horizontais com $N_H = 720$ (em ciclos/grau).	3
Tabela B.3 – Frequências espaciais verticais com $N_V = 576$ (em ciclos/grau).	3

Capítulo 1

Introdução

Actualmente, as tecnologias digitais encontram-se presentes em diversos aspectos do nosso quotidiano. De facto, nas duas últimas décadas, o desenvolvimento da informática e das telecomunicações digitais deu-se de uma forma extremamente rápida. Muitos frutos desse desenvolvimento foram (e continuam a ser) introduzidos na sociedade, influenciando o modo como vivemos, agimos e nos relacionamos, e conduzindo a um novo conceito – o da *sociedade da informação*.

Como expoente máximo deste desenvolvimento, registe-se o aparecimento e a rápida evolução da *Internet*, das comunicações móveis e da televisão digital, em paralelo com novos suportes para gravação digital¹, possibilitando a troca de enormes quantidades de informação, a uma escala nunca antes vista. Esta informação, para além dos tradicionais sinal de voz ou dados alfanuméricos, poderá também consistir em imagens fixas, som e vídeo, designados no seu conjunto como *informação multimédia*.

Embora se reconheçam inúmeras vantagens na representação da informação em formato digital e no recurso aos meios de transmissão digital, torna-se também mais difícil o desenvolvimento

¹ Com destaque para os discos ópticos, como o CD (*Compact Disk*) e o DVD (*Digital Versatile Disk*).

de mecanismos eficazes para controlo do uso da informação. Com efeito, se os acessos à informação e a própria informação são por vezes protegidos através de técnicas criptográficas, uma vez concedido o acesso e descriptada a informação, esta fica vulnerável a inúmeras formas de manuseamento, muitas vezes fora do âmbito legal. Embora este problema se tenha também colocado no “mundo analógico”, é de resolução mais delicada para a informação em formato digital, facilmente manipulável e possibilitando a produção de cópias indistintas do original, com consequentes prejuízos para os detentores de direitos (e.g., autores, editoras).

Como exemplo, refira-se o famoso caso *Napster* que, permitindo a livre troca de música em formato MP3² entre cibernautas de todo o mundo, se viu alvo de um processo judicial por não serem respeitados os direitos de autor. Outros sistemas, semelhantes ao *Napster*, surgiram entretanto e para os quais a troca de informação não se limita ao som, extendendo-se também ao vídeo³. Este panorama, motivou o aparecimento de várias iniciativas a nível internacional, com o intuito de desenvolver novos mecanismos de protecção e que salvaguardem os direitos de autores e distribuidores legais.

No conjunto das técnicas possíveis para protecção dos direitos de autor, têm merecido particular atenção as baseadas em *marcas-de-água digitais*⁴. O objectivo destas técnicas é inserir de forma invisível (i.e., perceptualmente e estatisticamente indetectável) e eventualmente robusta, informação adicional – *marca-de-água* – sobre os dados multimédia que se pretendem proteger. Esta informação, apenas acessível a entidades autorizadas, poderá consistir na identificação do autor ou detentor do produto original, e/ou permitir a verificação da integridade do produto em causa (i.e., verificar se o produto foi sujeito a manipulações ilegais). Por *robusta* entende-se a possibilidade de a marca permanecer no produto, mesmo quando este é sujeito a manipulações que de alguma forma o alterem. Não existem, nem provavelmente existirão, técnicas que sejam robustas a todas as manipulações possíveis. No entanto, é possível o desenvolvimento de técnicas que sejam robustas a manipulações que não degradem demasiado a qualidade do produto em causa, o que em princípio será suficiente, já que um produto demasiado degradado perde o seu valor comercial.

² MP3 (MPEG 1 - Layer 3): Sistema de compressão usado para informação áudio no formato digital. Devido à sua elevada taxa de compressão, permite por exemplo 120 faixas musicais num CD e transferências via *Internet* mais rápidas, tornando-se por esses motivos um formato habitual para informação áudio. Os elevados factores de compressão são obtidos à custa de uma redução pouco significativa da qualidade.

³ Actualmente é possível obter na *Internet* um filme com qualidade razoável, sem que este tenha chegado aos ecrãs de cinema.

⁴ Na literatura Inglesa, *digital watermarking*.

A relevância deste tema motivou, ao longo dos últimos anos, um aumento significativo da investigação nesta área, e a consequente criação de diversas empresas que desenvolvem aplicações comerciais com recurso a técnicas de marcas-de-água, de que são exemplos a norte-americana *Digimark Corp.*, e as europeias *Digital Copyright Technologies* e *Alpha Tec. Ltd.*. De realçar o consórcio *VWM Group*⁵, recentemente criado por empresas líderes na electrónica de consumo, bem como os projectos científicos europeus *TALISMAN*⁶ e *MIRADOR*⁷ – já concluídos e pioneiros na investigação nesta área – e *CERTIMARK*, actualmente em curso e cujo objectivo é a definição de um conjunto de testes de referência (*benchmarks*) para a avaliação do desempenho dos algoritmos de marca-de-água.

De referir também, a especificação de normas internacionais relacionadas com a identificação e protecção dos direitos de autor sobre produtos em formato digital. Entre estas, destaque para a norma ISO (*International Organization for Standardization*) que irá definir um identificador para conteúdos áudio-visuais constituído por oito caracteres ASCII (i.e., 64 bits), designado por *número audiovisual internacional normalizado* (ISAN)⁸. A especificação final desta norma deverá estar concluída no corrente ano (2002). A tecnologia das marcas-de-água é considerada a forma ideal de associar o código ISAN ao produto respectivo. A questão da protecção de direitos de autor tem também merecido particular atenção no grupo MPEG-4, onde se prevê a definição de uma *interface* genérica para um conjunto de ferramentas de gestão de direitos de autor e onde será possível a aplicação de diversas técnicas de protecção, entre as quais se incluem as marcas-de-água.

Não é certo que todas as questões relacionadas com a protecção dos direitos de autor sejam, ou venham alguma vez a ser, totalmente resolvidas com o recurso exclusivo à tecnologia das marcas-de-água. No entanto, esta poderá certamente dar um contributo importante na resolução deste tipo de problemas, quer como elemento dissuasivo, quer contribuindo para a identificação de produtos multimédia que foram sujeitos a cópia ou manipulações ilegais. De salientar que, para além da protecção dos direitos de autor, as marcas-de-água podem ainda ser utilizadas com outras finalidades, já que os dados inseridos podem conter informação diversa.

⁵ Este consórcio, formado em Abril de 2001 sob a designação de *Video Watermarking Group (VWM Group)* resulta da fusão dos grupos *Galaxy* (Hitachi, NEC, Pioneer e Sony) e *Millennium* (Digimark, Macrovision e Philips), tendo como objectivo o desenvolvimento de um sistema para protecção de vídeo digital, baseado na tecnologia das marcas-de-água.

⁶ O principal objectivo do projecto *ACTS-TALISMAN* (1995-98) foi providenciar meios de protecção de direitos de autor sobre produtos multimédia. Entre estes meios, encontram-se as marcas-de-água.

⁷ O projecto *ACTS-MIRADOR* (1998-99) pode ser considerado uma extensão do projecto *TALISMAN*, orientado para a aplicação de técnicas de marcas-de-água a vídeo comprimido segundo a norma MPEG-4.

⁸ ISAN – *International Standard Audiovisual Number*.

Como se tornará evidente ao longo deste texto, a utilização da tecnologia das marcas-de-água sobre um produto multimédia pode ser considerada como um processo de comunicação, no qual se pretende enviar uma mensagem – a marca-de-água – através de um canal ruidoso – o produto a proteger. Nesta medida, conceitos e resultados da teoria da comunicação digital, podem ser naturalmente estendidos aos sistemas de marca-de-água, com o intuito de melhorar o desempenho destes sistemas. Avaliar o impacto que a utilização de diversas formas de *modulação, codificação de canal e combinação de sinal*, pode ter no desempenho dos algoritmos de marca-de-água baseados em espalhamento de espectro, constituiu o principal objectivo desta tese. Embora esta nova tecnologia possa ser utilizada em vários tipos de produtos multimédia, o estudo aqui apresentado restringiu-se a imagens fixas e vídeo. O trabalho desenvolvido conduziu a uma estruturação deste relatório em seis capítulos.

Após o capítulo de introdução (capítulo 1) apresentam-se, no capítulo 2, os princípios de funcionamento da tecnologia das marcas-de-água, bem como alguns conceitos e terminologia utilizados nesta área. De forma a ser feito o enquadramento histórico das marcas-de-água, é descrita a sua evolução desde as suas primeiras utilizações em papel, no século XIII, até à sua aplicação ao domínio digital, na época actual. Descrevem-se também as diferentes áreas de aplicação (para além da protecção dos direitos de autor) que podem beneficiar com a utilização desta nova tecnologia e sintetizam-se os requisitos que estas aplicações impõem aos algoritmos de marcação.

No capítulo 3, apresenta-se um estudo relativo à aplicação de modulação multinível, baseada em espalhamento de espectro, à assinatura digital⁹ de imagens fixas. Após uma breve introdução ao espalhamento de espectro, apresentam-se formas de aplicar esta modulação na assinatura de imagens e descrevem-se os esquemas gerais para inserção e extracção da marca-de-água utilizados ao longo desta tese. Prossegue-se com um estudo analítico e experimental do desempenho da modulação multinível, seguindo duas abordagens distintas: inserção / extracção da marca no espaço da imagem (ou domínio espacial) e inserção / extracção no domínio da frequência. A proposta de um esquema de modulação / desmodulação multinível para a inserção / extracção da marca-de-água no domínio da frequência, constitui a principal contribuição deste capítulo.

No capítulo 4 avalia-se o desempenho de alguns códigos de correcção de erro, usuais em sistemas de telecomunicações, quando associados às técnicas de assinatura de imagens

⁹ O termo “assinatura digital” é muitas vezes utilizado para designar o processo de autenticação do emissor de uma mensagem ou do conteúdo da mesma, com recurso a técnicas criptográficas. Ao longo deste texto, esse termo será utilizado para referir a inserção de informação adicional – marca-de-água – sobre imagens ou vídeo, independentemente da finalidade dessa informação.

analisadas no capítulo 3. O texto deste capítulo inicia-se com uma breve taxonomia dos códigos correctores de erro mais utilizados em sistemas de telecomunicações. Prossegue-se com o estudo analítico e experimental da melhoria do desempenho na extracção da marca-de-água, resultante da utilização das várias técnicas de correcção de erro estudadas. Por fim, comparam-se os resultados obtidos com os diferentes códigos, extendendo-se essa comparação aos dois domínios de inserção da marca – espacial e frequência. A avaliação das várias soluções para correcção de erros tendo por base o espalhamento de espectro [6], incluindo a associação de códigos correctores não binários com a modulação multinível [7], constituem as principais contribuições deste capítulo.

A tecnologia das marcas-de-água em vídeo pode ser analisada como um sistema multi-canal, se em cada trama for inserida a mesma marca e se cada trama puder ser considerada como um canal independente. Nesta situação, a detecção da marca-de-água pode ser melhorada considerando simultaneamente um grupo de tramas consecutivas, aplicando técnicas de combinação de sinal, amplamente utilizadas no domínio das comunicações rádio com diversidade. No capítulo 5, efectua-se um estudo analítico e avalia-se experimentalmente o desempenho de duas alternativas de combinação de sinal – *lógica de maioria* e *combinação linear* – quando usadas em conjunto com o sistema de marca-de-água analisado nos capítulos anteriores [4,5,6]. No caso da estratégia *lógica de maioria*, determina-se qual o ponto, no sistema de recepção, em que deverá ter lugar a combinação. Para a *combinação linear*, a análise teórica é efectuada no sentido de obter os pesos que maximizam a relação sinal-ruído do sinal combinado – *pesos óptimos* – e de avaliar o efeito da utilização de pesos não óptimos, de que é exemplo a utilização de *pesos constantes e unitários*. O estudo analítico é complementado com uma avaliação experimental dos dois métodos, utilizando três sequências de vídeo CCIR-601, codificadas em MPEG-2 a 2, 4 e 6 Mbit/s.

Para finalizar, no capítulo 6 apresentam-se as principais conclusões sobre o trabalho desenvolvido e sugerem-se alguns tópicos de pesquisa para trabalho futuro.

Capítulo 2

Assinatura de imagens com marcas-de-água

2.1 Introdução

Ao longo dos últimos anos verificou-se um aumento significativo da investigação na área das marcas-de-água, existindo actualmente diversas aplicações comerciais. Embora o objectivo inicial e propulsor desta nova tecnologia tenha sido assegurar a protecção dos direitos de autor em meio digital, não é certo que todos os problemas relacionados com o “controlo do uso da informação” sejam, ou venham alguma vez a ser, totalmente resolvidos com o recurso às marcas-de-água. No entanto, as marcas-de-água podem ser utilizadas com outras finalidades para além da protecção dos direitos de autor, já que podem conter informação diversa. Esta possibilidade, juntamente com a constatação das limitações das marcas-de-água como forma única de assegurar a protecção da informação, motivou a investigação de novas áreas de aplicação e conduziu a novos e importantes desenvolvimentos na assinatura digital de imagens.

Neste capítulo, que se encontra estruturado em seis secções, faz-se uma introdução à assinatura digital de imagens com recurso às técnicas de marca-de-água, descrevendo-se a terminologia e os princípios utilizados nesta área. Após a secção introdutória apresenta-se, na secção 2.2, uma breve resenha histórica das marcas-de-água, desde as suas primeiras utilizações em papel, até à sua utilização mais recente, na informação multimédia. Na secção 2.3 descreve-se o princípio de

funcionamento desta tecnologia e alguns conceitos relacionados. Na secção 2.4, apresentam-se áreas de aplicação que podem beneficiar da utilização das marcas-de-água e sintetizando-se os requisitos exigidos pelas aplicações descritas. Termina-se, na secção 2.5, com uma breve referência ao problema da ambiguidade criada pela existência de múltiplas marcas num mesmo produto e às formas possíveis de minimizar o seu efeito.

2.2 Breve historial das marcas-de-água

As primeiras utilizações das marcas-de-água remontam aos finais do séc. XIII, em simultâneo com o aparecimento da manufacturação do papel. O primeiro registo de um papel marcado com esta técnica data de 1292 e tem a sua origem na localidade de *Fabriano*, em Itália [17]. No final do séc. XIII existiam, em *Fabriano*, cerca de 40 oficinas de fabrico de papel. Cada oficina produzia papel com características específicas (formato, qualidade e preço), mas ainda demasiado rugoso para poder ser utilizado na escrita. Este papel era depois tratado por artesãos, que o poliam e dividiam em folhas. Para completar o circuito, o produto final era vendido aos retalhistas, que por sua vez o vendiam ao público. A concorrência em cada uma das etapas do percurso do papel, desde o seu estado em bruto até à venda ao público era grande e, por passar por tantas mãos, tornava-se difícil saber ao certo a sua proveniência. As marcas-de-água surgiram como um meio de garantir a origem e autenticidade do papel.

Depressa apareceram tentativas de imitação, falsificação e mesmo anulação das marcas-de-água. No entanto, ainda hoje o papel marcado com esta técnica é usado em documentos de grande importância, como bilhetes de identidade, notas ou papel timbrado¹⁰. As marcas-de-água continuam a ser usadas na indústria do papel com o objectivo de datar, autenticar ou determinar a sua origem, fornecendo desta forma uma segurança que evitou, durante os últimos 700 anos, a fácil falsificação e duplicação de documentos. Deste modo, pode-se afirmar que as marcas-de-água constituem ainda nos dias de hoje, uma forma muito simples, mas efectiva, para protecção contra cópias de documentos impressos.

A crescente utilização de informação em formato digital e a fácil produção de cópias desta informação, indistintas da original, está a provocar uma situação semelhante à ocorrida com o papel nos séculos XIII / XIV. Com base na ideia originada em *Fabriano* surgiram, em 1993, as primeiras propostas para inserção de marcas-de-água em conteúdos digitais [43]. Desde então, o interesse por este assunto tem crescido de forma assinalável, sendo objecto de intensa investigação, tanto por entidades académicas, como por empresas comerciais.

¹⁰ De referir que estes documentos são usados frequentemente em tribunal.

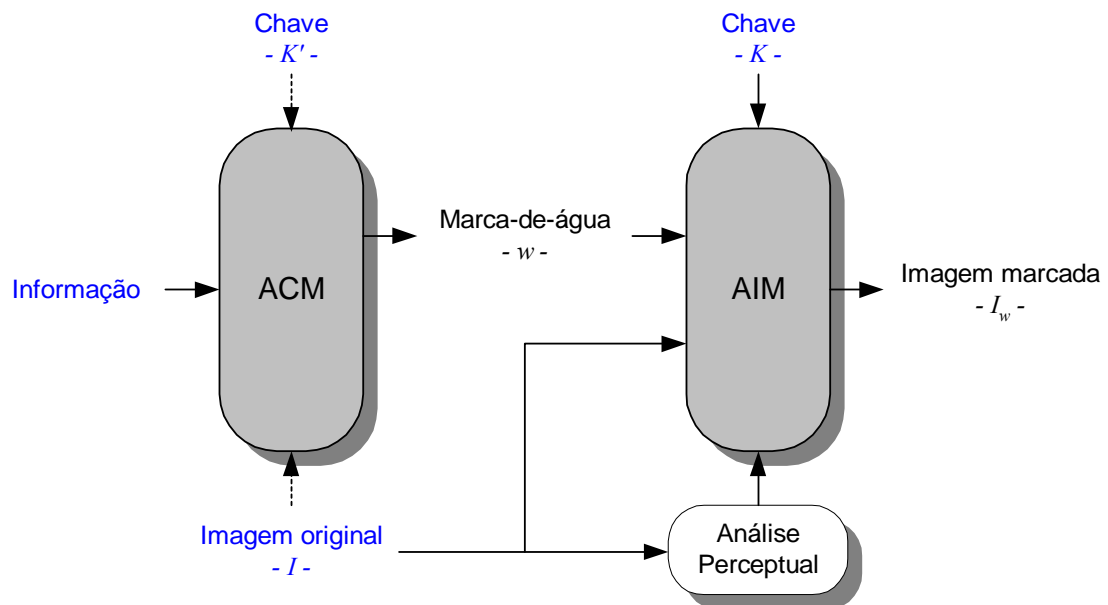


Figura 2.1 – Assinatura digital de uma imagem.

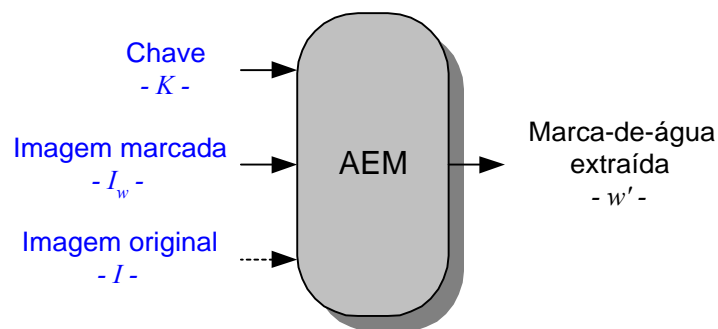


Figura 2.2 – Extração de uma marca-de-água digital.

2.3 Princípios de funcionamento e conceitos relacionados

A investigação científica em técnicas de marcas-de-água, motivou o desenvolvimento de conceitos específicos para esta área e a extensão a esta área de conceitos já existentes, que serão sintetizados nesta secção.

2.3.1 Esquema geral de um sistema de assinatura de imagens

Num esquema geral de assinatura digital de imagens (figuras 2.1, 2.2 e 2.3) distinguem-se três algoritmos [27]:

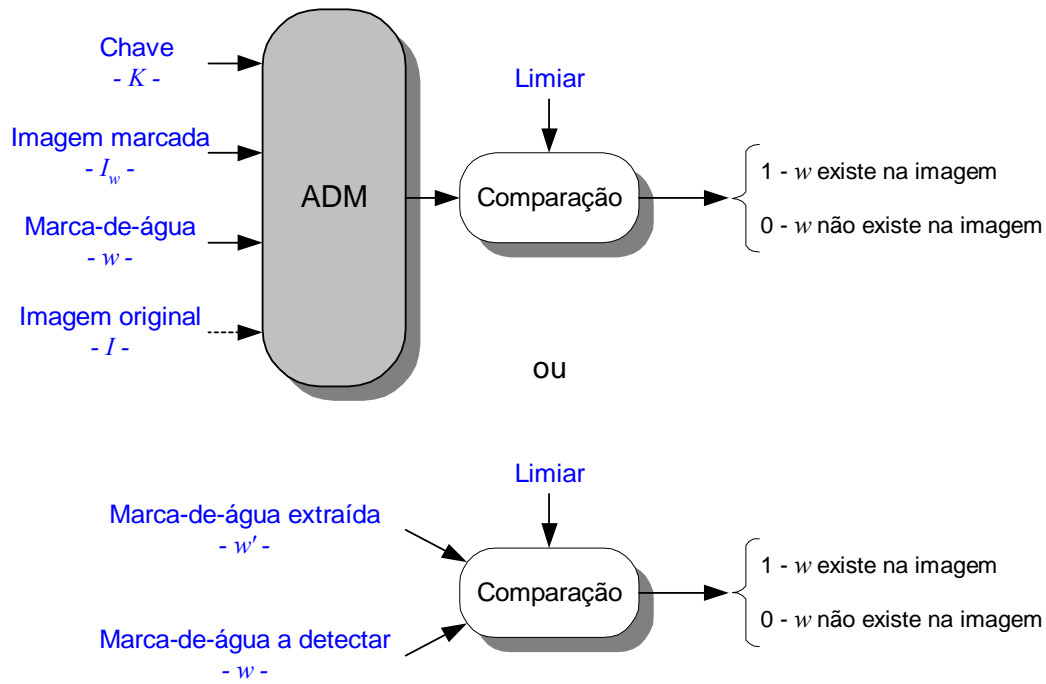


Figura 2.3 – Detecção de uma marca-de-água digital.

- *Algoritmo de criação da marca (ACM)* – produz uma sequência binária w (marca-de-água), com base na informação a inserir na imagem e, caso se pretenda proteger a informação contida na marca-de-água, com base numa chave secreta K' . Eventualmente, a marca-de-água poderá também ser dependente de algumas características da imagem a marcar – I ;
- *Algoritmo de inserção da marca (AIM)* – insere, sobre a imagem original I , a sequência binária w criada pelo ACM, obtendo-se a imagem marcada I_w . Para evitar a visibilidade da marca-de-água, este algoritmo deverá ter em consideração as características do sistema visual humano. As posições de imagem onde é realizada a inserção de um determinado bit da marca são usualmente determinadas por uma chave secreta K ;
- *Algoritmo de extracção/detecção da marca (AEM/ADM)* – extrai (AEM) a marca-de-água da imagem marcada e, eventualmente manipulada – I'_w – ou avalia (ADM) se uma determinada marca-de-água está presente nessa imagem. Ambos os processos requerem o conhecimento da chave K utilizada na inserção, sendo também necessário o conhecimento da marca se se pretender efectuar a detecção. Dependendo do algoritmo de extracção/detecção pode ou não ser requerida a imagem original. No processo de detecção são possíveis duas abordagens distintas: extracção da marca e comparação com a marca w a detectar; correlação entre a marca w e a imagem assinada I'_w seguida de uma comparação do resultado com um valor de limiar.

2.3.2 Classes de marcas-de-água

As técnicas de marca-de-água podem ser classificadas segundo diferentes critérios, nomeadamente: o domínio em que é feita a inserção e extracção/detecção da marca; a robustez da marca face a manipulações da imagem; o tipo de aplicações a que se destinam (o que condiciona o método de extracção/detecção).

Quanto ao domínio em que é feita a inserção, considera-se usualmente:

- *Domínio espacial* – se a marca for inserida directamente no espaço da imagem, por alteração de valores das componentes da cor;
- *Domínio da frequência* – se a marca for inserida através da alteração de coeficientes espectrais resultantes de transformações em frequência. Entre as transformações mais utilizadas encontram-se a transformada DCT (Transformada de Coseno Discreta), a transformada *Wavelet*, ou a transformada DFT (Transformada de *Fourier* Discreta).

Utilizando como critério a robustez, as marcas-de-água podem ser classificadas como:

- *Robustas* – se persistirem na imagem após esta ser sujeita a manipulações (intencionais ou não) que não reduzam o valor comercial da imagem;
- *Frágeis* – se forem destruídas qualquer que seja a alteração da imagem marcada. Destinam-se normalmente a aplicações cujo objectivo é a avaliação da integridade estrita (i.e., ao nível do pixel) da imagem;
- *Semi-frágeis* – se persistirem na imagem apenas quando esta é sujeita a modificações autorizadas (e.g., compressão). Destinam-se normalmente à avaliação da integridade do conteúdo semântico da imagem, permitindo detectar manipulações indesejáveis, tais como inserção, remoção ou modificação de objectos visuais.

Relativamente à forma como é realizada a extracção ou detecção da marca-de-água, podem-se agrupar as várias técnicas em quatro classes [10,15,26]:

- *Extracção / detecção privada* [9,33] – se for necessária a presença do original para realizar a extracção (sistema *não-cego*¹¹). A extracção pode ser descrita como $w' = f_e(I'_w, I, K)$; a detecção é vista como uma função $f_d(I'_w, I, w, K)$, sendo o seu resultado uma decisão binária, i.e., “ w presente” ou “ w não presente”. As marcas-de-água *privadas* são as mais robustas uma vez que utilizam o original na extracção;
- *Detecção semi-privada* [13,40] – se não for requerida a presença do original para realizar a extracção da marca-de-água (sistema *cego*¹²). A detecção é feita em função de I'_w , w e K ;
- *Extracção / detecção semi-pública* [3,13,37] – se a chave necessária para a extracção depender da própria imagem, i.e., $K = f_k(I)$. À semelhança do caso anterior, o sistema é também *cego*. A extracção será uma função de I'_w e $K = f_k(I)$, e a detecção uma função de I'_w , w e $K = f_k(I)$;
- *Extracção pública* [13,40] – se a extracção for realizada apenas com base na imagem recebida I'_w e na chave secreta K (extracção *cega*). Esta é a classe de técnicas de marcas-de-água com maior número de aplicações.

2.3.3 Qualidade visual

Nos sistemas de marca-de-água são usualmente referidos dois tipos de qualidade visual: a qualidade visual resultante após inserção da marca-de-água – QVM – e a qualidade visual resultante após manipulações da imagem marcada que provoquem a destruição da marca-de-água – QVD.

A qualidade da imagem após inserção da marca-de-água deve ser tão boa quanto a da imagem original, i.e., a marca deve ser imperceptível. Para que este requisito seja alcançado, o método de inserção deve ter em consideração as características do sistema visual humano.

Por outro lado, na maioria das aplicações pretende-se que a degradação que é necessária causar à imagem, de forma a destruir a marca, seja o mais elevada possível, conduzindo a uma qualidade visual – QVD – demasiado baixa e, consequentemente, a um valor comercial nulo. As marcas-de-água frágeis usadas na *autenticação* e *avaliação da integridade*, aplicações onde é desejável a destruição da marca qualquer que seja a manipulação, constituem uma excepção a esta regra.

¹¹ Na literatura Inglesa, *non-blind watermark*.

¹² Na literatura Inglesa, *blind* ou *oblivious watermark*.

2.3.4 Fiabilidade na extracção

Em aplicações que requerem a extracção da marca-de-água, são tipicamente utilizados dois indicadores do desempenho do sistema:

- *Probabilidade de erro da marca (P_e)* – probabilidade de ser extraída uma marca-de-água diferente da que foi previamente inserida;
- *Probabilidade de erro de bit (P_b)* – probabilidade de erro de bit na marca-de-água extraída.

Embora idealmente ambas as probabilidades devam ser nulas, na prática não o são devido essencialmente a dois factores: erros que ocorrem em sistemas de marcas-de-água cegos, provocados pela interferência da imagem sobre a marca; erros resultantes da manipulação (intencional ou não) da imagem marcada, manipulação essa que provoca a degradação da marca inserida.

2.3.5 Fiabilidade na detecção

Como referido atrás, a detecção de uma dada marca-de-água consiste em determinar se essa marca se encontra ou não presente no produto marcado. Esta decisão pode ser vista como um teste de hipóteses:

H_0 : A imagem I_w não contém a marca w , inserida com a chave K ;

H_1 : A imagem I_w contém a marca w , inserida com a chave K .

Sendo a decisão \mathbf{D} limitada ao conjunto $\{H_0, H_1\}$, definem-se as seguintes probabilidades como indicadores do desempenho do sistema:

- *Probabilidade de detecção (P_d)* – probabilidade da decisão ser feita a favor da hipótese H_1 , quando de facto foi inserida a marca w , com chave K , i.e., $P_d = P(\mathbf{D}=H_1 | H_1)$;
- *Probabilidade de falso alarme (P_F)* – probabilidade de ser feita uma decisão em favor de H_1 quando na realidade não foi inserida a marca w , com chave K , i.e., $P_F = P(\mathbf{D}=H_1 | H_0)$.

Numa sistema ideal ter-se-ia $P_d=1$ e $P_F=0$. Na prática, estes valores ideais não são atingidos, pelas razões apontadas anteriormente.

2.3.6 Payload

No contexto das marcas-de-água, designa-se por *payload* o comprimento (em bits) da marca-de-água, sem ser contabilizada eventual informação redundante introduzida pela utilização de códigos correctores de erro ou de qualquer outra forma de codificação.

O valor do *payload* varia consoante o tipo de aplicação a que se destinam as marcas-de-água, podendo ser apenas um bit (aplicações em que apenas é feita detecção por correlação), 8 bits (controlo de cópias em DVD), 64 bits (código *ISAN*) ou um número elevado de bits, e.g., 256 bits ou mais (comunicação secreta).

2.3.7 Complexidade

Por *complexidade* entende-se o número de operações aritméticas necessárias para executar um determinado algoritmo e a natureza destas operações: operações inteiras, com vírgula fixa ou vírgula flutuante. Em determinadas aplicações, é necessária a inserção, extracção e detecção da marca-de-água em tempo real, pelo que a complexidade do algoritmo assume um papel extremamente importante. De referir, no entanto, que maior complexidade nem sempre significa maior tempo de processamento do algoritmo, pois este pode ser reduzido com a utilização de técnicas de *pipelining* ou paralelização. No entanto, a utilização deste tipo de técnicas implica a utilização de mais recursos (e.g., processadores, memórias) o que pode conduzir a um aumento de custos.

2.3.8 Formas de ataque e sua caracterização

No contexto das marcas-de-água, designa-se por *ataque* a manipulação de uma imagem marcada, modificando de alguma forma o seu conteúdo. Um ataque que vise a destruição da marca-de-água é designado por *ataque intencional*, enquanto que um ataque resultante da utilização de técnicas comuns de processamento de imagem, sem a finalidade de destruir a marca-de-água (e.g., compressão), é designado por ataque *não-intencional*. Tendo em conta as suas características fundamentais, os ataques podem ser classificados em cinco categorias distintas (A1 a A5) [10,15]:

- A1 – Manipulações na imagem ou vídeo, frequentes na sua distribuição:
 - Compressão segundo as normas JPEG e MPEG;
 - Conversões digital/analógico e analógico/digital;
 - Impressão e digitalização (em imagem fixa);

- Conversões de formatos em vídeo (alteração da resolução espacial ou da frequência de imagem, passagem de formato progressivo a entrelaçado e vice-versa);
 - Marcação múltipla do mesmo produto;
 - Erros devido a perdas de pacotes em transmissões via *Internet*.
- A2 – Ataques resultantes do processamento de imagem, usualmente não-intencionais:
 - Filtragem passa-baixo e de mediana;
 - Adição de ruído;
 - Melhoramento dos contornos;
 - Correção do factor gama e modificação do histograma (equalização).
- A3 – Ataques devido a transformações geométricas simples, intencionais ou não:
 - Translação;
 - Mudança de escala (*zoom*);
 - Cortes (*cropping*).
- A4 – Ataques com transformações geométricas mais complexas, usualmente intencionais:
 - Rotação;
 - Transformações geométricas generalizadas;
 - Transformações geométricas locais, aleatórias.
- A5 – Ataques estatísticos, intencionais:
 - Métodos estatísticos que permitem estimar e remover a marca-de-água;
 - Geração de uma nova imagem realizando a média sobre várias cópias da mesma imagem, mas com marcas-de-água diferentes.

2.3.9 Segurança

Uma técnica de marca-de-água só é verdadeiramente segura se o conhecimento do algoritmo de inserção não permitir, a uma entidade não autorizada, detectar a presença da marca e conseguir proceder à sua extracção. Para que isso possa ser alcançado, é necessário fazer depender de uma chave secreta a inserção e extracção da marca-de-água. Adicionalmente, o conteúdo da marca-de-água poderá ser protegido através de técnicas criptográficas. Para que a segurança não seja facilmente quebrada, é conveniente que o universo das chaves seja grande, desencorajando-se assim a busca exaustiva das chaves por parte de entidades fraudulentas.

2.4 Áreas de aplicação das marcas-de-água

O ano de 1993 pode ser considerado como o início da investigação na área de assinatura digital de imagens, embora algumas publicações anteriores a essa data [41,42] tivessem já introduzido a ideia de inserir informação em imagens, de forma a assegurar os direitos de autor. Desde então, os algoritmos de marca-de-água ganharam especial relevo, tanto por parte da comunidade académica, como de entidades privadas, sendo este interesse crescente acompanhado por uma rápida e enorme evolução no desempenho das técnicas desenvolvidas.

No que respeita à informação visual, a grande motivação para o desenvolvimento de algoritmos foi, numa fase inicial, a protecção dos direitos de autor sobre imagens fixas, como consequência do grande número e variedade de imagens com acesso livre, obtidas através da *Internet*. No entanto, novas áreas de aplicação têm sido sugeridas em publicações recentes e que se descrevem sinteticamente nesta secção.

2.4.1 Prova do proprietário

Neste tipo de aplicação, a marca-de-água é utilizada com o intuito de identificar o detentor de direitos do produto assinado. Pretende-se, com esta aplicação, substituir o processo tradicional de protecção/gestão de direitos de autores efectuado pelas Sociedades de Autores. Existem dois modelos de gestão de direitos de autor:

- *Gestão centralizada* – se cada proprietário (autor ou detentor dos direitos) registar o seu trabalho numa sociedade de protecção de direitos de autor. Essa sociedade fica então incumbida de zelar pelos interesses do proprietário, fornecendo as provas de propriedade sempre que surjam questões relacionados com os direitos de autor;
- *Gestão distribuída* – se a gestão dos direitos de autor for assegurada pelo próprio detentor de direitos. É neste contexto que as marcas-de-água poderão contribuir de forma significativa para a protecção dos direitos. Este modelo de gestão, mais liberal, poderá adequar-se a cenários para os quais não seja indicado (ou seja impossível) registar o produto numa sociedade de autores.

A utilização de marcas-de-água para protecção de direitos de autor, num sistema de gestão distribuída, tem como requisito principal uma elevada robustez das marcas-de-água a ataques que visem a sua remoção e que não diminuam o valor comercial do produto. Uma vez que nenhuma técnica de marca-de-água proposta até hoje satisfaz inteiramente este requisito, o

recurso a estas técnicas não poderá constituir uma prova suficiente numa decisão judicial. Esta limitação conduziu a uma reformulação no uso das marcas-de-água para prova do proprietário:

- As marcas-de-água (robustas) podem contribuir para a prova de propriedade quando combinadas com outros elementos, constituindo um importante auxílio na resolução de questões judiciais;
- As marcas-de-água podem fornecer elementos adicionais durante um processo de investigação, nomeadamente contribuindo para a identificação de imagens que foram sujeitas a manipulações e quais os tipos de manipulações a que foram sujeitas (*avaliação da integridade*).

2.4.2 Identificação

A identificação universal e unívoca de um produto multimédia com recurso a marcas-de-água possibilitará a gestão automática da comercialização dos produtos e dos direitos de autor correspondentes. Para tal, é fundamental a definição de identificadores normalizados para imagem fixa e vídeo. Várias normas ISO foram (ou estão actualmente a ser) definidas com vista à identificação de produtos multimédia. Entre os identificadores a serem normalizados, destaca-se o *número audiovisual internacional normalizado* – ISAN – cujo conteúdo consiste numa sequência de 8 caracteres ASCII (i.e., 64 bits) e que possibilitará a indexação de produtos áudio-visuais a uma base de dados universal onde estará contida informação relativa a esse produto (descrição do produto; definição das entidades com direito legal para o comercializar; descrição das entidades detentoras dos direitos de autor, etc.).

2.4.3 Autenticação

As marcas-de-água para autenticação destinam-se a garantir, ao consumidor, a proveniência e genuinidade do produto. Para tal, é necessário que os receptores, através da marca-de-água extraída, identifiquem univocamente a origem do material. Existem duas sub-classes dentro deste tipo de aplicações:

- Autenticação tolerante a certo tipo de manipulações de imagem, inerentes à distribuição e difusão da imagem (e.g., compressão). Neste caso, a solução para autenticação passa pela inserção de uma marca-de-água semi-frágil, identificando a origem do produto;

- Autenticação intolerante a qualquer forma de manipulação do produto. Este tipo de aplicação sugere a utilização de marcas-de-água frágeis.

2.4.4 Monitorização da difusão/distribuição

Este grupo de aplicações surge fundamentalmente no contexto de distribuição em larga escala de produtos televisivos. Existem essencialmente duas classes de aplicações de marcas-de-água neste campo, que visam:

- Medir o impacto do produto no público consumidor (medidas de audiência);
- Detectar a utilização ilegal dos produtos televisivos.

Ambos os casos requerem técnicas de marcas-de-água com características semelhantes, sendo a extracção ou detecção das marcas-de-água realizada em equipamento receptor da difusão.

No primeiro caso, após a extracção da marca-de-água é enviada uma mensagem para uma central responsável pelo tratamento dos dados extraídos, através de um canal de retorno. Torna-se deste modo possível identificar a estação difusora, o programa, o tempo de visualização do programa e a localização geográfica dos consumidores, dados que poderão ser utilizados para realizar estatísticas referentes às audiências, de uma forma automática e eficiente.

O objectivo da segunda classe de aplicações é garantir, aos detentores de direitos sobre o material difundido, de que não há abuso relativamente aos contratos estabelecidos com as difusoras. Estas violações aos contratos são geralmente de dois tipos:

- O material televisivo é exibido mais vezes, ou em mais canais do que o estipulado pelo contrato. Este tipo de abuso pode surgir quando o material televisivo consiste, por exemplo, em reportagens de agências noticiosas, material desportivo ou filmes;
- O material televisivo é exibido menos vezes que o acordado no contrato, o que pode suceder quando o produto televisivo consiste em publicidade.

Actualmente, a detecção das situações descritas exige a monitorização exaustiva de todas as emissões realizadas por parte das emissoras com as quais existe contrato, o que é muitas vezes impraticável. A utilização de marcas-de-água poderá garantir esta monitorização. À semelhança do descrito para medições de audiências, as entidades detentoras dos direitos poderão colocar

receptores (localizados nas áreas de interesse) que detectem as marcas-de-água nos produtos negociados e enviem para uma central todos os dados referentes à emissão do produto em causa, podendo as irregularidades serem detectadas de forma automática.

2.4.5 Avaliação da integridade

Actualmente existem no mercado e a baixo custo, ferramentas poderosas de edição e processamento de imagem (“Photoshop”, “Paintshop”, “Corel Draw”), que possibilitam alterar o conteúdo de uma cena com facilidade, movendo, apagando, ou acrescentando objectos visuais, com resultados idênticos aos dos profissionais do meio fotográfico tradicional.

Neste contexto, têm sido investigados algoritmos de assinatura de imagens que permitam avaliar a integridade do conteúdo da imagem e que podem ser divididos em três classes:

- Inserção de uma marca-de-água frágil, se se pretender garantir integridade estrita, i.e., ao nível do pixel da imagem. Neste caso, a informação inserida resulta tipicamente da aplicação de uma função de *hash*¹³ à imagem a proteger;
- Utilização de uma marca-de-água semi-frágil, caso se admitam alterações inerentes à distribuição ou difusão de imagens ou vídeo (e.g., compressão);
- Inserção de uma marca-de-água robusta que contenha um apontador, endereço, ou chave de indexação para o local onde se encontra informação relativa ao conteúdo da imagem. Esta informação deverá ser gerida por uma entidade certificadora e deverá possibilitar a validação do conteúdo da imagem.

2.4.6 Impressões digitais (*fingerprinting*)

No processo de venda de uma cópia de uma imagem ou de um vídeo, poderá ser útil inserir, na cópia transaccionada, a identificação do cliente. Deste modo, sempre que for detectada uma cópia ilegal, torna-se possível identificar qual o cliente que a originou e obter provas para um possível processo judicial. Uma marca-de-água que contenha a identificação do receptor é designada por impressão digital (*fingerprint*).

¹³ Uma função de *hash* efectua a projecção de pontos de um domínio para outro, de menores dimensões. Uma função de *hash* muito simples é uma soma de referência (*checksum*) – a soma lógica dos valores de uma sequência binária. Alterando um único bit da sequência de entrada obtém-se uma sequência de saída distinta. Efectuar o inverso, i.e., obter a sequência de entrada a partir da de saída, é virtualmente impossível.

A grande motivação das técnicas de impressão digital é o combate à pirataria, permitindo monitorizar e identificar de forma automática as ilegalidades ocorridas no trajecto percorrido por um dado produto.

2.4.7 Controlo de utilização

Por controlo de utilização entende-se a permissão dada aos utilizadores no acesso aos produtos, nomeadamente a realização de cópias e a sua visualização/impressão. A protecção contra cópias não autorizadas é de extrema importância nos dias de hoje, dada a crescente comercialização de vídeo em formato digital. A protecção contra visualização/impressão não autorizadas acaba por ser uma extensão deste problema, aplicado a imagens fixas com carácter confidencial e que encontra aplicação principalmente nos sectores militar e médico.

Protecção contra cópias

A gravação de vídeo em suporte digital – DVD – é uma realidade que se espera vá substituir, a breve trecho, o sistema analógico VHS¹⁴. No sistema DVD, o vídeo encontra-se comprimido segundo a norma MPEG-2. Este novo meio de distribuição, para além de permitir um armazenamento mais eficiente, conduz também uma maior durabilidade dos produtos. No entanto, com o aparecimento de gravadores DVD, surge naturalmente o receio, por parte das produtoras cinematográficas e de vídeo, da cópia ilegal com uma qualidade idêntica à do produto original.

Uma forma eficaz de impedir a realização de cópias ilegais é inserir, no próprio conteúdo vídeo, informação de controlo referente a permissões de gravação. Obviamente, para que este sistema seja efectivo, é necessário que os gravadores DVD sejam normalizados de acordo com este objectivo, extraíndo a informação inserida e dando-lhe a interpretação correcta.

Os requisitos para este tipo de aplicações são particularmente exigentes – a marca-de-água terá de ser imperceptível, extremamente robusta a compressões elevadas e o ruído introduzido pela presença da marca não deverá prejudicar o desempenho da compressão MPEG. Uma vez que a maioria dos leitores DVD permite efectuar transformações geométricas (e.g., *zoom*), a marca-de-água deverá também ser robusta a estas transformações. Para além da robustez a técnicas de processamento próprias do formato DVD, a marca deverá ainda ser robusta a ataques intencionais que visem a sua alteração ou remoção.

¹⁴ VHS – *Video Home System*.

Permissão para visualização/impressão

A visualização ou impressão de documentos, imagens ou vídeo, podem ser condicionadas com recurso às marcas-de-água. Neste caso, cada produto é classificado em categorias, desde altamente confidencial, para a qual não é permitido nem a visualização, nem a impressão, até uma consulta totalmente livre, passando por categorias intermédias, para as quais é possível efectuar cópias, modificar o conteúdo, etc. Os dispositivos responsáveis pela visualização ou impressão deverão verificar o grau de confidencialidade do documento em causa, por extracção da marca-de-água e, de acordo com a informação extraída, proporcionar ao utilizador as permissões a que este tem direito.

2.4.8 Transporte de informação adicional

A utilização de marcas-de-água permite estabelecer um processo comunicação em que a marca-de-água é a mensagem e o produto assinado (imagem, vídeo, som ou texto), constitui o canal de transmissão. As aplicações deste tipo variam consoante a categoria da informação transportada [10]:

- *Pública* – se a informação for acessível a todos os que acederem ao produto marcado;
- *Privada* – se o produto marcado puder circular livremente, mas a informação adicional que transporta só for acessível a determinadas entidades;
- *Escondida ou secreta* – se o produto em que se encontra a informação adicional servir apenas para encobrir a mensagem; na prática, estabelece-se um canal secreto sobre um canal aparente, o que constitui um dos princípios da “informação escondida” (*steganography*).

Informação pública

Como exemplos de informação pública acessível através das marcas-de-água, tem-se:

- Informação adicional que tradicionalmente é transportada em cabeçalhos e que geralmente não permanece associada à imagem após conversões de formatos;
- Apontadores para páginas em bases-de-dados universais, onde se encontra armazenada informação de interesse público e relacionada com o produto;

- Informação pública sobre a origem do produto e que se enquadra em algumas das aplicações atrás descritas, como a identificação e autenticação, casos em que o utilizador tem interesse em saber a proveniência dos produtos que adquire.

Informação privada

Neste caso, a informação transportada pela marca-de-água poderá consistir em bits de paridade que permitam a detecção ou correcção de erros ocorridos na transmissão das imagens ou, no caso das sequências de vídeo, poderá consistir em informação para sincronização de som e imagem.

Informação escondida (*steganography*)

A utilização de marcas-de-água como um canal secreto pode ter inúmeras aplicações. Como exemplos mais significativos, encontram-se as comunicações militares ou que envolvam segredos de estado, comunicações entre instituições financeiras em que o sigilo deve ser preservado, ou dentro do sector médico, mantendo segredos dados médicos pessoais. Para este tipo de aplicações, não existem normalmente preocupações quanto a ataques intencionais, pois torna-se difícil para um atacante conseguir prever esta forma de comunicação. No entanto, é exigida robustez em relação a ataques não intencionais como a compressão, já que a informação audiovisual circula habitualmente em formato comprimido.

2.4.9 Síntese de requisitos

De acordo com o atrás exposto, existe um vasto conjunto de aplicações que podem beneficiar com a utilização de técnicas de marca-de-água. Dependendo da aplicação, os requisitos exigidos aos algoritmos de marca-de-água variam. A tabela 2.1 sintetiza os requisitos necessários para cada uma das aplicações descritas [10,15]. Nesta tabela, o símbolo “X” assinala as formas de ataque (designadas pelas siglas A1 a A5 e descritas em 2.3.8) a que a marca deve ser robusta. O símbolo “?” assinala requisitos que podem ou não ser exigidos.

2.5 “Ambiguidades” nas marcas-de-água

Uma das formas mais simples de “atacar” uma imagem marcada é inserir uma outra marca, dificultando-se assim a identificação da marca original e, no contexto da protecção dos direitos de autor, a identidade do detentor dos direitos.

Aplicações	Ataques não intencionais		Ataques intencionais			Payload (bits)	Extracção / Detecção cega (C) vs. não-cega (NC)
	A1	A2	A3	A4	A5		
Prova de propriedade	X	X	X	X	X	1*; 48-64	NC
Identificação	X	X	X	X	X	64	?
Autenticação	X				X	48-64	?
Monitorização	X	X			X	64-72	C
Verificação da integridade	?	?	?		X	48-64	C
Impressões digitais	X	X	?			16-64	?
Controlo de utilização	X	X	X	X	X	4-8	C
Informação adicional pública	X	X	?			48-64	C
Informação adicional privada	X	X	?		X	48-64	?
Informação adicional secreta	X	X	?	?	?	≥ 256	?

* No caso de *detecção privada*.

Tabela 2.1 – Síntese de requisitos comuns.

Uma possível solução para este problema é a utilização de selos temporais (*time-stamping*), atribuídos por uma entidade certificadora (e.g., sociedades de autores). A marca-de-água correspondente ao selo mais antigo será considerada a original [36,47].

Uma outra solução passa pelo registo, numa entidade certificadora, do produto assinado e/ou do produto original e, por comparação entre o produto em questão e o registado, apura-se o legítimo proprietário. No entanto, este registo só será suficiente para levantar a ambiguidade se o algoritmo de inserção da marca não for *invertível* [12], o que pode ser conseguido fazendo depender a marca-de-água a inserir, da imagem original.

Com efeito, suponha-se que a entidade A (o autor/detentor) assina uma imagem I com uma marca w e uma chave K , obtendo a imagem marcada I_w , ou seja $I_w = f_A(I, w, K)$; suponha-se que posteriormente uma entidade fraudulenta B consegue gerar um falso original I_f e uma falsa marca w_f de tal modo que após inserção da falsa marca no falso original é obtida a imagem I_w , ou seja, $I_w = f_B(I_f, w_f, K_f)$. Como a diferença entre I e I_w é a assinatura de A , então I conterà também a assinatura de B . Por outro lado, se a técnica de marca-de-água utilizada por A for suficientemente robusta, então I_f conterà também a assinatura de A . Em suma, o verdadeiro

original, na posse de A , conterá a assinatura de B ; o falso original, na posse de B , conterá a assinatura de A . A resolução deste problema passa pela utilização de uma marca-de-água dependente da imagem original, dependência essa que pode ser conseguida, por exemplo, utilizando funções de *hash* (não invertíveis) que façam depender a marca a inserir da imagem a marcar.

Capítulo 3

Espalhamento de espectro em marcas-de-água

3.1 Introdução

A assinatura digital de imagens pode ser considerada como um processo de comunicação em que se pretende enviar uma mensagem – a marca-de-água – através de um canal¹⁵ ruidoso – a imagem. Nesta medida, conceitos e resultados da teoria da comunicação, tais como modulação, codificação de canal ou detecção óptima, podem ser extendidos aos sistemas de marca-de-água, com o intuito de melhorar o seu desempenho.

A figura 3.1 apresenta um esquema genérico da assinatura digital de imagens, visto numa perspectiva de comunicação digital e onde se evidenciam os procedimentos, ao nível do emissor e receptor, que foram estudados nesta tese: a inserção da marca-de-água (mensagem a enviar) na imagem, é precedida por um processo de codificação de canal e modulação da marca, sendo a extracção da marca precedida pelos processos inversos (desmodulação e decodificação de canal).

¹⁵ Neste caso, o canal não é uma realidade física, tal como nos sistemas de comunicação usuais, já que todo o processo de comunicação se desenrola num domínio puramente digital.

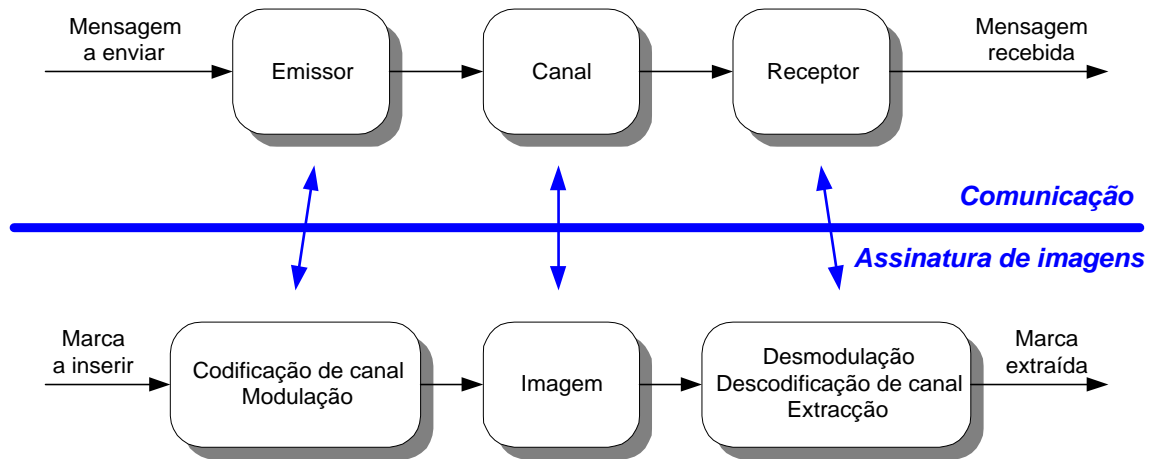


Figura 3.1 – A tecnologia das marcas-de-água numa perspectiva de comunicação digital.

Este capítulo centra-se fundamentalmente no processo de modulação. Sendo o alfabeto da mensagem a transmitir constituído por um conjunto de M ($M \geq 2$) símbolos distintos, a modulação consiste na transformação (unívoca) de cada símbolo da mensagem, num sinal (ou forma de onda) de um conjunto de M formas de onda distintas. Existem diversas formas de modulação, sendo a sua escolha condicionada pela aplicação em vista. No caso da assinatura de imagens, os requisitos *robustez* (i.e., resistência a formas comuns de processamento de sinal) e *segurança* (i.e., resistência a esforços intencionais para remover a marca) levaram a que *I. J. Cox* sugerisse em [11] a utilização da modulação por *espalhamento de espectro com sequência directa*. Sumariamente, esta técnica de modulação, no domínio digital, consiste na representação de cada símbolo do alfabeto da mensagem por uma sequência pseudo-aleatória, designada por *sequência de espalhamento*, com uma largura de banda muito superior à da mensagem original.

Em muitas das técnicas de assinatura digital de imagens, o alfabeto da mensagem é constituído por apenas dois símbolos: o símbolo 0 e o símbolo 1. Neste caso, diz-se que é utilizada sinalização binária e a modulação por espalhamento de espectro requer a existência de pelo menos duas sequências de espalhamento distintas: uma para representar o símbolo 0 e outra para representar o símbolo 1 (podem ser duas sequências simétricas). Se o alfabeto da mensagem for constituído por M símbolos distintos, com $M > 2$, está-se em presença de um sistema multinível e a sinalização (e a modulação subsequente) é designada por *M-ária*.

Em sistemas de comunicação, a utilização de alfabetos com vários símbolos e a modulação *M-ária* correspondente conduzem, para certos esquemas de modulação, a uma redução na probabilidade de erro de bit na recepção [34]. Como demonstrado em [25], este tipo de modulação pode também contribuir para uma melhoria do desempenho dos sistemas de marca-de-água.

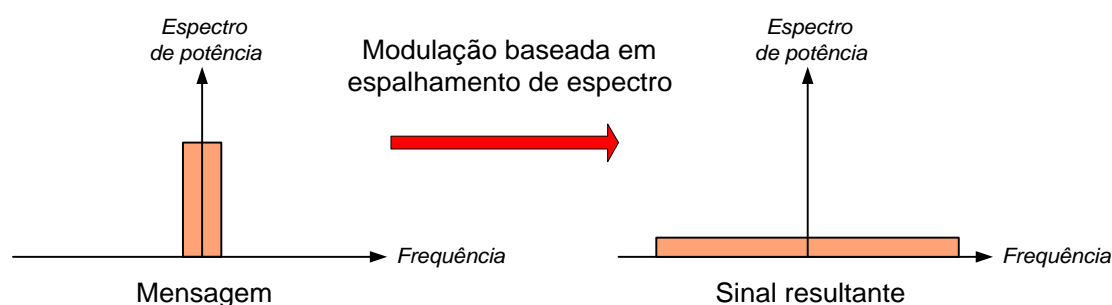


Figura 3.2 – Espalhamento do espectro de um sinal.

Ao longo deste capítulo ir-se-á estudar a aplicação de modulação multinível baseada em espalhamento de espectro à assinatura digital de imagens. A utilização desta forma de modulação em sistemas de marca-de-água foi proposta em [25], para inserção no domínio espacial. O estudo apresentado em [25] é revisto neste capítulo, sendo extendido para aplicação a esquemas de marcas-de-água com inserção no domínio da frequência. O sistema de inserção/extracção de marca-de-água implementado (incluindo modulador e desmodulador) constitui a base de estudo para os capítulos 4 e 5 desta tese.

Este capítulo encontra-se estruturado em oito secções. Após a introdução efectua-se, na secção 3.2, uma breve introdução ao espalhamento de espectro e apresentam-se algumas formas de aplicar esta modulação à assinatura digital de imagens. Estabelecidos os princípios do espalhamento de espectro descrevem-se, nas secções 3.3 e 3.4 respectivamente, os esquemas gerais para inserção e extracção da marca-de-água utilizados ao longo desta tese. Nas secções 3.5 e 3.6 apresenta-se um estudo analítico e experimental da aplicação de modulação multinível em marcas-de-água, seguindo duas abordagens distintas: inserção/extracção no domínio espacial e inserção/extracção no domínio da DCT. Em 3.7 comparam-se os resultados obtidos nestes dois domínios e em 3.8 sumarizam-se as principais conclusões do estudo efectuado neste capítulo.

3.2 Espalhamento de espectro: princípios básicos e aplicação à assinatura de imagens

O espalhamento de espectro é uma técnica de modulação especialmente concebida para combater elevados níveis de interferência no canal de comunicação e/ou para evitar que a mensagem seja recebida por outros receptores que não o desejado [34]. Como é evidenciado pela sua designação, esta técnica “espalha” o espectro do sinal a transmitir, numa banda de frequências muito mais larga do que a banda do sinal da mensagem, tal como ilustrado na figura 3.2.

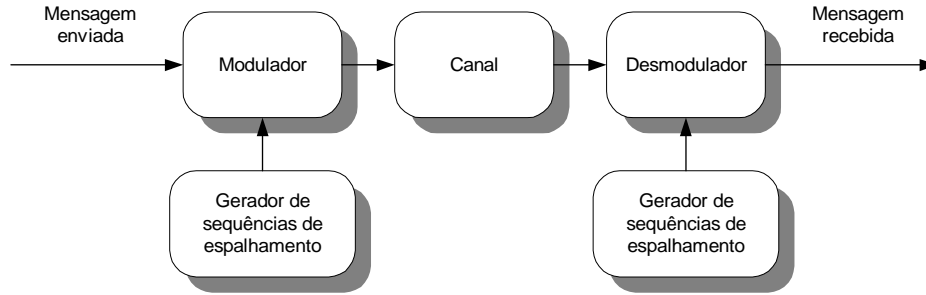


Figura 3.3 – Comunicação digital com modulação baseada em espalhamento de espectro.

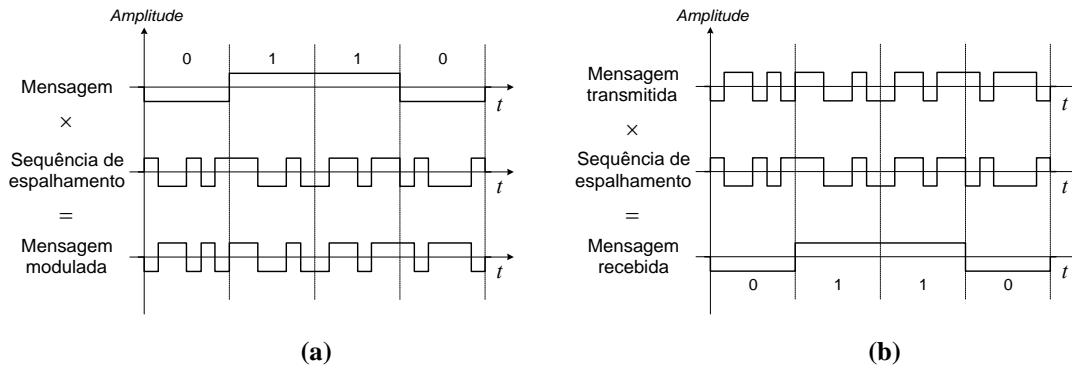


Figura 3.4 – Exemplo de modulação (a) e de desmodulação (b) por espalhamento de espectro.

Na figura 3.3 representa-se, em diagrama de blocos, um sistema de comunicação com espalhamento de espectro. Como se pode constatar desta figura, o espalhamento de espectro envolve a geração de sequências pseudo-aleatórias designadas por *sequências de espalhamento*¹⁶. Uma vez que as sequências de espalhamento utilizadas na modulação são também necessárias para uma correcta desmodulação, um receptor necessita de as conhecer a fim de receber a mensagem.

A figura 3.4 apresenta um exemplo simples, ilustrativo da modulação baseada em espalhamento de espectro com *sequência directa*, quando aplicada à mensagem binária “0110”, representada utilizando sinalização antipodal¹⁷. A designação de espalhamento de espectro com sequência directa, deriva do facto de o sinal modulado resultar directamente da multiplicação da sequência de espalhamento pelos bits da mensagem (figura 3.4-a)). Na recepção, a mensagem é recuperada multiplicando-se a sequência modulada recebida pela sequência de espalhamento (figura 3.4-b)). No exemplo apresentado, é de realçar o facto do ritmo binário do sinal modulado – R_s – ser bastante superior ao do sinal da mensagem – R_b – do que resulta o aumento da largura de banda atrás referida. O quociente R_s/R_b é usualmente designado por *factor de expansão da banda* (*pulse size* ou *chip rate*).

¹⁶ É também habitual na literatura utilizar-se o termo *sequências de pseudo-ruído*.

¹⁷ Em sinalização binária antipodal os símbolos 1 e 0 são representados por +1 e -1, respectivamente.

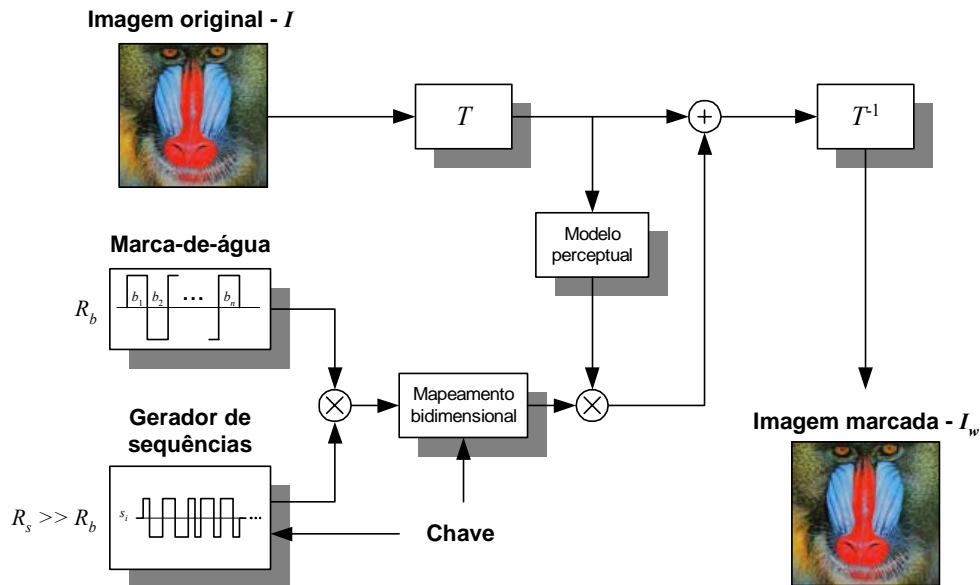


Figura 3.5 – Assinatura de imagens baseada em espalhamento de espectro.

A modulação baseada em espalhamento de espectro é especialmente atraente quando se pretende [34]:

- Transmitir uma mensagem através de um canal sujeito, potencialmente, a interferência propositada (*jamming*)¹⁸;
- “Esconder” a mensagem, transmitindo-a com baixa potência de modo a ser confundida com ruído no canal¹⁸;
- Assegurar a privacidade da mensagem, na presença de outros receptores.

No caso da assinatura digital de imagens, vista como um processo de comunicação, pretende-se transmitir a marca-de-água de uma forma imperceptível, sem diminuir o valor comercial da imagem assinada. Para tal, a marca deverá ser inserida com um nível de energia muito inferior ao da imagem. Pretende-se ainda que a marca inserida seja correctamente detectada, mesmo após a imagem ter sido sujeita a determinadas manipulações, tais como compressão (JPEG ou MPEG) ou filtragens e ainda que a marca seja apenas correctamente detectada por um receptor autorizado. Estes requisitos sugerem ser a assinatura digital de imagens uma aplicação na qual se poderá tirar o máximo partido da utilização de modulação baseada em espalhamento de espectro.

¹⁸ De facto, pelo teorema de *Shannon* [38] um aumento na largura de banda de transmissão permite uma menor relação sinal-ruído na recepção, para a mesma capacidade do canal. A utilização de espalhamento de espectro expande a largura de banda do sinal e, consequentemente, aumenta a robustez ao ruído ou, para o mesmo nível de ruído, permite potências de emissão mais baixas.

Na figura 3.5 apresenta-se um esquema genérico de um sistema de marcas-de-água que utiliza espalhamento de espectro por sequência directa, com $M=2$. Supõe-se que a marca é uma sequência binária, constituída por N_b símbolos antipodais – $B = \{b_1 \dots b_{N_b}\}$, $b_i = \pm 1$, \forall_i . A imagem I é definida como um conjunto de *pixels* que representam a informação visual. Numa imagem policromática, definida no espaço RGB (Vermelho-Verde-Azul), cada *pixel* é um conjunto de três valores:

$$I(m,n) = \{ R(m,n), G(m,n), B(m,n) \}, \quad (3.1)$$

em que $(m,n) \in \wedge^2$ é a localização espacial num sistema de coordenadas cartesiano e $R(m,n)$, $G(m,n)$, $B(m,n)$ são as componentes de cor, na posição (m,n) , no sistema de cores RGB. Os valores das componentes de cor são habitualmente discretizados utilizando 8 bits.

A inserção da marca na imagem é efectuada de acordo com as seguintes etapas:

1. Geração de sequências pseudo-aleatórias (*sequências de espalhamento*) – s_i – dependentes de uma chave secreta e com um débito binário – R_s – muito superior ao débito binário da mensagem – R_b ;
2. Multiplicação (*espalhamento*) dos bits da marca pelas sequências de espalhamento;
3. Mapeamento das sequências obtidas na etapa 2 num espaço bidimensional (2D) e de uma forma que garanta a ausência de sobreposição espacial entre as várias sequências. Este mapeamento deve depender de uma chave secreta e garantir que cada sequência seja uniformemente dispersa pelo espaço 2D;
4. Multiplicação do sinal 2D obtido na etapa 3 por um factor local (ou *factor perceptual*) obtido de acordo com um modelo do sistema visual humano (ou *modelo perceptual*);
5. Adição do sinal resultante na etapa 4, que funciona neste contexto como ruído aditivo, ao espaço de inserção da marca-de-água, obtido a partir da imagem original I através de uma transformação T ;
6. Aplicar a transformação inversa de T ao resultado obtido na etapa 5, obtendo-se assim a imagem marcada I_w .

O factor perceptual deverá tomar valores baixos em posições de imagem onde eventuais alterações sejam perceptíveis, como em zonas homogéneas, e deverá tomar valores mais elevados em posições de imagem onde eventuais alterações sejam menos perceptíveis, como em zonas texturadas.

Relativamente ao espaço de inserção da marca-de-água, existem essencialmente duas abordagens: inserção no *domínio espacial* e inserção no *domínio da frequência*.

No primeiro caso – domínio espacial – a marca-de-água é inserida directamente no espaço da imagem, mais precisamente numa ou mais componentes da cor. Como exemplos, em [25] o autor sugere a utilização da componente *B* (azul) e em [20] é sugerida a inserção na componente da luminância. A inserção na componente da luminância é a mais utilizada, por esta componente da cor estar directamente acessível em muitos formatos de imagem digital. Para além disso, verifica-se experimentalmente que a inserção na componente da luminância é mais robusta do que a inserção na componente *B*. Por estes motivos, o espaço das marcas-de-água (referente ao domínio espacial) utilizado neste trabalho corresponde à luminância da imagem.

A assinatura digital de imagens no domínio da frequência, tal como a sua designação indica, pressupõe que é realizada uma transformação em frequência sobre a imagem, sendo a marca-de-água inserida no espaço resultante da transformação. Entre as transformações de frequência mais utilizadas encontram-se:

- A transformada *Wavelet* discreta (*DWT*) – A principal vantagem desta transformação é efectuar uma decomposição em bandas de frequência “semelhante” à efectuada pelo sistema visual humano, o que possibilita o desenvolvimento e aplicação de modelos perceptuais adequados e otimizar a energia da marca inserida, para uma dada qualidade visual. Este facto é explorado em [2], onde os autores apresentam um esquema de marcas-de-água baseado em espalhamento de espectro, com inserção no domínio da transformada *DWT*. Esta e outras vantagens, com destaque para a reduzida complexidade da *DWT*, são sintetizadas em [30], onde é efectuado um estudo comparativo de vários algoritmos de marca-de-água baseados na *DWT*.
- A transformada de *Fourier* discreta (*DFT*) – Apesar de pesada computacionalmente, é utilizada em esquemas de marcas-de-água robustos a manipulações geométricas simples e a cortes (*Crop*) na imagem marcada [11,29].

- A transformada do coseno discreta global (*DCT* global) – Embora computacionalmente menos complexa que a *DFT*, apresenta características similares, nomeadamente na robustez a cortes na imagem marcada. No trabalho desenvolvido em [11], onde é também utilizado o espalhamento de espectro, a inserção das marcas-de-água é efectuada neste domínio.
- A transformada *DCT* orientada ao bloco de dimensões 8×8 pixels (*DCT* 8×8) – Amplamente utilizada em esquemas de assinatura digital de imagens [18], por ser usada também nas normas de compressão actuais (JPEG e MPEG), evitando-se assim o cálculo de transformações adicionais para inserção da marca. É ainda de salientar o facto de existir actualmente um estudo muito completo sobre os modelos perceptuais referentes a aplicações com este tipo de transformada, tanto para compressão, como para marcas-de-água.

Devido às vantagens apresentadas, nomeadamente o facto de ser computacionalmente simples e de se enquadrar bem no actual panorama de armazenamento e distribuição de imagens fixas e vídeo [16], neste trabalho optou-se pela inserção no domínio da transformada *DCT* orientada ao bloco.

3.3 Inserção da marca-de-água

3.3.1 Esquema geral de inserção

O esquema de inserção da marca-de-água por espalhamento de espectro a analisar pode ser observado na figura 3.6. Embora semelhante ao esquema básico apresentado na figura 3.5 evidencia, através do bloco “Modulação multinível”, inexistente na figura 3.5, o assunto a ser estudado ao longo deste capítulo. Admite-se que se pretende inserir a marca-de-água no espaço bidimensional X , resultante da aplicação da transformação T , à imagem original. A transformação T é, no contexto desta tese, o simples cálculo da componente de luminância da imagem, ou a transformada *DCT* da imagem (orientada ao bloco de dimensão 8×8 pixels).

Como mencionado na secção anterior, a marca a inserir é uma sequência binária constituída por N_b bits de informação – $B = \{b_1 \dots b_{N_b}\}$. Supondo que o sistema utiliza modulação M -ária, com M níveis de sinalização, os N_b bits que constituem a marca são mapeados para N_s símbolos, obtendo-se a sequência $B_s = \{a_1 \dots a_{N_s}\}$. Cada símbolo a_i pode tomar um valor entre M valores diferentes – A^1, \dots, A^M . Este mapeamento obtém-se através do simples agrupamento de $\log_2(M)$ bits da mensagem original e utilizando o valor decimal resultante como um índice para seleccionar o símbolo correspondente. Deste modo, a relação entre N_s e N_b é dada por:

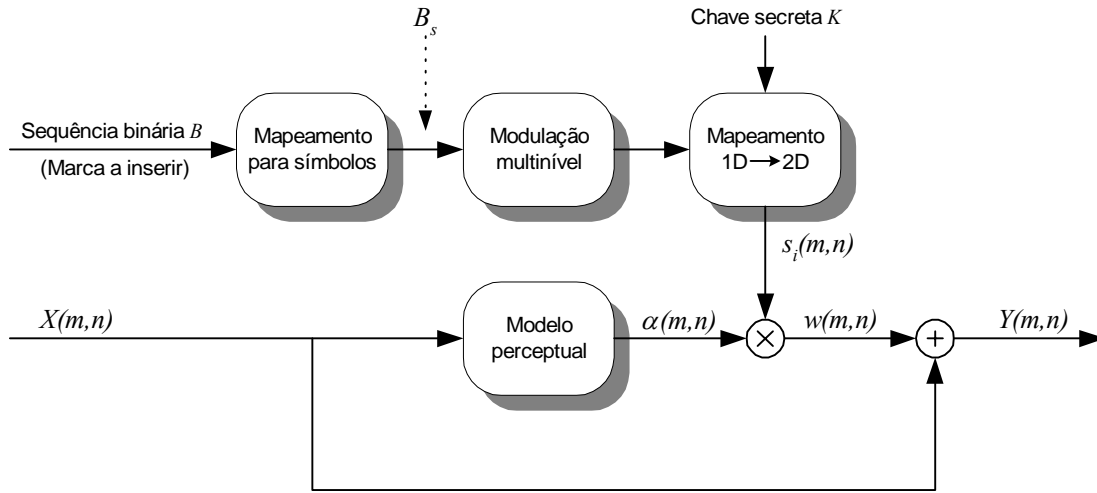


Figura 3.6 – Esquema geral de inserção da marca-de-água.

$$N_s = \frac{N_b}{\log_2 M}. \quad (3.2)$$

Seguindo a abordagem proposta em [25], a sequência B_s é modulada utilizando M sequências bi-ortogonais $-s^1$ a s^M . Estas sequências devem possuir certas propriedades, nomeadamente terem média nula e variância unitária, como será demonstrado na secção 3.5.2. A utilização de M sequências bi-ortogonais requer a geração de $M/2$ sequências ortogonais, que são utilizadas para modular os símbolos A^1 a $A^{M/2}$, e de $M/2$ sequências antipodais das anteriores, que são usadas para modular os símbolos restantes $-A^{M/2}$ a A^M . O modulador multinível tem como função associar uma sequência de modulação $s_i \in \{s^1, \dots, s^M\}$, $i = \{1 \dots N_s\}$, a cada símbolo a_i da mensagem B_s .

A saída do modulador é aplicada a um bloco cuja função é mapear as sequências de modulação para posições físicas da imagem (*pixels*). Este mapeamento é pseudo-aleatório, dependendo de uma chave secreta K e a sua inversão só é possível se a chave K for conhecida. Deste modo, é garantida a segurança do sistema. Doravante, a notação $s_i(m,n)$ designará o elemento da sequência s_i mapeado para a posição (m,n) da imagem. Será ainda assumido que $s_i(m,n)=0$ caso nenhum elemento de s_i seja mapeado para a posição (m,n) . Assim, se S_i representar o conjunto de todas as posições da imagem para as quais s_i foi mapeada, então $S_i \cap S_j = \emptyset$, $\forall i \neq j$, o que significa que os mapeamentos espaciais das sequências não se sobrepõem.

Após o mapeamento espacial, os valores de $s_i(m,n)$ são multiplicados por um factor $\alpha(m,n)$, cujo propósito é adaptar a inserção da marca-de-água ao sistema visual humano. A marca-de-água w

pode ser então definida como a sobreposição de todas as sequências de modulação s_i , mapeadas para posições bidimensionais e multiplicadas por $\alpha(m, n)$, isto é:

$$w(m, n) = \sum_{i=1}^{N_s} \alpha(m, n) s_i(m, n). \quad (3.3)$$

Para finalizar o processo de inserção da marca-de-água, a marca w é adicionada ao espaço X , resultando o espaço marcado Y :

$$Y(m, n) = X(m, n) + w(m, n). \quad (3.4)$$

A imagem marcada é obtida aplicando-se a transformação inversa de T a Y .

Descrito de uma forma sucinta o processo de inserção da marca-de-água, aprofundam-se nos pontos seguintes o funcionamento dos blocos mais relevantes no esquema de inserção: modulação multinível, mapeamento bidimensional e modelo perceptual.

3.3.2 Modulador multinível

Como referido na secção anterior, as sequências de espalhamento deverão ser ortogonais ou bi-ortogonais, de modo a que a distância entre estas seja maximizada. No sistema em estudo, optou-se pela utilização de sequências de espalhamento bi-ortogonais, já que conduzem a esquemas de desmodulação menos complexos [34].

Um conjunto de M funções ortogonais, com M potência de 2, pode ser obtido recursivamente a partir da matriz de *Hadamard* de ordem 2, utilizando o produto de *Kronecker* (\otimes):

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; \quad (3.5)$$

$$\mathbf{H}_M = \mathbf{H}_2 \otimes \mathbf{H}_{M/2} = \begin{bmatrix} \mathbf{H}_{M/2} & \mathbf{H}_{M/2} \\ \mathbf{H}_{M/2} & -\mathbf{H}_{M/2} \end{bmatrix}, \quad (3.6)$$

em que \mathbf{H}_2 é a matriz de *Hadamard* de ordem 2 e \mathbf{H}_M a matriz de *Hadamard* de ordem M , cujas linhas (ou colunas) servem de base para um conjunto de M funções ortogonais¹⁹. Como exemplo, a matriz de *Hadamard* de ordem 4 – \mathbf{H}_4 – que servirá de base a um sistema de 4 níveis, construída recursivamente a partir do método referido, é a seguinte:

¹⁹ Também designadas por funções ortogonais de *Hadamard-Walsh*.

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \quad (3.7)$$

Para se obter um conjunto de M sequências bi-ortogonais basta juntar, ao conjunto formado por $M/2$ sequências ortogonais, o simétrico das mesmas. Utilizando a representação matricial, a matriz \mathbf{B}_M que contém a base para a construção de M sequências bi-ortogonais, pode-se escrever como:

$$\mathbf{B}_M = \begin{bmatrix} \mathbf{H}_{M/2} \\ -\mathbf{H}_{M/2} \end{bmatrix}. \quad (3.8)$$

Como exemplo, a matriz \mathbf{B}_4 contendo a base para 4 sequências de espalhamento bi-ortogonais (as suas linhas) é:

$$\mathbf{B}_4 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (3.9)$$

Observando a matriz \mathbf{B}_M , para M arbitrário, verifica-se que existem duas linhas cuja média não é nula. De modo a anular o valor médio na construção das M sequências de espalhamento, cada uma com comprimento λ_s , a linha de \mathbf{B}_M que serve de base à sequência é replicada alternadamente com o seu simétrico, até ser atingido o comprimento λ_s . Designando por $\mathbf{S}_M(\lambda_s)$ o conjunto de sequências de espalhamento de ordem M e com comprimento λ_s , ter-se-á:

$$\mathbf{S}_M(\lambda_s) = \overbrace{[\mathbf{B}_M \quad -\mathbf{B}_M \quad \cdots \quad \mathbf{B}_M \quad -\mathbf{B}_M]}^{\lambda_s \text{ Colunas}}. \quad (3.10)$$

De acordo com o descrito, um conjunto possível de sequências de espalhamento bi-ortogonais com comprimento 8, necessárias para um sistema de 4 níveis, é:

$$\mathbf{S}_4(8) = \begin{bmatrix} 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}. \quad (3.11)$$

Resta acrescentar que a pares de símbolos antipodais, isto é, pares de símbolos cuja representação binária é complementar, fazem-se corresponder pares de sequências de

espalhamento simétricas. Deste modo, assegura-se que símbolos cuja representação binária está a maior distância, têm representações em termos de sequências de espalhamento também mais distantes.

3.3.3 Mapeamento bidimensional

Após a modulação dos símbolos que constituem a marca pelas sequências de espalhamento, obtém-se um sinal unidimensional (1D) que deverá ser mapeado num espaço bidimensional (2D) coincidente com o espaço da imagem. Este mapeamento deverá garantir que cada sequência fique uniformemente espalhada pela imagem, ou seja, que posições da imagem onde um mesmo símbolo da marca é inserido não sejam vizinhas. Esta medida serve para evitar que posições da imagem contíguas sejam atribuídas a um mesmo símbolo, o que poderia ser prejudicial caso essas posições correspondessem a uma zona homogênea, onde a força de inserção terá de ser reduzida.

De modo a garantir a segurança do sistema, a atribuição de posições de imagem a cada elemento das sequências de espalhamento é feita de forma pseudo-aleatória e de acordo com uma chave secreta K . Essa chave é a semente de um gerador de números pseudo-aleatórios, base da construção de uma tabela de atribuição de posições de imagem aos elementos das sequências de espalhamento.

Na figura 3.7 pode ser observado um pequeno exemplo do funcionamento do mapeamento bidimensional para 4 símbolos a inserir (sequências s_0 a s_3) e uma imagem de dimensão 4×4 pixels. Com base numa dada chave K é gerada uma tabela de atribuições de posições de imagem (figura 3.7-b)). O resultado final do mapeamento é obtido tendo em conta os índices (m,n) de posições de imagem atribuídos a cada um dos elementos das sequências de espalhamento. Pode observar-se, como exemplo, que ao terceiro elemento da sequência s_1 (-1) é atribuída a posição de imagem (2,3).

3.3.4 Modelo perceptual

Como já referido, as modificações provocadas pela inserção da marca-de-água no produto a proteger não devem degradar a sua qualidade de forma perceptível. Dois fenómenos autorizam que, até certa medida, a adição de “ruído”²⁰ a uma imagem não seja apercebida pela vista humana:

²⁰ Com efeito, a diferença entre as imagens marcadas e as imagens originais pode ser considerada como “ruído” introduzido nas imagens.

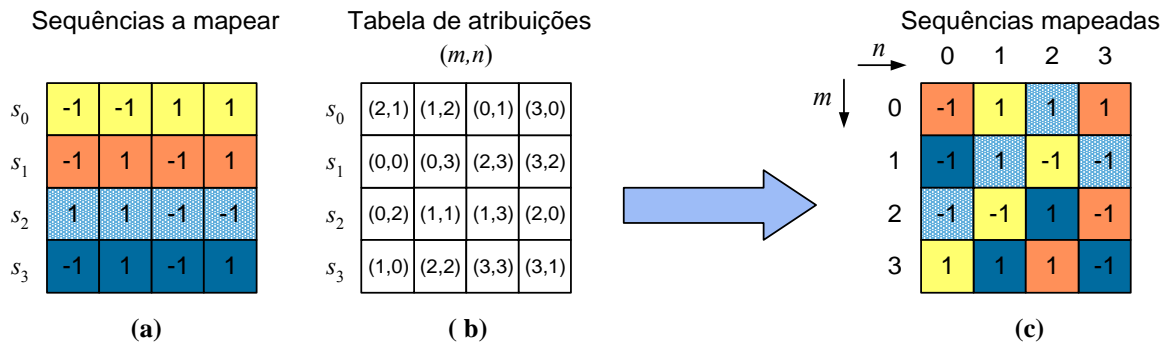


Figura 3.7 – Exemplo ilustrativo do mapeamento bidimensional:
a) Sequências a mapear; b) Tabela de atribuições; c) Sequências mapeadas.

- A percepção do ruído depende da sua distribuição em frequência, sendo o ruído de alta frequência menos visível que o de baixa frequência. De facto, o olho humano funciona aproximadamente como um filtro passa-baixo, tanto no domínio das frequências espaciais como temporais;
- Estímulos visuais podem-se “mascarar” (no sentido que a sua percepção pela vista humana é reduzida) uns aos outros. Este fenómeno (*masking*) garante a redução de visibilidade do ruído por outro estímulo existente na zona onde esse ruído é introduzido.

O *modelo perceptual* do sistema de inserção da marca-de-água tem a função de adaptar a força de inserção da marca, de modo a que as alterações causadas à imagem original sejam imperceptíveis ao sistema visual humano (SVH). De acordo com os fenómenos acima descritos, a marca pode ser inserida com mais energia em posições de imagem que apresentem muita textura, dado que o SVH é menos sensível a alterações em zonas cuja actividade espacial é intensa (altas frequências espaciais). Pelo contrário, em posições de imagem correspondentes a zonas homogêneas, onde o SVH é mais sensível a alterações, a marca deverá ser inserida com menos energia. Verifica-se também que a sensibilidade do SVH depende do valor médio da luminância, sendo a sensibilidade mais elevada quanto mais baixo for esse valor.

Existe uma vasta literatura em modelos perceptuais. Na escolha do modelo a usar, deverá atender-se à forma como é inserida a marca-de-água e ao domínio da imagem em que é inserida. No caso concreto do sistema em estudo foram utilizados dois modelos diferentes, um para inserção no domínio espacial e outro para inserção no domínio dos coeficientes DCT.

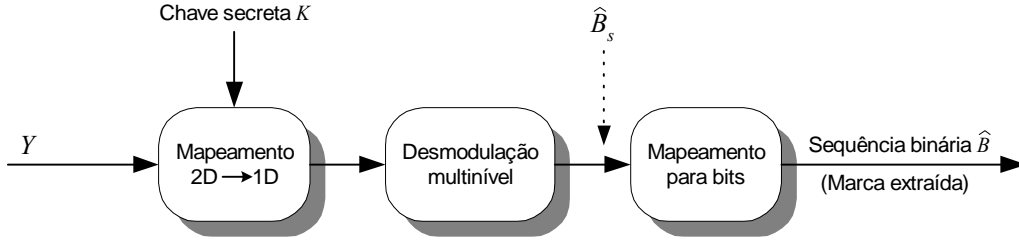


Figura 3.8 – Esquema geral de extracção da marca-de-água.

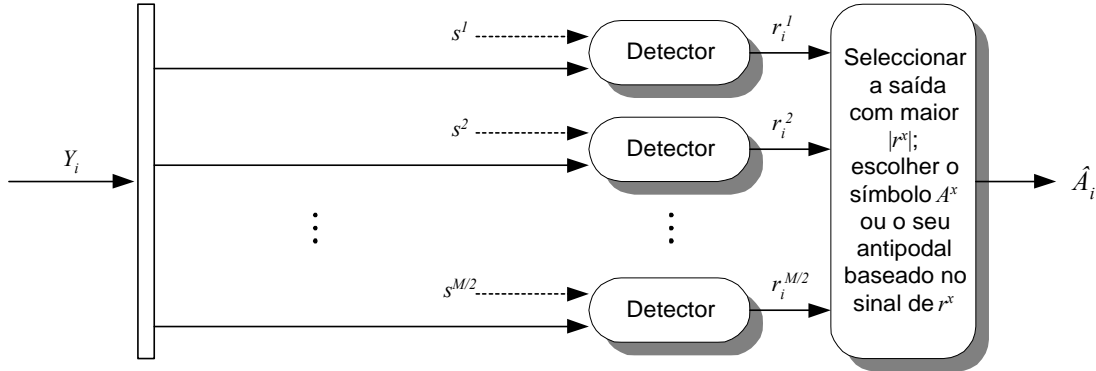


Figura 3.9 – Esquema geral do desmodulador multinível.

3.4 Extracção da marca-de-água

A extracção da marca sem o conhecimento da imagem original é possível com a utilização do sistema representado na figura 3.8. Admite-se que a imagem marcada é previamente transformada para o espaço bidimensional Y , no qual foi inserida a marca.

Na primeira etapa, o sinal bidimensional Y é mapeado para um espaço unidimensional. Este processo consiste na operação inversa do mapeamento bidimensional realizado na inserção da marca (secção 3.3.3). Utilizando uma chave idêntica à usada para inserir a marca, o mapeador unidimensional tem a função de gerar sequências unidimensionais – Y_i – constituídas pelos valores que Y toma nas posições onde se inseriu a_i .

A desmodulação do sinal recebido (figura 3.9) é feita através da combinação das sequências unidimensionais Y_i com as primeiras $M/2$ sequências do conjunto de sequências de espalhamento – $\{s^1, \dots, s^{M/2}\} \in \mathbf{S}_M$. A decisão sobre cada símbolo é dependente dos resultados de cada combinação – $r_i^j, j \in \{1, \dots, M/2\}$ – tal como indicado na figura 3.9.

Para completar o processo de extracção da marca, a sequência de símbolos extraídos é mapeada para uma sequência binária. A sequência binária resultante – \hat{B} – é a marca-de-água recebida.

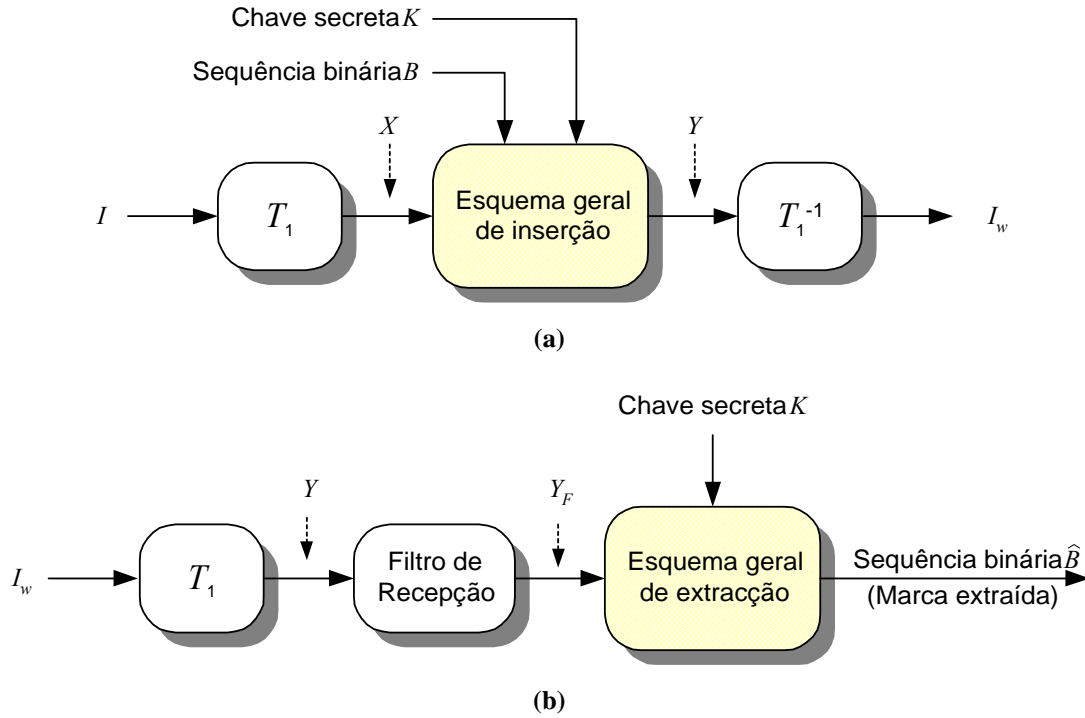


Figura 3.10 – Esquemas de inserção e extração de marcas-de-água no domínio espacial:
a) Inserção; b) Extração.

3.5 Marcas-de-água no domínio espacial

Nesta secção apresenta-se o estudo referente a marcas-de-água com inserção no domínio espacial, utilizando modulação multinível baseada em espalhamento de espectro. No ponto 3.5.1 desta secção descrevem-se os esquemas de inserção/extração da marca-de-água, particularizados para o domínio espacial. Em 3.5.2 é feito um estudo analítico, cujo objectivo é determinar o impacto da utilização da modulação multinível na probabilidade de erro de bit da marca extraída. Para finalizar, no ponto 3.5.3 são apresentados os resultados obtidos, teóricos e experimentais, incluindo resultados experimentais de resistência à compressão JPEG.

3.5.1 Inserção e extração da marca-de-água no domínio espacial

A figura 3.10 apresenta as alterações introduzidas nos esquemas de inserção e extração da marca-de-água descritos anteriormente (secções 3.3 e 3.4), de modo a que estas operações se processem no domínio espacial.

A diferença fundamental em relação aos esquemas gerais atrás apresentados reside no esquema de extração da marca, que inclui neste caso um filtro de recepção, cuja função é aproximar o canal (imagem) de um canal com ruído aditivo, branco e gaussiano (AWGN), para o qual é

conhecido o detector óptimo. Como já demonstrado por outros autores [14], a utilização deste filtro melhora significativamente o desempenho do detector.

Para além do filtro de recepção, os esquemas de inserção/extracção no domínio espacial contêm um modelo perceptual e um desmodulador particularizados para este domínio. Nos pontos seguintes descreve-se, com mais detalhe, cada um destes componentes.

Espaço das marcas-de-água

O espaço das marcas-de-água é, para a inserção espacial, a componente de luminância – L – da imagem. Supondo que a imagem I se encontra definida no sistema de cores RGB, a transformação de coordenadas para o sistema LUV (luminância L – crominância U – crominância V) é obtida através de [35]:

$$\begin{bmatrix} L \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (3.12)$$

pelo que a transformação do espaço da imagem para o espaço das marcas-de-água é dada por:

$$T_1(I(m,n)) = L(m,n) = 0.299R(m,n) + 0.587G(m,n) + 0.114B(m,n). \quad (3.13)$$

A passagem do sistema LUV para o sistema RGB obtém-se por inversão de (3.12):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & -0.000 & 1.139 \\ 1.000 & -0.394 & -0.581 \\ 0.998 & 2.028 & -0.001 \end{bmatrix} \begin{bmatrix} L \\ U \\ V \end{bmatrix}. \quad (3.14)$$

Deste modo ficam definidas as transformações necessárias para operação no domínio espacial, com os esquemas de inserção/extracção da marca-de-água propostos.

Modelo perceptual

Uma forma simples, mas eficaz, de adaptar localmente a força de inserção ao sistema visual humano, é utilizar para medida dessa força o resultado de uma filtragem da imagem do tipo *Laplaciana*. Com efeito, os filtros de Laplace constituem uma classe de filtros diferenciais de 2ª ordem, originando valores mais elevados (em módulo) nas zonas da imagem com maior actividade espacial e valores mais baixos (em módulo) em zonas homogéneas. Estes filtros

podem ser implementados na forma digital de diversos modos, sendo o mais frequente o descrito por (3.15):

$$H = \frac{1}{4} \cdot \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}, \quad (3.15)$$

solução sugerida em [24] e também adoptada nesta tese. O resultado da filtragem é ainda multiplicado por um factor β que regula a força de inserção global. O peso perceptual relativo a cada posição da imagem pode ser escrito como:

$$\alpha(m, n) = \beta \cdot X(m, n) * H(m, n) = \beta \sum_{k,l} X(m-k, n-l) H(k, l), \quad (3.16)$$

onde $*$ representa a operação convolução.

Filtro de recepção

Nos sistemas de marca-de-água que requerem a imagem (ou vídeo) originais para efectuar a extracção da marca (*extracção privada*), a estimativa da marca inserida pode ser facilmente obtida por subtracção da imagem original à imagem marcada, isto é:

$$w(m, n) = Y(m, n) - X(m, n). \quad (3.17)$$

Nos sistemas em que o produto original não é conhecido (*deteccção semi-privada; extracção pública e semi-pública*), torna-se imperativo utilizar um esquema alternativo para detectar a marca inserida.

Segundo a teoria da deteção, o detector do tipo correlador é o óptimo para canais caracterizados por ruído aditivo, branco e Gaussiano (AWGN). Sabe-se também da teoria da deteção, que é possível efectuar a deteção óptima, no caso de ruído não branco, através de um filtro “branqueador” colocado à entrada do correlador.

Embora o espectro de potência das imagens reais esteja longe de poder ser considerado branco e de o projecto de um filtro “branqueador” requerer o conhecimento da imagem original, é possível obter uma filtragem com boas aproximações a uma filtragem “branqueadora”, utilizando um filtro predictor, dada a elevada correlação existente entre posições vizinhas na imagem [14]. Com efeito, a diferença entre a imagem original e o resultado da filtragem por um

simples filtro predictor de 1ª ordem (i.e., considerando a vizinhança-4 de um ponto para efectuar a predição desse ponto) conduz a um erro de predição que aproxima bem o ruído branco e gaussiano.

O filtro utilizado como aproximação ao filtro branqueador é dado por:

$$F = \frac{1}{4} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \quad (3.18)$$

o que é equivalente a subtrair à imagem o resultado da filtragem com um filtro predictor de primeira ordem (3×3) com pesos idênticos e iguais a 1/4.

A imagem filtrada – Y^F – obtém-se efectuando a convolução entre o filtro F e a imagem recebida:

$$Y^F = Y * F. \quad (3.19)$$

Desmodulador multinível

O desmodulador multinível (figura 3.11) consiste em $M/2$ correladores lineares onde, para detecção de cada um dos símbolos que compõem a marca, é realizada a correlação entre Y_i^F (sequência de amostras de Y^F correspondentes ao símbolo a_i) com cada uma das $M/2$ sequências ortogonais utilizadas durante o processo de inserção da marca-de-água. A saída que apresenta maior valor absoluto limita a decisão a um dado símbolo (A^x) ou ao seu antipodal (A^{x*}): se o sinal do valor da correlação for positivo, a decisão é feita a favor do símbolo A^x ; caso contrário, é escolhido o antipodal A^{x*} .

3.5.2 Análise do desempenho

Seguindo a abordagem em [25] e assumindo que o filtro “branqueador” na recepção garante uma boa aproximação ao modelo de canal com ruído branco e gaussiano, a probabilidade de erro de bit na extracção da marca pode ser obtida a partir do conhecimento da média e variância do sinal à saída dos correladores do desmodulador multinível.

O valor da saída do correlador j correspondente à extracção do i -ésimo símbolo da marca, dada a sequência de modulação s^j , pode ser escrito como:

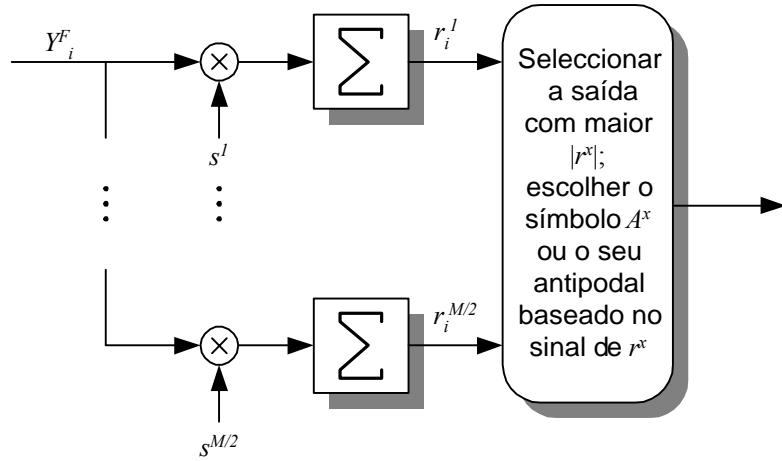


Figura 3.11 – Esquema do desmodulador multinível para o domínio espacial.

$$\begin{aligned}
 r_i | s^j &= \langle Y_F, s^j \rangle = \langle Y * F, s^j \rangle = \\
 &= \langle (X + w) * F, s^j \rangle = \\
 &= \langle X * F + w * F, s^j \rangle = \\
 &= \langle X * F, s^j \rangle + \langle w * F, s^j \rangle = \\
 &= \langle X * F, s^j \rangle + \langle \alpha \cdot s_i * F, s^j \rangle,
 \end{aligned} \tag{3.20}$$

onde \langle , \rangle representa a operação produto interno. Expressando o produto interno como somatório de produtos e escrevendo matematicamente a expressão da convolução, tem-se:

$$\begin{aligned}
 r_i | s^j &= \sum_{(m,n) \in S_i} \left[\left(\sum_{(k,l)} X(m-k, n-l) F(k, l) \right) s^j(m, n) \right] + \\
 &+ \sum_{(m,n) \in S_i} \left[\left(\sum_{(k,l)} \alpha(m-k, n-l) s_i(m-k, n-l) F(k, l) \right) s^j(m, n) \right],
 \end{aligned} \tag{3.21}$$

em que S_i designa o conjunto de posições onde foi inserido o símbolo i da marca. O valor esperado de $r_i | s^j$ é dado por:

$$\begin{aligned}
 E[r_i | s^j] &= E \left[\sum_{(m,n) \in S_i} \sum_{(k,l)} X(m-k, n-l) F(k, l) s^j(m, n) \right] + \\
 &+ E \left[\sum_{(m,n) \in S_i} \sum_{(k,l)} \alpha(m-k, n-l) s_i(m-k, n-l) F(k, l) s^j(m, n) \right].
 \end{aligned} \tag{3.22}$$

Se as sequências de espalhamento tiverem média nula e variância unitária, podem ser efectuadas várias simplificações em (3.22). O primeiro termo em (3.22) anula-se, pois s^j tem média nula.

No segundo termo, o produto $s_i(m-k, n-l)s^j(m, n)$, para $(k, l) = (0, 0)$, tem valor 1, se $s_i = s^j$, -1, se $s_i = -s^j$, e nulo nos restantes casos; para valores de (k, l) diferentes de $(0, 0)$, e tendo em conta a correlação cruzada entre s_i e s^j , o segundo termo anula-se. Atendendo ainda à estrutura do filtro F definido na secção 3.5.1, tem-se $F(0, 0) = 1$. Nestas condições, resulta:

$$E[r_i | s^j] = \begin{cases} \pm \sum_{(m, n) \in S_i} E[\alpha(m, n)], & \text{para } s_i = \pm s^j \\ 0 & , \text{ para } s_i \neq \pm s^j \end{cases} \quad (3.23)$$

Para a variância, tem-se:

$$\begin{aligned} Var[r_i | s^j] &= E\left[\left(r_i | s^j\right)^2\right] - E^2[r_i | s^j] \\ &= E\left[\left(\sum_{(m, n) \in S_i} \sum_{(k, l)} X(m-k, n-l)F(k, l)s^j(m, n) + \right.\right. \\ &\quad \left.\left. + \sum_{(m, n) \in S_i} \sum_{(k, l)} E[\alpha(m-k, n-l)s_i(m-k, n-l)F(k, l)s^j(m, n)]\right)^2\right] - \\ &\quad - \left(\sum_{(m, n) \in S_i} E[\alpha(m, n)]\right)^2. \end{aligned} \quad (3.24)$$

Representando por X_F o resultado da convolução de X com o filtro F e atendendo às características das sequências de espalhamento e à estrutura do filtro F , (3.24) simplifica-se para:

$$\begin{aligned} Var[r_i | s^j] &= \sum_{(m, n) \in S_i} E[X_F^2(m, n)] + \sum_{(m, n) \in S_i} E[\alpha^2(m, n)] + \\ &\quad + \sum_{(m, n) \in S_i} \sum_{\substack{(k, l) \neq (0, 0) \\ (m-k, n-l) \in S_i}} E[\alpha^2(m-k, n-l)]F^2(k, l) + \\ &\quad + \sum_{(m, n) \in S_i} \sum_{\substack{(k, l) \neq (0, 0) \\ (m-k, n-l) \in S_i}} E[\alpha(m, n)\alpha(m-k, n-l)]F^2(k, l) - \left(\sum_{(m, n) \in S_i} E[\alpha(m, n)]\right)^2. \end{aligned} \quad (3.25)$$

As expressões (3.23) e (3.25) representam, respectivamente, a média e variância de r_i condicionadas ao conjunto de pontos S_i . A média e variância de r_i podem ser obtidas recorrendo às propriedades do valor esperado e ao teorema da variância total incondicional [32]:

$$E[r_i] = E_S[E[r_i | s^j]], \quad (3.26)$$

$$Var[r_i] = E_S[Var[r_i|s^j]] + Var_S[E[r_i|s^j]], \quad (3.27)$$

em que $E_S[x]$ e $Var_S[x]$ são, respectivamente, o valor esperado e variância da variável aleatória x no conjunto $S = \bigcup_i S_i$. Aplicando estas propriedades, obtém-se:

$$\begin{aligned} E[r_i] &= \pm \frac{D}{N_s} \sum_{(m,n)} E[\alpha(m,n)], \\ Var[r_i] &= \frac{D}{N_s} \sum_{(m,n)} E[X_F^2(m,n)] + \frac{N_s D - D^2}{N_s^2} \sum_{(m,n)} E[\alpha^2(m,n)] + \\ &\quad + \frac{D^2}{N_s^2} \sum_{(m,n)(k,l) \neq (0,0)} E[\alpha^2(m-k, n-l)] F^2(k,l) + \\ &\quad + \frac{D^2}{N_s^2} \sum_{(m,n)(k,l) \neq (0,0)} E[\alpha(m,n) \alpha(m-k, n-l)] F^2(k,l), \end{aligned} \quad (3.29)$$

em que D representa a densidade de inserção da marca-de-água – quociente entre o número de posições de imagem marcadas e o número total de posições de imagem – e N_s representa o número de símbolos inseridos. Definindo μ e σ^2 como o valor esperado e a variância de r_i , respectivamente, e considerando que $D/N_s \gg D^2/N_s^2$, estes valores podem ser aproximados por:

$$\mu = E[r_i] \approx \frac{DHV}{N_s} E[\alpha(m,n)], \quad (3.30)$$

$$\sigma^2 = Var[r_i] \approx \frac{DHV}{N_s} (E[X_F^2(m,n)] + E[\alpha^2(m,n)]), \quad (3.31)$$

onde H, V são, respectivamente, as dimensões horizontal e vertical da imagem (em *pixels*).

Para analisar o desempenho da sinalização *M-ária* bi-ortogonal, começa-se por derivar a expressão da probabilidade de ser detectado um símbolo correcto [34]. Uma vez que esta análise é semelhante para qualquer símbolo transmitido, vai-se assumir que foi transmitido o símbolo A_1 (com sequência de modulação correspondente s^1). A probabilidade – P_S – de uma decisão correcta é dada por [34]:

$$P_S = \int_0^{+\infty} P\left(r^1 > |r^2|, r^1 > |r^3|, \dots, r^1 > |r^{M/2}| \middle| r^1 > 0\right) p(r^1) dr^1, \quad (3.32)$$

onde $p(r^1)$ é a função densidade de probabilidade da variável aleatória r^1 . Assumindo que $p(r^1)$ pode ser aproximada por uma distribuição normal com média e variância definidas pelas

expressões (3.30) e (3.31) e que as densidades de probabilidade – $p(r^m)$ – para as restantes $M/2-1$ correlações r^m são também normais com a mesma variância, mas com média nula, tem-se:

$$p(r^1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r^1-\mu)^2}{2\sigma^2}} \quad (3.33)$$

$$p(r^m) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{r^{m2}}{2\sigma^2}} \quad (3.34)$$

pelo que:

$$P\left(r^1 > |r^m| \middle| r^1 > 0\right) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-r^1}^{r^1} e^{-\frac{r^{m2}}{2\sigma^2}} dr^m = \frac{1}{\sqrt{2\pi}} \int_{-\frac{r^1}{\sigma}}^{\frac{r^1}{\sigma}} e^{-\frac{x^2}{2}} dx. \quad (3.35)$$

Assumindo que as variáveis aleatórias r^m são independentes, a probabilidade conjunta de $r^1 > |r^m|$ dado que $r^1 > 0$ para $m=2,3\dots M/2$ é dada pelo resultado obtido em (3.35) elevado à potência $M/2-1$ [34]. Deste modo, a expressão (3.32) pode ser escrita como:

$$P_S = \int_0^{+\infty} p(r^1) \cdot \left(\frac{1}{\sqrt{2\pi}} \int_{-\frac{r^1}{\sigma}}^{\frac{r^1}{\sigma}} e^{-\frac{x^2}{2}} dx \right)^{\frac{M}{2}-1} dr^1. \quad (3.36)$$

Substituindo $p(r^1)$ pela expressão (3.33), mudando a variável de integração e manipulando algebricamente a expressão, chega-se a:

$$P_S = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{v^2}{2}} \cdot \text{erf}\left(\frac{1}{\sqrt{2}}\left(v + \frac{\mu}{\sigma}\right)\right)^{\frac{M}{2}-1} dv. \quad (3.37)$$

A probabilidade de erro de símbolo – P_M – é

$$P_M = 1 - P_S = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{v^2}{2}} \cdot \text{erf}\left(\frac{1}{\sqrt{2}}\left(v + \frac{\mu}{\sigma}\right)\right)^{\frac{M}{2}-1} dv. \quad (3.38)$$

No caso particular da utilização de sinalização binária antipodal ($M=2$), a equação (3.38) simplifica-se para:

$$P_b = 1 - P_S = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{v^2}{2}} dv = Q\left(\frac{\mu}{\sigma}\right), \quad (3.39)$$

resultado conhecido para a expressão de probabilidade de erro de bit em canais AWGN quando é utilizada sinalização binária antipodal.

Para a dedução da probabilidade de erro de bit P_b a partir da probabilidade de erro de símbolo P_M , para $M > 2$, deverá ter-se em conta as seguintes situações de erro, que podem ocorrer quando é transmitido um símbolo A^x qualquer:

- Detecta-se o símbolo antipodal de A^x . Neste caso, todos os bits do símbolo em questão são recebidos com erro. Esta situação é a mais rara, e tanto mais quanto maior for o número de níveis M ;
- A decisão é feita erradamente a favor de um dos $M-2$ símbolos restantes. Esta será a situação de erro mais comum ao utilizar-se sinalização bi-ortogonal.

Tendo em conta estas duas situações, pode-se mostrar [34] que a probabilidade de erro de bit é limitada por:

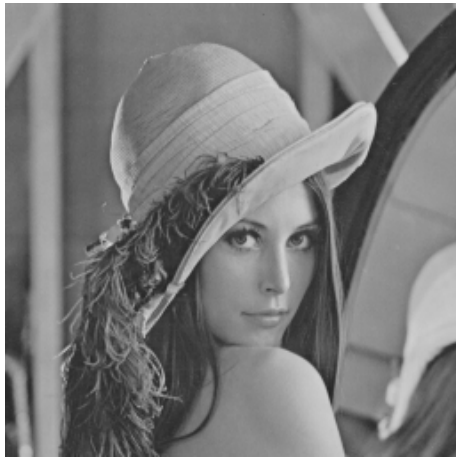
$$\frac{P_M}{2} < P_b \leq P_M, \quad (3.40)$$

e que à medida que o número de níveis aumenta, a probabilidade P_b vai-se aproximando do limite inferior. Assim, para M suficientemente grande, tem-se:

$$P_b \approx \frac{P_M}{2}. \quad (3.41)$$

3.5.3 Resultados

De modo a avaliar experimentalmente o desempenho da sinalização multinível efectuaram-se vários testes, utilizando nas simulações as imagens apresentadas na figura 3.12. Estas imagens foram escolhidas com base nas suas diferenças em termos de actividade espacial, nomeadamente: maior frequência de zonas texturadas na imagem *Mandrill*, zonas essencialmente homogéneas na imagem *02*, representando *Lena* uma imagem natural típica em termos de textura. Os resultados apresentados referem-se às curvas da probabilidade de erro de bit, obtidas para diversos valores do número de níveis utilizados na sinalização (M) e em função do número de pixels da imagem utilizados para inserir um bit da marca. Estas curvas foram inicialmente obtidas utilizando as expressões teóricas derivados na secção anterior, tendo sido posteriormente confirmadas com recurso a simulações experimentais.



Lena (512 × 512 pixels)



Mandrill (520 × 480 pixels)



02 (768 × 512 pixels)

Figura 3.12 – Imagens utilizadas nas simulações.

Para finalizar, apresentam-se os resultados experimentais do desempenho do sistema em presença de compressão *JPEG*, um tipo de compressão que introduz uma degradação na imagem difícil de modelizar teoricamente. As curvas experimentais obtidas com compressão representam a probabilidade de erro de bit em função do factor de qualidade da compressão²¹ *JPEG* (*Q*), o que é usual em estudos deste tipo, na área da assinatura digital de imagens. O factor de qualidade é tipicamente expresso em percentagem (0 a 100%). De notar que, ao ser realizada a compressão *JPEG*, são sempre introduzidas perdas na imagem, mesmo que o factor de qualidade seja de 100%.

²¹ Quanto maior for o factor de qualidade menor é o factor de compressão e vice-versa.

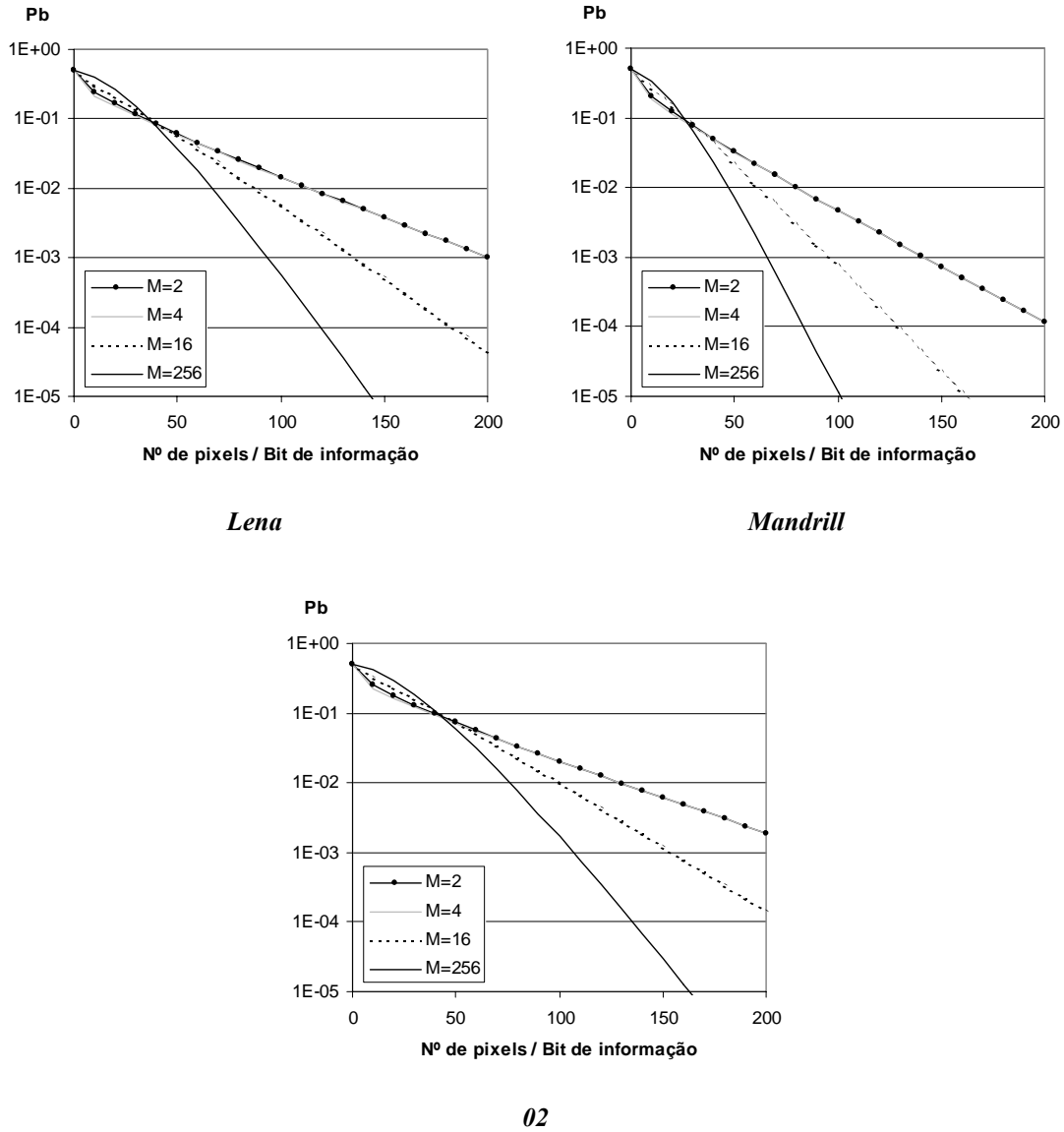


Figura 3.13 – Resultados teóricos – P_b vs. N° de Pixels / Bit de informação.

Resultados teóricos

Na figura 3.13 podem-se observar as curvas teóricas de P_b obtidas para 2,4,16 e 256 níveis de sinalização, em função do número de posições do espaço da marca-de-água utilizados para a transmissão de um bit da marca. Esta grandeza, equivalente ao factor de expansão da banda (*pulse size* ou *chip rate*), pode ser obtida a partir da densidade de inserção – D – das dimensões horizontal e vertical da imagem – H, V – e do número de bits de informação útil da marca-de-água – N_b – de acordo com:

$$\text{nº de posições de inserção por bit de informação} = \frac{DHV}{N_b}. \quad (3.42)$$

O parâmetro que regula a força de inserção da marca-de-água (β) foi fixado no valor 0.4 em todos os testes.

Como se pode observar, os resultados teóricos obtidos para cada imagem são semelhantes: em todas as situações verifica-se que um aumento no número de níveis utilizados na sinalização conduz a um melhor desempenho na extracção dos bits que compõem a marca-de-água. Em todos os casos, a curva da probabilidade de erro de bit decresce mais rapidamente à medida que se aumenta o número de níveis utilizados na sinalização. A única excepção é a utilização de 4 níveis de sinalização, para a qual a evolução da curva de P_b é idêntica à do caso em que se utiliza sinalização binária ($M=2$).

Para valores muito baixos do número de pixels por bit de informação, a utilização de sinalização multinível não é vantajosa. Com efeito, e independentemente do número de níveis usados, a relação μ / σ é muito baixa, conduzindo a probabilidades de erro de bit superiores a 0.1, situação sem qualquer interesse prático.

O ponto de intersecção das diversas curvas, relativas a cada imagem, corresponde ao limite de Shannon [38]. Este limite estabelece que, para uma dada capacidade do canal, e com banda de transmissão infinita, é possível obter uma probabilidade de erro de bit arbitrariamente pequena na recepção, desde que o valor da relação sinal-ruído por bit seja superior a -1.6 dB. Como poderá ser observado no anexo A, o valor do *nº de pixels por bit de informação* em que ocorre o cruzamento entre as curvas corresponde a um valor da relação sinal-ruído próximo de -1.6 dB, independentemente do tipo de imagem.

Resultados experimentais

Com a intenção de se comprovar os resultados previstos teoricamente, foram realizadas simulações experimentais utilizando o método de *Monte Carlo*. Cada ponto de cada curva experimental resulta de 1000 testes se $M=16$ ou 256, e 250 testes se $M=2$ ou 4. Esta diferença justifica-se pelo facto de serem necessárias mais amostras para se obterem os pontos correspondentes aos valores mais baixos de P_b , para os casos $M=16$ e $M=256$. O valor de β é idêntico ao utilizado nas curvas teóricas.

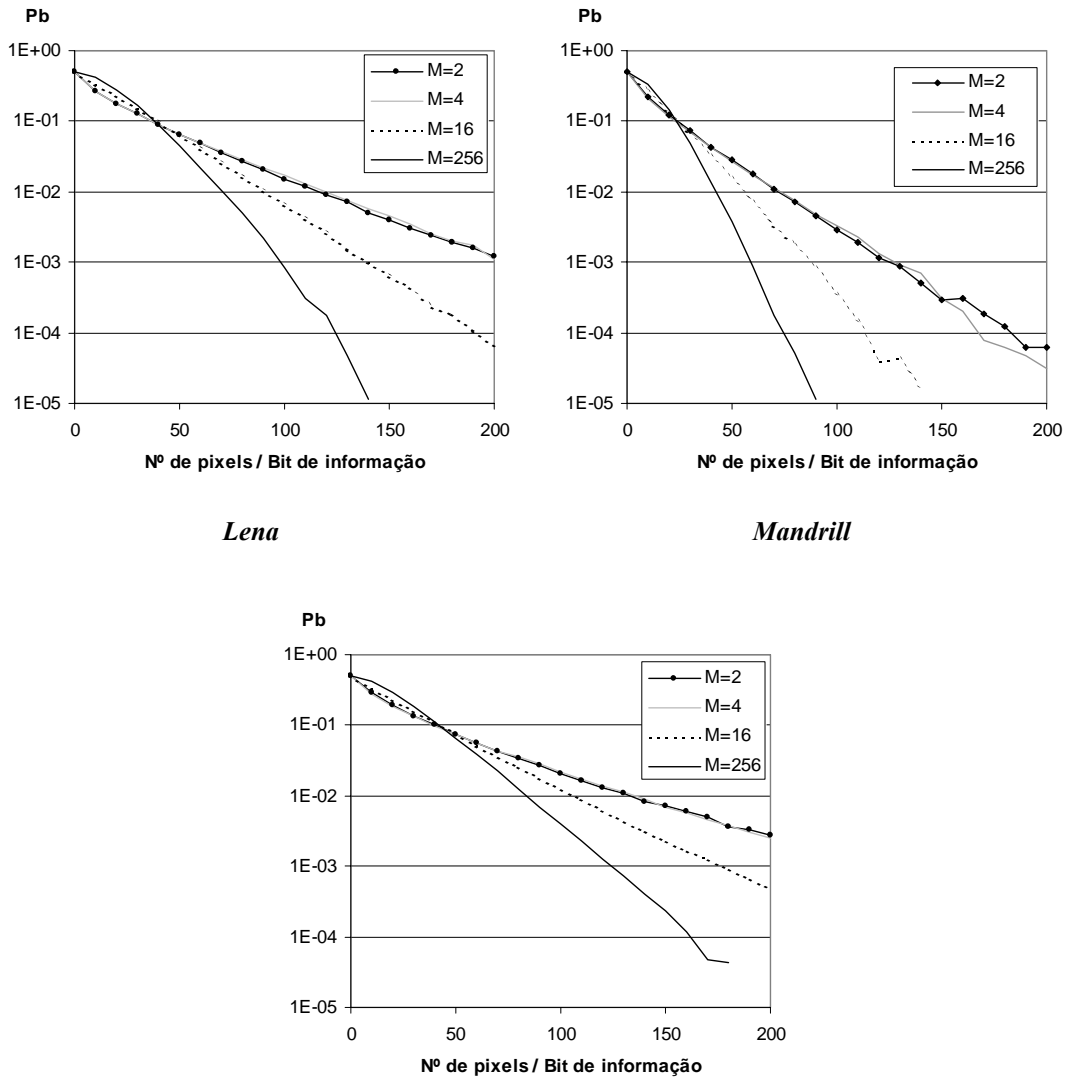


Figura 3.14 – Resultados experimentais – P_b vs. N° de Pixels / Bit de informação.

O conjunto de gráficos apresentados na figura 3.14 representa a evolução experimental de P_b em função do número de pixels utilizados para transmitir um bit da marca. Como se pode observar, os resultados obtidos experimentalmente são bastante próximos dos previstos teoricamente, o que confirma a validade do desenvolvimento analítico realizado na secção 3.5.2 para obtenção de curvas de probabilidade de erro de bit.

Verifica-se também a existência de alguma instabilidade nos resultados experimentais, para valores mais baixos de P_b ($<10^{-4}$), devido ao número de testes ser insuficiente para esta gama de probabilidades. Com efeito, numa simulação *Monte Carlo*, para ser atingida uma probabilidade de erro de bit P_b com um erro relativo de 10%, é necessário que o número de bits em teste seja de pelo menos $100/P_b$ [23,39]. Nesta experiência o número de testes foi de 1000 (no máximo)

com 256 bits por teste, o que conduz a 256 000 bit inseridos. Nestas condições, resulta uma probabilidade de erro de bit mínima de cerca de 4×10^{-4} , com um erro relativo de 10%.

Resultados experimentais em presença de compressão JPEG

De forma a verificar-se o comportamento do sistema em presença de compressão JPEG, realizou-se um novo conjunto de testes equivalentes aos descritos na secção anterior. Em cada teste foi gerada aleatoriamente uma marca-de-água com por 64 bits²² de comprimento e as forças de inserção da marca foram ajustadas para cada imagem: 0.4 na imagem *Lena*, 0.2 na imagem *02* e 0.1 na imagem *Mandrill*. Após a inserção da marca, a imagem marcada foi comprimida com diversos valores do factor de qualidade (Q). Os gráficos da figura 3.15 representam a probabilidade de erro de bit resultante para a marca extraída.

Verifica-se que quanto maior for M , mais baixo é o valor de Q a partir do qual se deixa de registar a ocorrência de erros de bit nas simulações. À semelhança dos casos anteriores, também em presença de compressão JPEG o ponto de cruzamento das diversas curvas ocorre para uma probabilidade de erro de bit de aproximadamente 0.1.

Para a imagem *Lena*, com $\beta=0.4$ e $M=256$, não ocorrem erros a partir de $Q = 45\%$, factor de qualidade que conduziu a uma taxa de compressão de 1:40. Os testes realizados sobre as imagens *Mandrill* e *02* apresentam desempenhos inferiores aos obtidos com *Lena*, o que se justifica pela menor força de inserção usada nestas imagens. Com efeito, mantendo $\beta=0.4$, não se registaram erros sobre *Mandrill* (para $Q>20\%$) e o número de erros sobre a imagem *02* foi pouco significativo.

Resultados experimentais em presença de ruído branco e gaussiano

Outro teste foi realizado com a finalidade de testar a robustez na extracção da marca-de-água, quando a imagem marcada se encontra corrompida por ruído branco e gaussiano de média nula e variância σ_r^2 . Na figura 3.16 pode observar-se o efeito da corrupção por ruído gaussiano, com $\sigma_r = 10$, sobre a imagem *Lena*.

²² Optou-se por um comprimento de 64 bits por ser este o *payload* exigido pela maioria das aplicações de marcas-de-água. Para além disso, o identificador ISAN terá também este comprimento, sendo exigida robustez a ataques não intencionais resultantes da distribuição e difusão de imagens e vídeo, como é o caso da compressão.

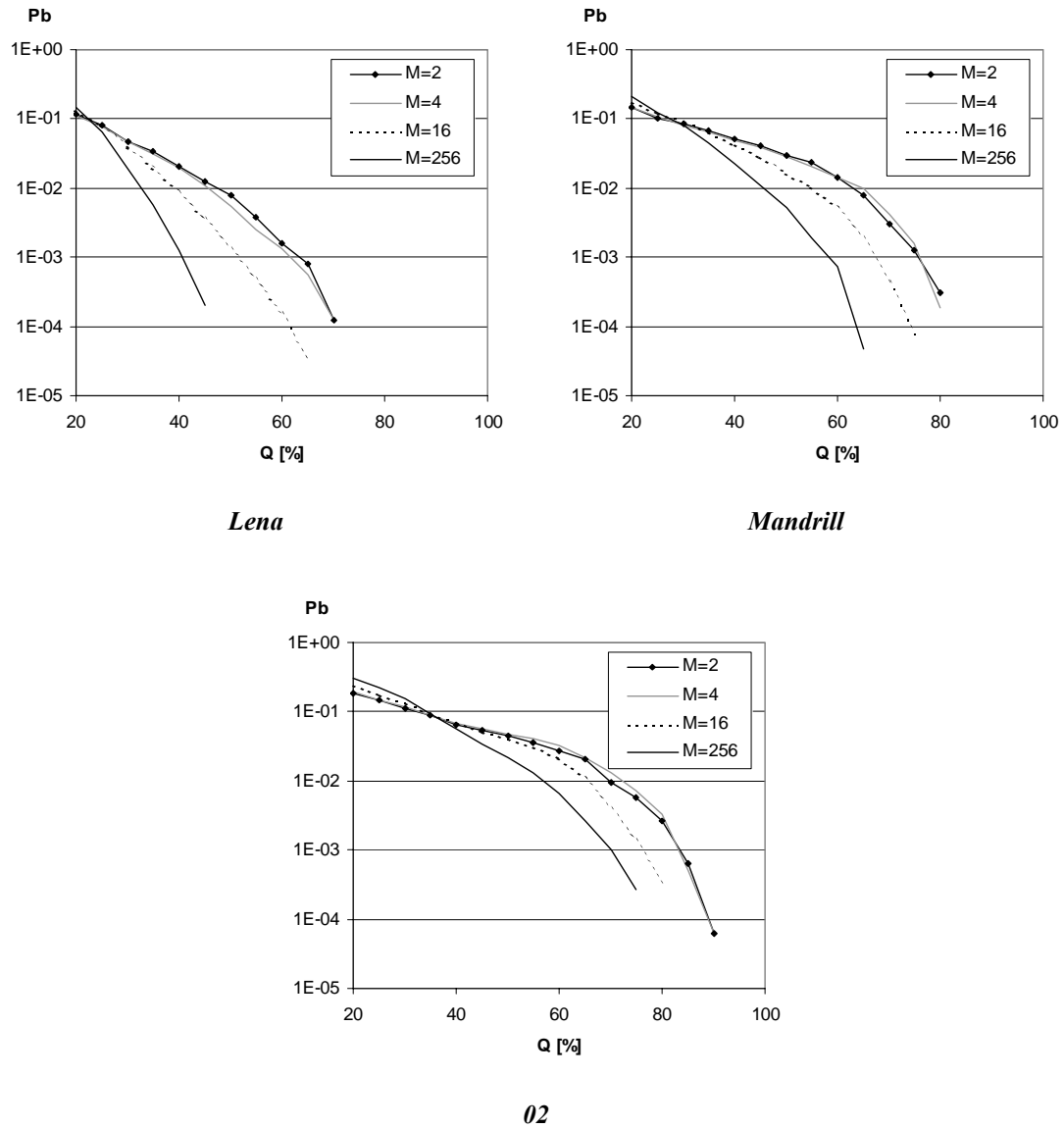


Figura 3.15 – Resultados experimentais – P_b em presença de compressão JPEG.



Figura 3.16 – Ruído gaussiano na imagem *Lena*:

- a) Imagem marcada; b) Imagem marcada corrompida por ruído gaussiano com $\sigma_r = 10$;
 c) Módulo da diferença entre (a) e (b) multiplicada por 8.

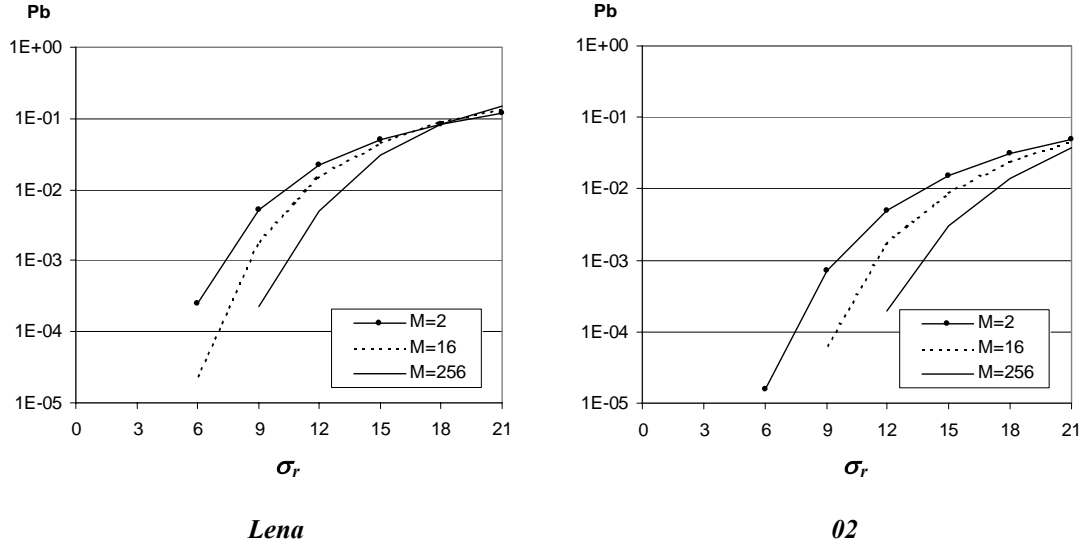


Figura 3.17 – Resultados experimentais – P_b em presença de ruído branco gaussiano.

Obtiveram-se curvas da probabilidade de erro de bit em função de σ_r – com $\sigma_r \in \{3, 6, \dots, 21\}$ – para $M=2, 16$ e 256 . As marcas-de-água utilizadas têm 256 bits de comprimento e a força de inserção foi ajustada para 0.4. O número de testes realizados para obter cada ponto de cada curva foi de 250 para o caso em que $M=2$ e 500 para os restantes casos. Na figura 3.17 apresentam-se os resultados obtidos²³.

Como pode ser observado, também em presença de corrupção por ruído gaussiano se verifica que a modulação multinível conduz a um melhor desempenho na extracção da marca. Quanto maior o número de níveis utilizados na modulação, maior o valor de σ_r a partir do qual se registam erros de bit na extracção da marca-de-água.

Resultados experimentais na presença de cortes (*crop*)

Para finalizar, realizaram-se testes de modo a avaliar o comportamento do sistema quando a imagem marcada sofre um corte. Neste estudo, foram considerados cortes quadrangulares com uma largura de L_{Crop} pixels e com centros coincidentes com o centro da imagem. Na figura 3.18 apresenta-se um exemplo de um corte deste tipo sobre a imagem *Lena*, com $L_{crop} = 320$. O efeito do corte sobre uma imagem marcada traduz-se pela perda de pontos que transportam informação referente à marca-de-água e na consequente redução de energia da marca.

²³ No intervalo de σ_r considerado, não foram registados erros de bit nos testes referentes à imagem *Mandrill*.



Figura 3.18 – Corte sobre a imagem *Lena* (marcada) com $L_{Crop} = 320$.

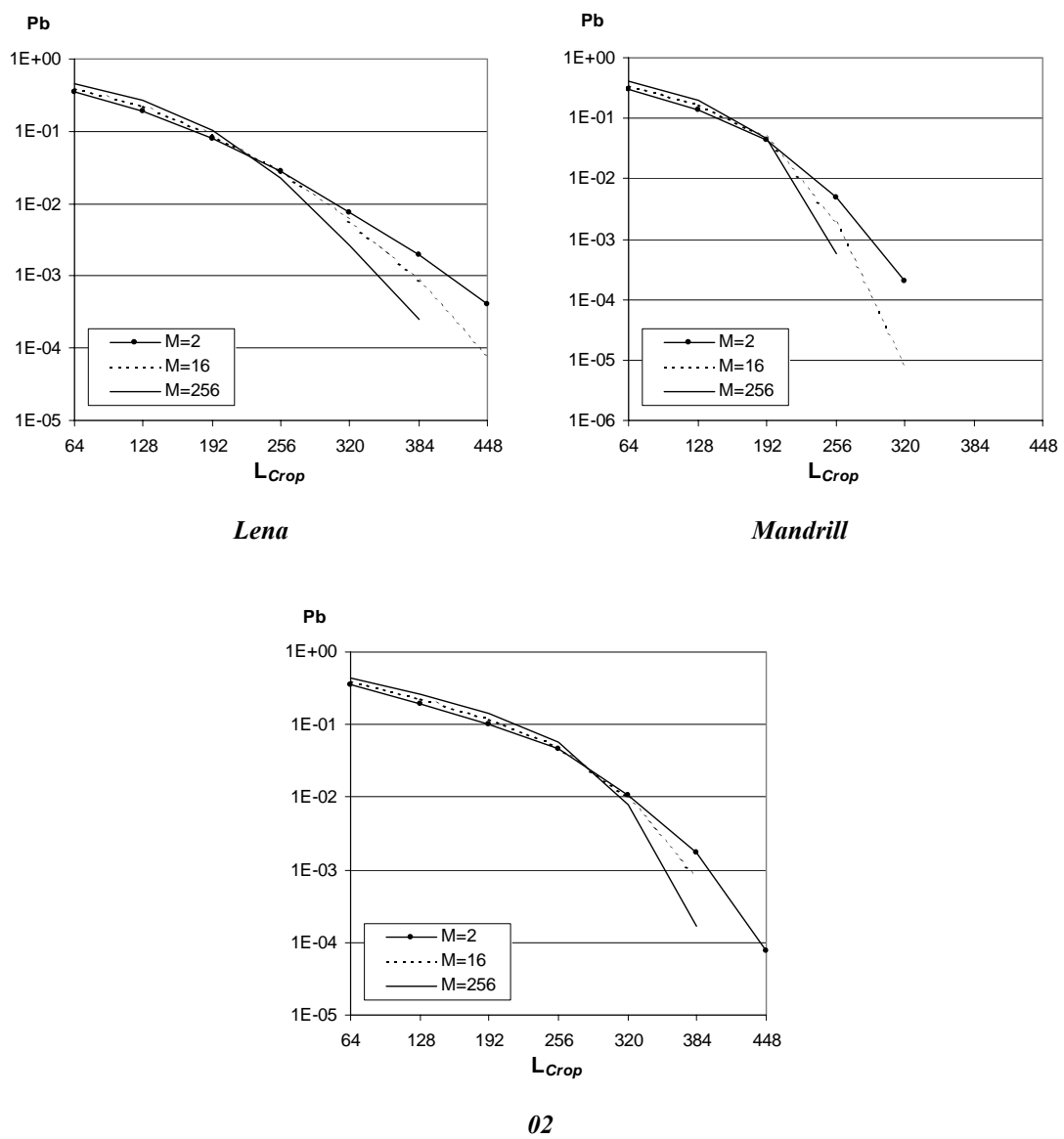


Figura 3.19 – Resultados experimentais – P_b em presença de cortes.

Na figura 3.19 apresentam-se os resultados obtidos para a probabilidade de erro de bit na extracção, em função da largura da janela de corte – L_{Crop} . As condições de teste utilizadas são idênticas às descritas anteriormente em presença de ruído branco e gaussiano. Mais uma vez se verifica que o desempenho melhora com o aumento de M .

De acordo com os resultados obtidos e para $\beta=0.4$, $M=16$ e $N_b=256$, o sistema em análise é robusto a cortes até 25% da imagem *Lena* e 60% sobre as imagens *Mandrill* e *02*, sendo de esperar percentagens de corte mais elevadas às indicadas aumentando β e/ou M , ou reduzindo N_b .

3.6 Marcas-de-água no domínio da frequência

Como referido atrás, existem vários tipos de transformação do domínio espacial para o domínio da frequência. Todas estas transformações têm em comum o facto do espaço resultante da transformação ser também um espaço bidimensional, constituído por diferentes componentes de frequência. Deste modo, o método de inserção atrás estudado pode ser facilmente extendido a este domínio.

Neste estudo optou-se pela utilização de um esquema de marcas-de-água que utiliza a transformada DCT, aplicada a blocos de imagem com dimensões 8×8 pixels, segundo os esquemas de inserção / extracção da marca-de-água propostos em [18]. Esta opção justifica-se por este tipo de transformada ser amplamente utilizada em técnicas de compressão de imagem, tanto em imagem fixa (norma JPEG), como em vídeo (normas MPEG-1, MPEG-2 e MPEG-4). Assim sendo, é possível tirar partido da forma como é feita a compressão nestas técnicas, de modo a tornar o sistema de marcas-de-água mais robusto à distorção (ou “ataque não intencional”) causado pela compressão.

3.6.1 Inserção e extracção da marca-de-água no domínio da frequência

A figura 3.20 apresenta as alterações introduzidas nos esquemas gerais de inserção e extracção da marca-de-água descritos anteriormente (secções 3.3 e 3.4), de modo a que o espaço da marca-de-água seja o espaço da transformada DCT.

Em relação aos esquemas gerais apresentados anteriormente, há a realçar o facto de a transformação T_2 representar duas operações consecutivas: o cálculo da componente de luminância da imagem e a transformada DCT, orientada ao bloco de dimensão 8×8 pixels, desta componente. Desta transformada resultam, por bloco, 64 coeficientes – $C(i,j)$ com $i,j \in \{0, \dots, 7\}$.

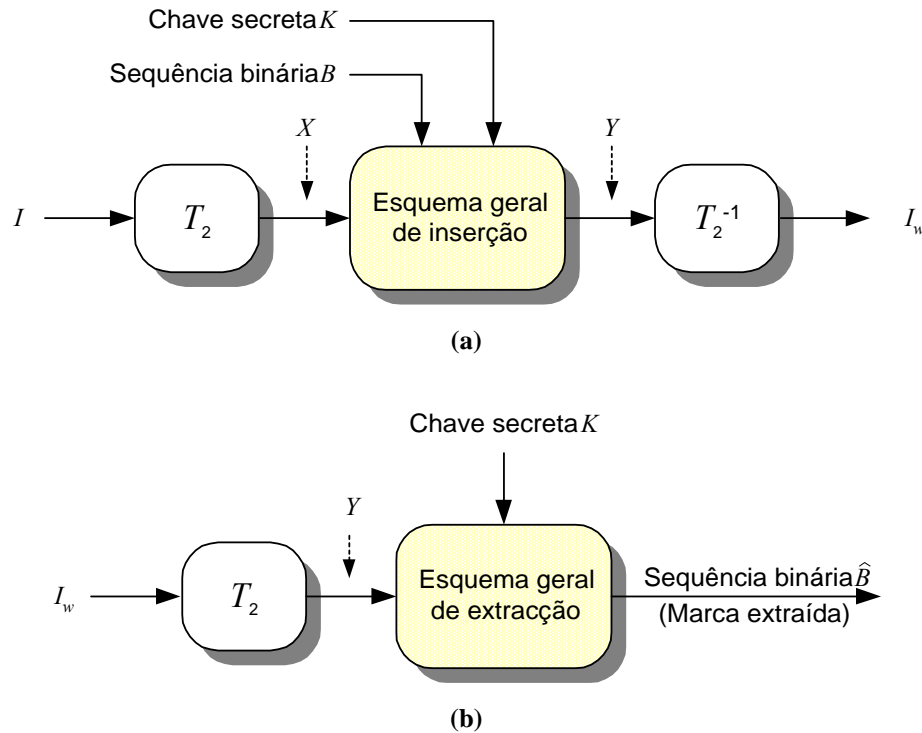


Figura 3.20 – Esquemas de inserção e extração de marcas-de-água no domínio da frequência:
a) Inserção; b) Extração.

Após inserção da marca nos coeficientes DCT, aplica-se transformada DCT inversa, obtendo-se a luminância da imagem marcada. Para a extração da marca, é necessário aplicar as operações inversas.

Resta acrescentar que na implementação utilizada não se utilizaram todos os coeficientes DCT para inserção da marca. A selecção dos coeficientes a usar é usualmente baseada nos seguintes critérios:

- O sistema visual humano é menos sensível às altas frequências espaciais, o que sugere a utilização dos coeficientes correspondentes a essas frequências para a introdução da marca;
- As componentes de alta frequência da imagem sofrem uma maior degradação quando em presença de compressão, o que sugere a utilização de coeficientes correspondendo a frequências intermédias e baixas.

O compromisso entre estes dois critérios, conduz à escolha do subconjunto de coeficientes DCT correspondentes à banda de frequências intermédias, representados na figura 3.21.

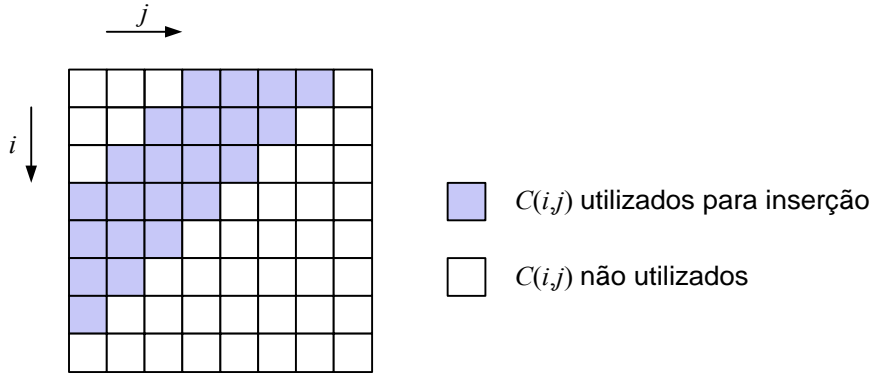


Figura 3.21 – Coeficientes DCT utilizados para inserção da marca.

À semelhança do descrito para o domínio espacial, o espaço Y é obtido adicionando a marca-de-água – w – ao espaço X , i.e.:

$$Y(m, n) = X(m, n) + w(m, n). \quad (3.43)$$

3.6.2 Modelo perceptual para o domínio da frequência

Para maximizar a energia da marca sem degradação visível da qualidade da imagem, é necessário determinar a máxima variação permitida em cada coeficiente DCT, o que pode ser obtido pela utilização de um modelo perceptual adaptado a este domínio. Neste estudo, e à semelhança do que foi feito em [18], utilizou-se o modelo perceptual proposto em [1].

Segundo [1], a máxima alteração possível no (i, j) -ésimo coeficiente DCT de um bloco 8×8 , de modo a ser imperceptível, é dada pelo *limiar de visibilidade* – $V(i, j)$. Este valor, expresso em unidades logarítmicas, pode ser aproximado por:

$$\log V(i, j) = \log \left(\frac{V_{\min} (f_{i,0}^2 + f_{0,j}^2)^2}{(f_{i,0}^2 + f_{0,j}^2)^2 - 4(1-r)f_{i,0}^2 f_{0,j}^2} \right) + K \left(\log \frac{\sqrt{f_{i,0}^2 + f_{0,j}^2}}{f_{\min}} \right)^2, \quad (3.44)$$

em que $f_{i,0}$ e $f_{0,j}$ são, respectivamente, as frequências espaciais (em ciclos por grau – ver anexo B) das funções de base da transformada DCT. V_{\min} é o valor mínimo de $V(i, j)$ associado à frequência espacial f_{\min} , r e K são valores empíricos, tipicamente ajustados para os valores 0.7 e 1.758, respectivamente. Este modelo não é válido para o coeficiente DC – $C(0,0)$.

Em cada bloco de coeficientes DCT, o valor de $V(i, j)$ é ajustado tendo em conta a luminância média do bloco, de acordo com:

$$V'(i, j) = V(i, j) \left(\frac{C(0,0)}{\bar{C}(0,0)} \right)^{\alpha_T}, \quad (3.45)$$

onde $\bar{C}(0,0)$ é o valor médio dos coeficientes DC da imagem e α_T é um parâmetro empírico. Na tabela 3.1 encontram-se sintetizados os valores dos diversos parâmetros utilizados na implementação deste modelo perceptual.

Parâmetro	Valor
K	1.758
V_{\min}	1.1548
f_{\min}	3.68
r	0.7
α_T	0.649

Tabela 3.1 – Parâmetros utilizados no modelo perceptual do domínio da frequência.

Uma vez calculado o limiar de visibilidade ajustado para cada coeficiente, o valor da máscara perceptual é obtido através de:

$$\alpha(m, n) = \frac{1}{16} \cdot \left(1 + (\sqrt{2} - 1) \cdot \delta(i)\right) \cdot \left(1 + (\sqrt{2} - 1) \cdot \delta(j)\right) \cdot \beta \cdot V'(i, j), \quad (3.46)$$

em que i e j são, respectivamente, os restos das divisões inteiras de m e n por 8, $\delta(\cdot)$ é o impulso de Dirac e β o valor que permite ajustar a força com que é inserida a marca-de-água.

3.6.3 Modelo estatístico dos coeficientes DCT

A caracterização da estatística dos coeficientes DCT é essencial para a derivação da estrutura do desmodulador. A distribuição estatística de cada coeficiente DCT é bem aproximada pela função densidade de probabilidade *gaussiana generalizada*, com média nula:

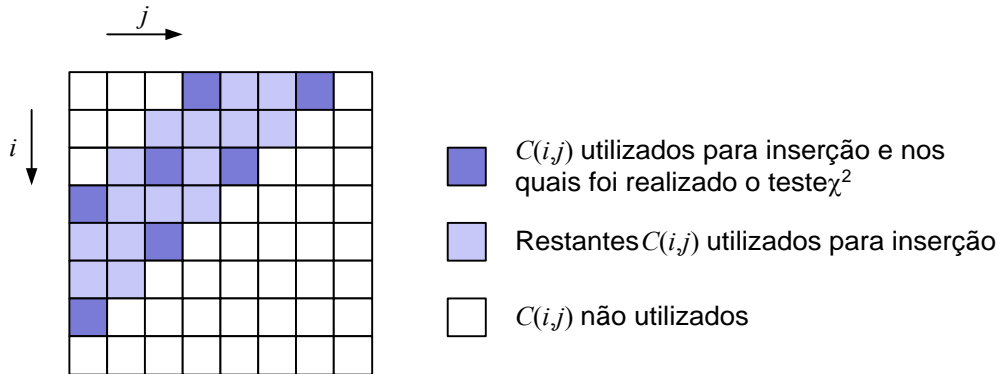
$$f_x(x) = A e^{-|Bx|^c}, \quad (3.47)$$

onde A e B dependem do parâmetro c e do desvio padrão da distribuição – σ – de acordo com:

$$B = \frac{1}{\sigma} \left(\frac{\Gamma\left(\frac{3}{c}\right)}{\Gamma\left(\frac{1}{c}\right)} \right)^{\frac{1}{2}}, \quad A = \frac{Bc}{2\Gamma\left(\frac{1}{c}\right)} \quad (3.48)$$

em que $\Gamma(\cdot)$ denota a função *gama*, definida como:

c	Coeficiente						
	(0,3)	(0,6)	(2,2)	(2,4)	(3,0)	(4,2)	(6,0)
$c=2$ (Gauss)	0.85	0.55	1.27	0.75	0.68	0.38	0.08
$c=1$ (Laplace)	0.32	0.09	0.53	0.23	0.23	0.09	0.02
$c=1/2$	0.09	0.31	0.12	0.21	0.14	0.29	0.73

Tabela 3.2 – Resultados do teste chi-quadrado (χ^2).Figura 3.22 – Coeficientes DCT para os quais foi realizado o teste χ^2 .

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \quad (3.49)$$

As distribuições de Gauss e Laplace são casos particulares da distribuição gaussiana generalizada, com $c=2$ e $c=1$, respectivamente.

De modo a avaliar a validade da distribuição gaussiana generalizada para modelizar a verdadeira distribuição dos valores dos coeficientes DCT, procedeu-se à realização do teste chi-quadrado²⁴, para os casos concretos de distribuições com $c=2$ (Gauss), $c=1$ (Laplace) e $c=1/2$. Na tabela 3.2 apresentam-se os resultados deste teste quando aplicado aos coeficientes DCT representados na figura 3.22 e para o caso particular da imagem *Lena*.

De entre os parâmetros c experimentados, verifica-se que com $c=2$ os resultados não são satisfatórios, pelo que a distribuição de Gauss não é uma boa modelização da distribuição dos coeficientes DCT. Com $c=1$ e $c=1/2$ as aproximações conseguidas pela distribuição generalizada são boas: a distribuição de Laplace aproxima melhor a distribuição dos coeficientes DCT referentes a frequências intermédias e a distribuição generalizada com $c=1/2$ é mais adequada para os coeficientes DCT correspondentes a frequências mais baixas.

²⁴ O teste chi-quadrado permite avaliar quão bem uma distribuição teórica (Gauss, Laplace, Poisson, etc.) se ajusta a uma distribuição empírica. O resultado deste teste é sempre positivo e quanto mais próximo de zero for, melhor a distribuição teórica se ajusta à empírica.

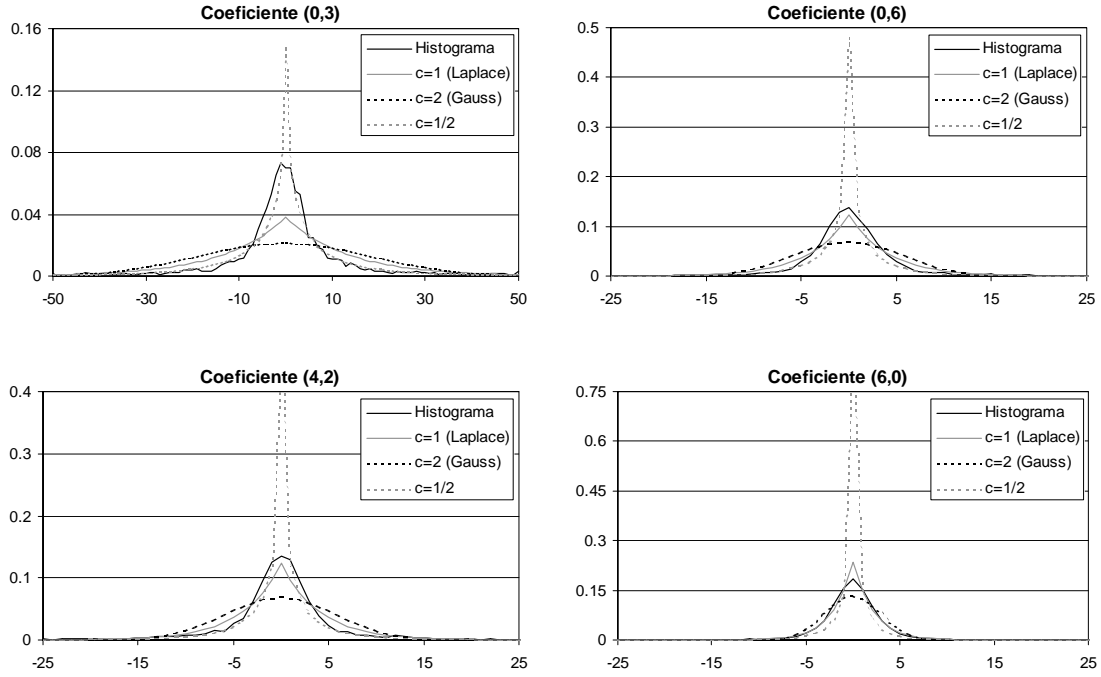


Figura 3.23 – Distribuição estatística de alguns coeficientes DCT – imagem *Lena*.

Para completar este estudo, encontram-se representados na figura 3.23 os histogramas de alguns coeficientes DCT da imagem *Lena*, juntamente com as curvas referentes à distribuição generalizada com $c=2$, $c=1$ e $c=1/2$. Como pode ser observado, as formas dos histogramas não variam significativamente, confirmando serem bem aproximadas pela distribuição generalizada.

3.6.4 Estrutura do desmodulador

Nesta secção, revê-se a estrutura do detector óptimo para o caso binário proposta em [18] e estende-se essa derivação para o caso multinível. De acordo com a equação (3.43) e designando cada posição (m,n) por $[m]$, tem-se:

$$X[m] = Y[m] - w[m]. \quad (3.50)$$

Atendendo a (3.47), a f.d.p.²⁵ conjunta dos valores $X[m]$ (coeficientes DCT), assumindo independência estatística entre coeficientes, é dada por:

$$f(X[m]) = \prod_m A[m] e^{-|B[m]X[m]|^{c[m]}}. \quad (3.51)$$

Após inserção da marca w e atendendo a (3.50) e (3.51), pode-se escrever:

²⁵ f.d.p. – função densidade de probabilidade.

$$f(Y[m]|w[m]) = \prod_{m \in S} A[m] e^{-|B[m](Y[m]-w[m])|^{c[m]}}, \quad (3.52)$$

onde S designa o conjunto de coeficientes onde se inseriu w . Considere-se agora que foi inserida na imagem a marca-de-água w_l . Designando por w_m todas as marcas-de-água diferentes de w_l e admitindo marcas equiprováveis, a decisão sobre a marca inserida que minimiza a probabilidade de erro de bit na extracção, é dada pelo teste de máxima verosimilhança [44], no qual a marca estimada é a que satisfaz:

$$f(Y[m]|w_l[m]) > f(Y[m]|w_m[m]), \quad \forall_{m \neq l}, \quad (3.53)$$

ou, aplicando a função logaritmo à expressão anterior:

$$\log \frac{f(Y[m]|w_l[m])}{f(Y[m]|w_m[m])} > 0, \quad \forall_{m \neq l}. \quad (3.54)$$

Substituindo (3.52) em (3.54) e após algumas manipulações algébricas, obtém-se:

$$\sum_{m \in S} B[m]^{c[m]} (|Y[m] - w_m[m]|^{c[m]} - |Y[m] - w_l[m]|^{c[m]}) > 0. \quad (3.55)$$

Para o caso binário ($M=2$) e modulação por sequências de espalhamento bi-ortogonais, existe apenas uma sequência de espalhamento – s . Designando por $s[m]$ o elemento de s referente à posição $[m]$ tem-se:

$$w_l[m] = \begin{cases} \alpha[m]s[m] & \text{se } b_l = 1 \\ -\alpha[m]s[m] & \text{se } b_l = 0 \end{cases}. \quad (3.56)$$

Designando por S_i o conjunto de coeficientes onde foi inserido o i -ésimo bit da marca (b_i), pode-se demonstrar (Anexo C) que o sinal da saída do detector

$$r_i = \sum_{m \in S_i} B[m]^{c[m]} (|Y[m] + \alpha[m]s[m]|^{c[m]} - |Y[m] - \alpha[m]s[m]|^{c[m]}), \quad (3.57)$$

é suficiente para realizar a decisão sobre o valor desse bit b_i , isto é:

$$\begin{cases} r_i > 0 \Rightarrow b_i = 1 \\ r_i < 0 \Rightarrow b_i = 0 \end{cases}. \quad (3.58)$$

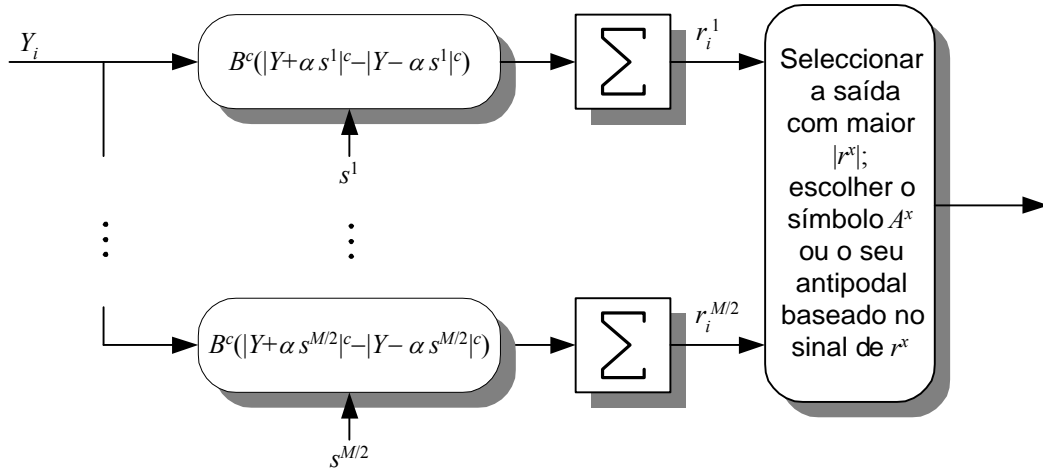


Figura 3.24 – Estrutura do desmodulador para inserção no domínio da frequência.

Pode-se também demonstrar (Anexo C) que a estrutura geral do desmodulador apresentado anteriormente (ver secção 3.4, figura 3.9), é também adequada para o caso multinível, com inserção nos coeficientes DCT. De acordo com essa estrutura, na desmodulação são utilizadas $M/2$ sequências de espalhamento bi-ortogonais e o desmodulador é constituído por $M/2$ detectores, a que se segue a decisão, como ilustrado na figura 3.24. Uma vez que a saída do j -ésimo detector, na extracção do i -ésimo símbolo da marca-de-água, é dada por:

$$r_i^j = \sum_{m \in S_i} B[m]^{c[m]} \left(\left| Y[m] + \alpha[m] s^j[m] \right|^{c[m]} - \left| Y[m] - \alpha[m] s^j[m] \right|^{c[m]} \right), \quad (3.59)$$

verifica-se que apenas um dos detectores apresentará à saída um valor médio não nulo, conduzindo a uma decisão entre um símbolo modulado com sequência de espalhamento $s_i = s^j$ ou um símbolo modulado com sequência de espalhamento $s_i = -s^j$. À semelhança do caso binário (expressão (3.57)), esta decisão é feita com base no sinal de r_i^j . Os restantes $M/2-1$ detectores apresentam nas suas saídas um valor médio nulo, sendo estes valores interpretados como ruído.

3.6.5 Análise do desempenho

Para se proceder à análise do desempenho, e seguindo o proposto em [18], há que começar por se caracterizar a estatística da saída do desmodulador, r_i . Uma vez que r_i resulta da soma de um número elevado de variáveis aleatórias identicamente distribuídas (expressões (3.57) e (3.59)) e atendendo ao teorema do limite central, a f.d.p. de r_i pode ser considerada como gaussiana. Deste modo, a sua estatística é completamente definida pelo valor médio e pela variância.

De acordo com o obtido na secção anterior, tem-se:

$$r_i = \sum_{\mathbf{m} \in S_i} B[\mathbf{m}]^{c[\mathbf{m}]} \left(|Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} \right). \quad (3.60)$$

Fazendo:

$$r[\mathbf{m}] = |Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]}, \quad (3.61)$$

e substituindo em (3.60) obtém-se:

$$r_i = \sum_{\mathbf{m} \in S_i} B[\mathbf{m}]^{c[\mathbf{m}]} r[\mathbf{m}]. \quad (3.62)$$

A média e variância de r_i dado o conjunto de posições S_i referentes ao símbolo i da marca são dadas por:

$$E[r_i | S_i] = \sum_{\mathbf{m} \in S_i} B[\mathbf{m}]^{c[\mathbf{m}]} E[r[\mathbf{m}]], \quad (3.63)$$

$$Var[r_i | S_i] = \sum_{\mathbf{m} \in S_i} B[\mathbf{m}]^{2c[\mathbf{m}]} Var[r[\mathbf{m}]]. \quad (3.64)$$

Tendo em conta que:

$$Y[\mathbf{m}] = X[\mathbf{m}] + w[\mathbf{m}] = X[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}], \quad (3.65)$$

e substituindo (3.65) em (3.61), resulta:

$$r[\mathbf{m}] = |X[\mathbf{m}] + 2\alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} - |X[\mathbf{m}]|^{c[\mathbf{m}]}. \quad (3.66)$$

A média e a variância de $r[\mathbf{m}]$, podem ser obtidas a partir de (3.66), admitindo que $s[\mathbf{m}] \in \{-1, 1\}$ e que $P(s[\mathbf{m}] = -1) = P(s[\mathbf{m}] = 1) = 1/2$. Neste caso resulta, para o valor esperado de $r[\mathbf{m}]$:

$$E[r[\mathbf{m}]] = \frac{1}{2} \left(|X[\mathbf{m}] + 2\alpha[\mathbf{m}]|^{c[\mathbf{m}]} + |X[\mathbf{m}] - 2\alpha[\mathbf{m}]|^{c[\mathbf{m}]} \right) - |X[\mathbf{m}]|^{c[\mathbf{m}]}. \quad (3.67)$$

Para a variância de $r[\mathbf{m}]$ tem-se, após manipulações algébricas simples:

$$Var[r[m]] = \frac{1}{4} \left(|X[m] + 2\alpha[m]|^{c[m]} - |X[m] - 2\alpha[m]|^{c[m]} \right)^2. \quad (3.68)$$

Até este ponto foram obtidas expressões para a média e variância de r_i condicionadas ao conjunto de posições S_i . Para se obter a média e variância de r_i não condicionadas, utilizam-se as propriedades do valor esperado, expressas em (3.26) e (3.27), vindo:

$$E[r_i] = E_S[E[r_i|S_i]] = \frac{D}{N} \sum_m B[m]^{c[m]} E[r[m]], \quad (3.69)$$

$$\begin{aligned} Var[r_i] &= E_S[Var[r_i|S_i]] + Var_S[E[r_i|S_i]] \\ &= \frac{D}{N} \sum_m B[m]^{2c[m]} Var[r[m]] + \frac{ND - D^2}{N^2} \sum_m B[m]^{2c[m]} E^2[r[m]]. \end{aligned} \quad (3.70)$$

Fica assim caracterizada a estatística de r_i . Exprimindo μ e σ^2 como:

$$\mu = E[r_i], \quad (3.71)$$

$$\sigma^2 = Var[r_i], \quad (3.72)$$

a probabilidade de erro de bit na extracção da marca é dada recorrendo às expressões (3.38), (3.39) e (3.41), substituindo μ e σ^2 pelos valores resultados obtidos nesta secção.

3.6.6 Resultados

À semelhança do que foi realizado para o caso de inserção no domínio espacial, efectuaram-se várias experiências com vista à verificação do desempenho da utilização de sinalização multinível, aplicada às marcas-de-água no domínio da frequência. Foram obtidos dois tipos de resultados: curvas de probabilidades de erro de bit, teóricas e experimentais, em função do número de pontos do espaço da marca utilizados para transmissão de um bit de informação da marca; evolução da probabilidade de erro de bit, em presença de compressão JPEG, em função do factor de qualidade da compressão. Antes da realização dos testes conducentes a estes resultados, foram efectuadas algumas simplificações ao modelo do detector analisado na secção anterior.

Simplificações realizadas ao modelo apresentado

Com o objectivo de reduzir a complexidade do esquema de extracção da marca-de-água apresentado, optou-se por simplificar o cálculo de r_i , dado por (3.57). Nesta expressão, o cálculo

de $B[\mathbf{m}]$ é computacionalmente pesado. No entanto, verificou-se experimentalmente que os resultados obtidos são pouco sensíveis ao valor de $B[\mathbf{m}]$ (figura 3.25). Deste modo, fazendo $B[\mathbf{m}]=1$, a expressão de r_i simplifica-se para:

$$r_i = \sum_{\mathbf{m} \in S_i} \left(|Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[\mathbf{m}]} \right). \quad (3.73)$$

Considerando ainda que a distribuição estatística dos coeficientes DCT utilizados para a inserção da marca-de-água é aproximada por uma distribuição de Laplace, i.e., $c[\mathbf{m}]=1$, resulta:

$$r_i = \sum_{\mathbf{m} \in S_i} \left(|Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]| - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]| \right), \quad (3.74)$$

constituindo-se assim o sinal considerado para a desmodulação da marca-de-água.

Este conjunto de simplificações tem reflexos nas expressões da média e variância de r_i , dadas respectivamente por (3.69) e (3.70), obtendo-se agora:

$$\mu = E[r_i] = \frac{D}{N} \sum_{\mathbf{m}} E[r[\mathbf{m}]], \quad (3.75)$$

$$\sigma^2 = Var[r_i] = \frac{D}{N} \sum_{\mathbf{m}} Var[r[\mathbf{m}]] + \frac{ND - D^2}{N^2} \sum_{\mathbf{m}} E^2[r[\mathbf{m}]], \quad (3.76)$$

com

$$r[\mathbf{m}] = |Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]| - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]|. \quad (3.77)$$

Todos os resultados, teóricos ou experimentais, foram obtidos considerando estas simplificações.

Resultados de P_b vs. número de pontos por bit de informação

Na figura 3.26 apresentam-se os resultados da probabilidade de erro de bit obtidos teoricamente e experimentalmente para cada uma das imagens, em função do número de pontos do espaço da marca-de-água utilizados para a inserção de um bit da marca. Os resultados apresentados referem-se à utilização de sinalização multinível com 2,4,16 e 256 níveis de sinalização. O valor da força de inserção (β na expressão (3.46)) foi ajustado para 1.5 em todos os testes.

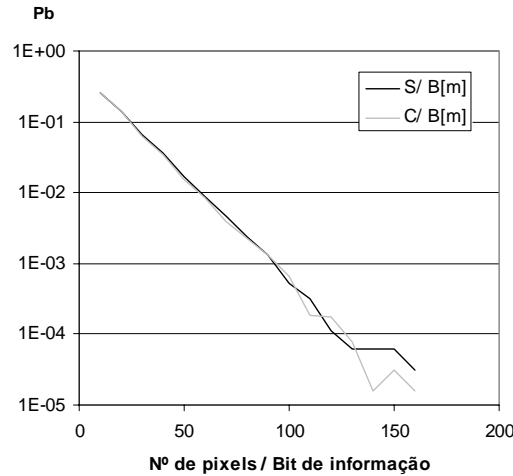


Figura 3.25 – Resultados experimentais para $M=2$, com e sem contabilização de $B[m]$, considerando que os coeficientes DCT seguem uma distribuição de Laplace ($c[m]=1$).

Para obtenção das curvas teóricas foram estimados os parâmetros μ e σ^2 , para cada valor do número de pontos por bit de informação, a partir das expressões (3.75) e (3.76), sendo aplicadas posteriormente as expressões que permitem obter P_b a partir desses parâmetros e do número de níveis utilizados na sinalização. Para realização dos testes experimentais, as condições impostas foram idênticas às referidas na secção 3.5.3, à parte o valor da força de inserção.

Tal como no caso de inserção no domínio espacial, o desempenho do sistema melhora à medida que aumenta o número de níveis utilizados na sinalização. Existe uma certa discrepância entre algumas curvas teóricas e experimentais, nomeadamente nas referentes à imagem *Lena*, nas quais o declive das curvas obtidas experimentalmente é mais acentuado do que o das curvas teóricas correspondentes. De notar, no entanto, que as relações existentes entre as diversas curvas não se altera, ou seja, as conclusões a retirar são as mesmas.

Resultados experimentais em presença de compressão JPEG

Na figura 3.27 encontram-se representados os resultados obtidos experimentalmente para a probabilidade de erro de bit P_b em função do factor de qualidade da compressão. O valor força de inserção relativo aos resultados apresentados é de 1.5. Quanto às restantes condições de teste, são idênticas às descritas na secção 3.5.3.

Para $\beta=1.5$ e $M=256$, verifica-se a inexistência de erros a partir de $Q=50\%$ para as imagens *Lena* e *02*, e a partir de $Q=35\%$ para a imagem *Mandrill*. A estes factores de qualidade, correspondem as taxas de compressão 1:14 (*Lena*), 1:35 (*02*) e 1:19 (*Mandrill*).

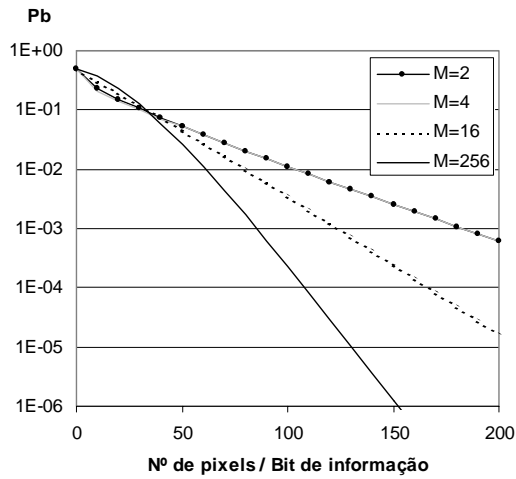
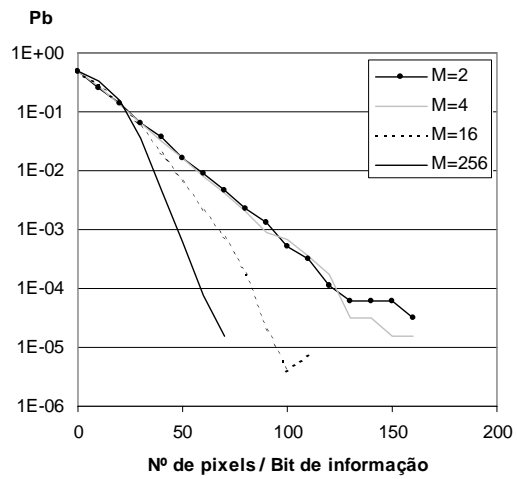
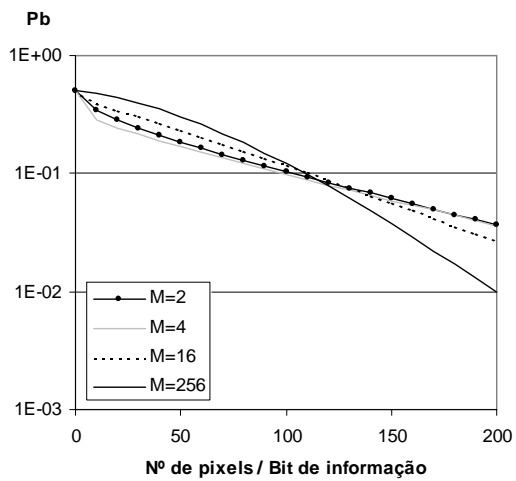
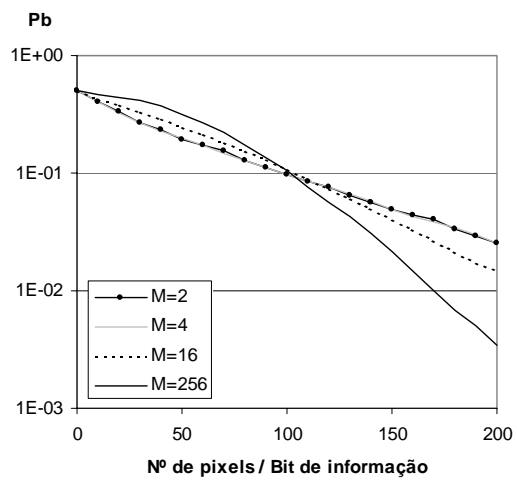
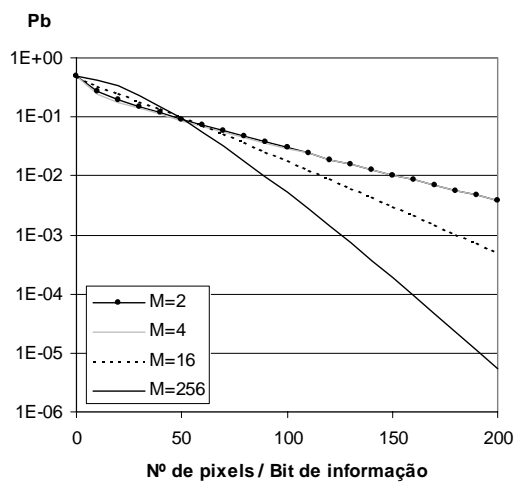
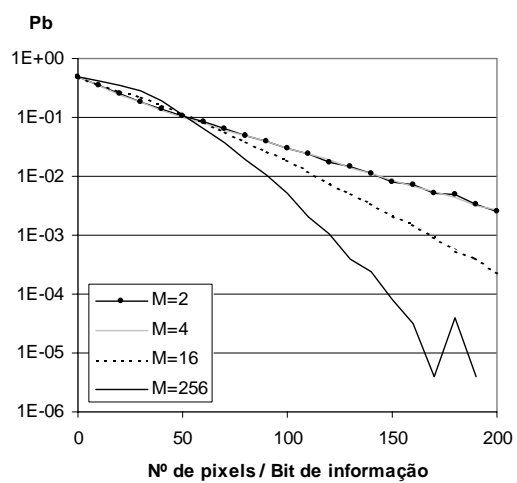
*Lena (Teo.)**Lena (Exp.)**Mandrill (Teo.)**Mandrill (Exp.)**02 (Teo.)**02 (Exp.)*

Figura 3.26 – Resultados teóricos e experimentais – P_b vs. N° de pontos / Bit de informação.

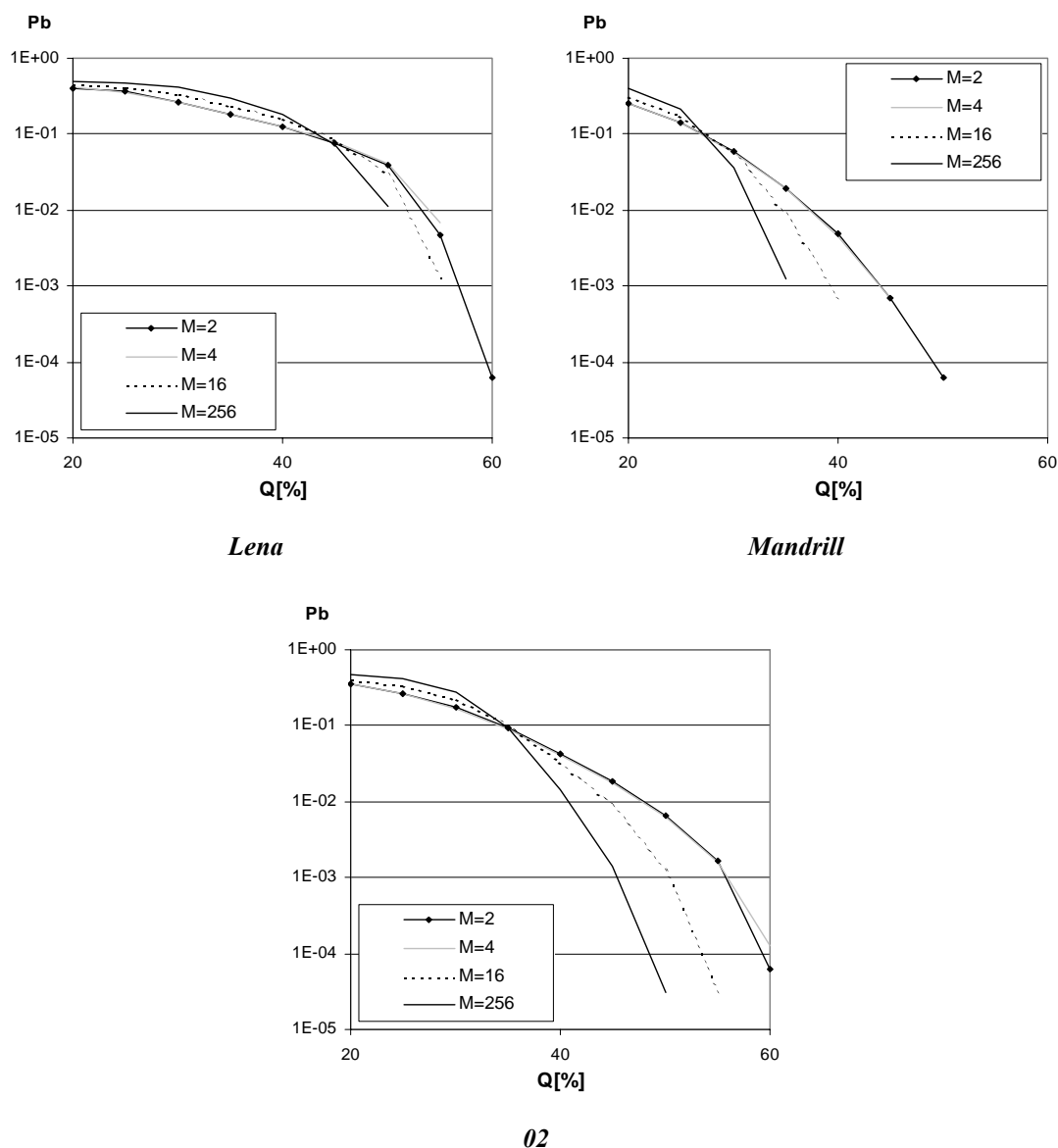


Figura 3.27 – Resultados experimentais – P_b em presença de compressão JPEG.

3.7 Comparação entre resultados referentes aos dois domínios

Pretende-se nesta secção estabelecer uma comparação entre os resultados obtidos nos dois domínios estudados: domínio espacial e domínio da frequência. De modo a poder ser efectuada esta comparação, procedeu-se a um ajuste dos valores das forças de inserção utilizadas em cada um dos domínios, de modo a que a degradação na imagem original causada pela inserção da marca-de-água, seja idêntica. Para se medir a degradação causada pela inserção da marca-de-água, foi utilizada a relação sinal-ruído de pico – PSNR (*Peak Signal to Noise Ratio*) – dada por:

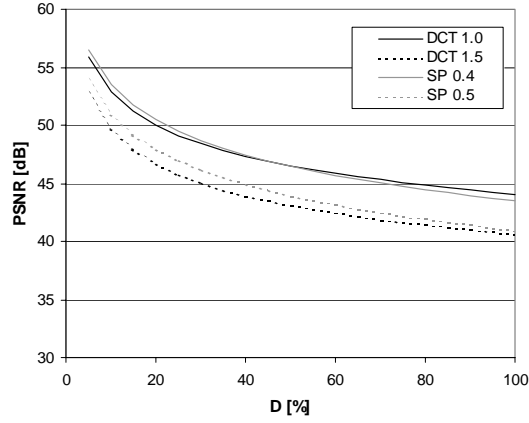


Figura 3.28 – PSNR resultante após inserção da marca – imagem *Lena*.

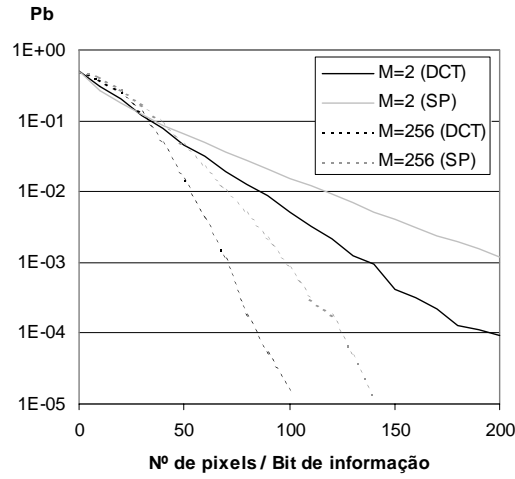


Figura 3.29 – Comparação de resultados nos dois domínios de inserção – imagem *Lena*.

$$PSNR_{[dB]} = 10 \log_{10} \left(HV \cdot \frac{255^2}{\sum_{(m,n)} [X_1(m,n) - X_2(m,n)]^2} \right), \quad (3.78)$$

onde X_1 e X_2 são as componentes de luminância da imagem original e da imagem marcada, respectivamente. Esta medida não é a mais adequada em marcas-de-água, dado que níveis equivalentes de degradação (mesmo PSNR) na mesma imagem, mas causados por dois processamentos diferentes, podem conduzir a resultados perceptuais distintos. Vai-se no entanto assumir que os modelos perceptuais utilizados nos dois domínios em estudo garantem a imperceptibilidade da marca, e que a PSNRs idênticos nos dois domínios correspondem degradações com resultados perceptuais idênticos.

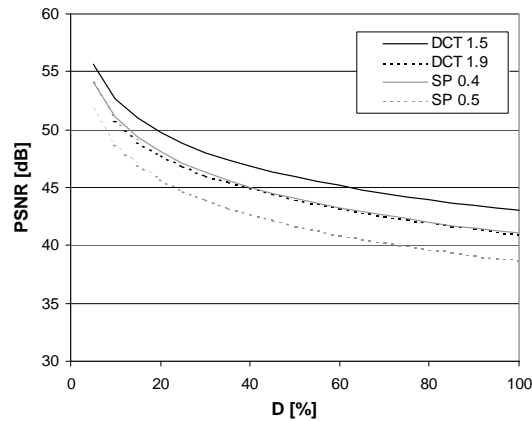


Figura 3.30 – PSNR resultante após inserção da marca – imagem 02.

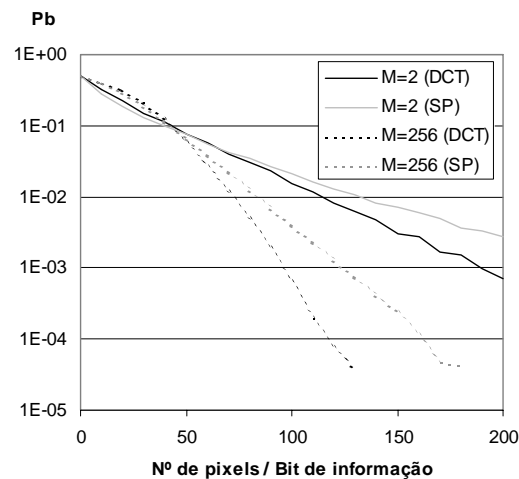


Figura 3.31 – Comparação de resultados nos dois domínios – imagem 02.

Na figura 3.28 representam-se as curvas da PSNR resultante após inserção da marca-de-água, em função da densidade de inserção, para a imagem *Lena*. Por observação deste gráfico, pode concluir-se que, para esta imagem, a inserção no domínio espacial com uma força de inserção de 0.4 é equivalente, em termos de PSNR, à inserção no domínio da frequência com uma força de 1.0. A figura 3.29 apresenta os resultados experimentais obtidos nos dois domínios, com 2 e 256 níveis de sinalização e com as forças de inserção referidas no parágrafo anterior: 0.4 para o domínio espacial e 1.0 para o domínio da frequência.

Foi efectuada uma comparação semelhante para a imagem 02. Por observação da figura 3.30, conclui-se que uma comparação justa entre os resultados dos dois domínios conduz à escolha de uma força de inserção de 0.4 no domínio espacial e de uma força de inserção de 1.9 para no domínio da frequência. Na figura 3.31 encontram-se representadas as curvas de probabilidade de erro de bit com $M=2$, 256 para estas duas forças de inserção.

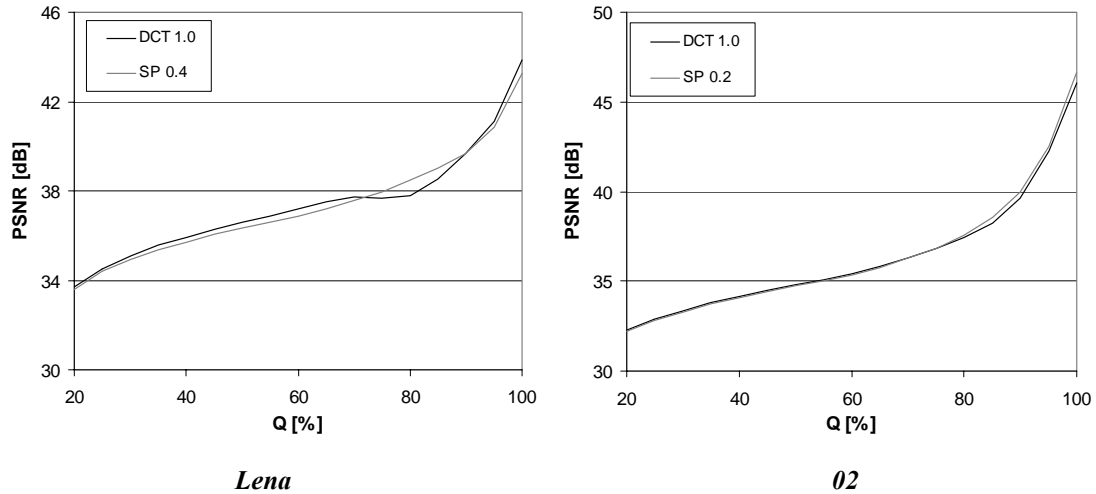


Figura 3.32 – PSNR resultante após inserção da marca e compressão JPEG.

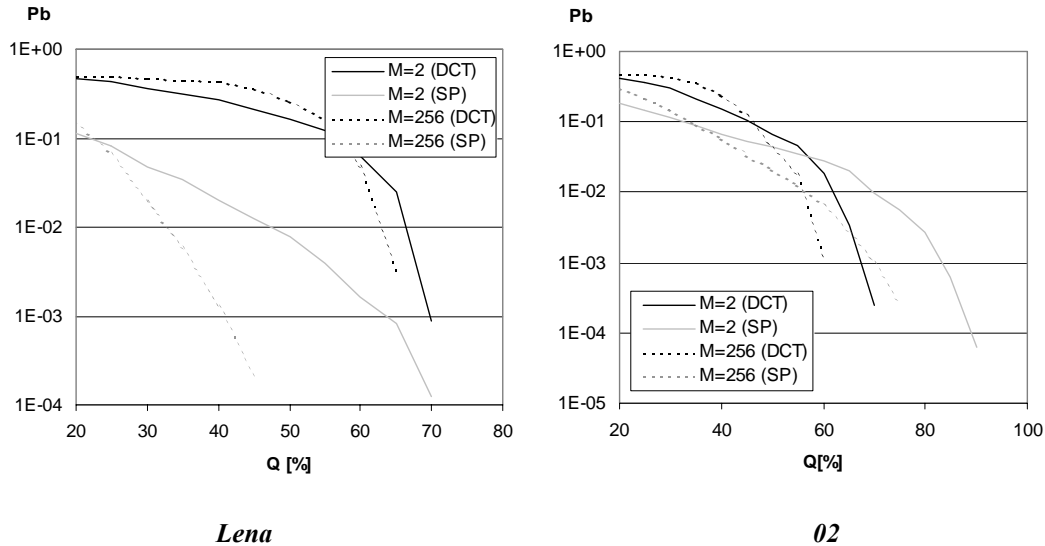


Figura 3.33 – Comparação de resultados nos dois domínios quando em presença de compressão.

Em ambos os casos observa-se que, com PSNRs semelhantes, o melhor desempenho é exibido pelo sistema de marcas-de-água no domínio da frequência: a partir de um dado valor do número de pixels por bit de informação útil, as curvas correspondentes ao domínio da frequência assumem valores mais baixos e decrescem com declive mais acentuado.

De um modo análogo, foi também realizada uma comparação dos resultados obtidos em ambos os domínios, quando em presença de compressão JPEG. Na figura 3.32 encontram-se representadas as evoluções da PSNR resultante, em função do factor de qualidade da compressão, após inserção da marca-de-água e posterior compressão da imagem marcada. A

densidade de inserção referente a este teste foi de 100%. Por observação da figura 3.32-a) pode-se concluir que, para a imagem *Lena*, a utilização de uma força de inserção de 0.4 no domínio espacial conduz a PSNRs semelhantes aos obtidos para uma força de inserção de 1.0 no domínio da frequência. Para a imagem *02*, e como comprova a figura 3.32-b), para obtenção de PSNRs semelhantes, deve-se ajustar a força de inserção no domínio espacial para o valor de 0.2 e a do domínio da frequência para 1.0.

Na figura 3.33 encontram-se representadas as curvas de probabilidades de erro de bit obtidas em cada domínio, utilizando as forças de inserção referidas. Verifica-se que os resultados obtidos para as imagens *Lena* e *02* são contraditórios – na imagem *Lena*, o domínio espacial apresenta melhores resultados, mas para a imagem *02* é no domínio da frequência que se verifica melhor desempenho. Estes resultados poderão ser devido a um modelo perceptual para inserção no domínio da frequência melhor adaptado à imagem *02*, ou a uma má escolha dos coeficientes para inserção no caso da imagem *Lena*. Determinar os melhores coeficientes DCT para inserção da marca-de-água, avaliar a dependência dessa escolha do tipo de imagem e ainda avaliar o desempenho de outros modelos perceptuais para este domínio, são certamente tópicos que merecem ser investigados em trabalho futuro.

3.8 Considerações finais

Ao longo deste capítulo comprovou-se, teoricamente e experimentalmente, que a utilização de sinalização multinível conduz a um aumento de desempenho dos sistemas de marca-de-água e de forma tanto mais significativa quanto maior for o número de níveis utilizados. Esta conclusão aplica-se tanto para inserção no domínio espacial como para inserção no domínio da frequência. O acréscimo de custo devido à utilização de sinalização multinível traduz-se num maior número de operações aritméticas realizadas, o que implica um aumento da complexidade do sistema. Estas operações são no entanto paralelizáveis, pelo que o aumento da complexidade não se deverá reflectir em aumento do tempo de processamento, crucial para aplicações em tempo real.

De modo análogo, o desempenho em presença de compressão JPEG também melhora quando o número de níveis utilizados na sinalização aumenta. De facto, por análise dos resultados obtidos experimentalmente, para probabilidades de erro de bit na extracção da marca-de-água inferiores a aproximadamente 0.1, a utilização de sinalização multinível revela-se sempre vantajosa.

Em presença de ruído branco e gaussiano, bem como em presença de cortes sobre as imagens de teste, as conclusões em relação à utilização de sinalização multinível são idênticas. No caso de

cortes, verifica-se no entanto que a diferença trazida pelo aumento do número de níveis de sinalização não é tão notória como nos restantes tipos de testes realizados.

Por último, a inserção no domínio da frequência aparenta ser mais robusta que a inserção no domínio espacial, sobretudo no caso em que a imagem marcada não sofre “ataques”. Em presença de compressão JPEG não é possível afirmar o mesmo, já que para as duas imagens utilizadas nos testes os resultados foram contraditórios. Investigar as razões de tais resultados, o que poderá conduzir à determinação de coeficientes e modelos perceptuais mais adequados para inserção no domínio DCT, será certamente um tópico com interesse para trabalho futuro.

Capítulo 4

Codificação para correcção de erros

4.1 Introdução

Como foi visto no capítulo anterior, os sinais modulados por espalhamento de espectro ocupam uma banda de frequências muito superior à do sinal de informação (sinal modulante). A redundância introduzida nesta expansão de banda confere, a esta forma de modulação, a capacidade de resistir a elevados níveis de interferência, que podem ocorrer durante a transmissão do sinal. Esta capacidade pode ser ainda aumentada se se associar o espalhamento de espectro a técnicas de codificação de canal [34].

A codificação de canal pode ser vista como uma forma eficiente de adicionar informação redundante à mensagem original, redundância essa que é utilizada na recepção de modo a possibilitar a recuperação da mensagem original, detectando e/ou corrigindo erros que tenham eventualmente ocorrido durante a transmissão. A utilização de codificação de canal conduz a um aumento da complexidade do sistema, em maior ou menor grau, consoante o tipo de código utilizado.

Existem alguns trabalhos de investigação centrados na aplicação de codificação correctora de erro à área da assinatura digital de imagens. Em [19], os autores apresentam um esquema de

marcas-de-água baseado em espalhamento de espectro com inserção espacial, no qual a marca-de-água é codificada recorrendo aos códigos de bloco binários BCH (*Bose-Chaudhuri-Hocquenhem*). No trabalho apresentado em [21] é feita uma comparação do desempenho demonstrado pelos códigos BCH e pelos códigos convolucionais binários, quando aplicados a um esquema de marcas-de-água com inserção no domínio espacial e orientado a blocos de dimensão 8×8 pixels (sem espalhamento de espectro).

Neste capítulo efectua-se um estudo comparativo do desempenho de códigos correctores de erros comuns em sistemas de comunicação digital, nomeadamente, códigos de bloco binários (BCH), códigos de bloco não binários (*Reed Solomon*) e códigos convolucionais binários, quando aplicados ao esquema de assinatura de imagens apresentado no capítulo 3. Avalia-se também, analiticamente e experimentalmente, a melhoria de desempenho obtida pela utilização de codificação de canal face à utilização de sinalização multinível simples.

Após a secção introdutória, apresentam-se, na secção 4.2, as principais classes de códigos de correcção de erro. Em 4.3 descrevem-se os esquemas de inserção e extracção da marca-de-água com utilização de codificação de canal, analisando-se, teórica e experimentalmente, o desempenho dos códigos implementados para inserção da marca no domínio espacial. Na secção 4.4 comparam-se os resultados obtidos. Nas secções 4.5 e 4.6 avalia-se experimentalmente o desempenho da codificação de canal em presença de compressão JPEG e em presença de cortes e ruído gaussiano, respectivamente. Em 4.7 apresentam-se resultados obtidos com inserção no domínio da frequência. Na secção 4.8 sumarizam-se as principais conclusões deste capítulo.

4.2 Principais classes de codificação de canal

Existem diversas classes de códigos para codificação de canal [31]. Uma revisão de todas estas classes seria demasiado exaustiva, saindo do âmbito dos objectivos desta tese. Deste modo, revêem-se nesta secção as classes de códigos habitualmente utilizadas em sistemas de telecomunicações, com particular destaque para os códigos em estudo nesta tese: *códigos de bloco* (binários ou não) e *códigos convolucionais binários*.

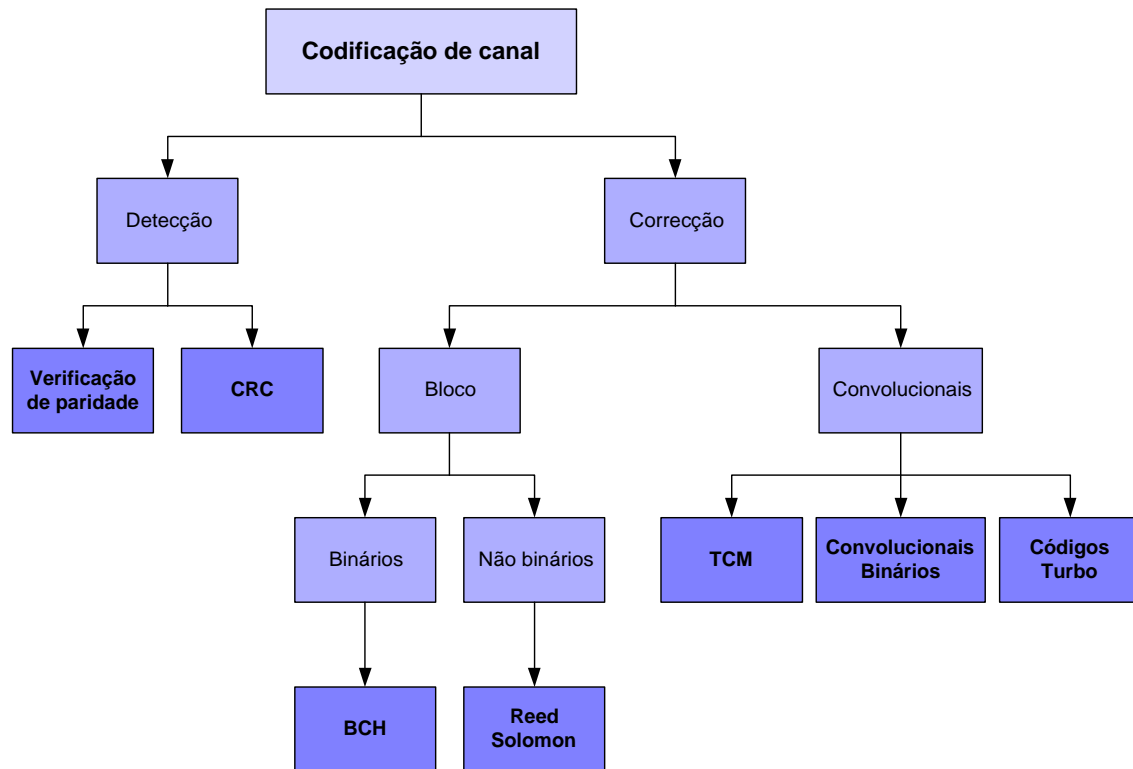


Figura 4.1 – Taxonomia das técnicas de codificação de canal usuais em telecomunicações.

4.2.1 Taxonomia

A figura 4.1 representa uma taxonomia das principais classes de códigos para codificação de canal [28]. Uma primeira divisão diferencia os *códigos de detecção de erro* dos *códigos de correcção de erro*.

Os *códigos de detecção de erro* acrescentam redundância à mensagem a transmitir e de modo a que o receptor possa detectar a ocorrência de erros na transmissão. No entanto, esta classe de códigos não efectua nenhum tipo de correcção de erro. São normalmente utilizados em sistemas de comunicações nos quais é retransmitida a mensagem, caso o receptor detecte a ocorrência de erros na transmissão. Os códigos de detecção de erro, em regra geral, adicionam menos redundância à mensagem que os códigos de correcção de erro e os seus descodificadores apresentam menor complexidade.

O código de detecção de erro mais simples é o *código de verificação de paridade*, que consiste em acrescentar à mensagem original um único bit, designado por *bit de paridade*. Se a este bit estiver atribuída uma paridade par, o seu valor é escolhido de modo a que o número de bits a “1” na mensagem codificada seja par. Como exemplo, se se pretendesse transmitir a mensagem

“10010100”, seria acrescentado o bit de paridade com valor “1”, constituindo a palavra de código “100101001”. Este tipo de códigos é muito simples de implementar mas tem o inconveniente de apenas permitir detectar um número ímpar de erros.

Uma classe de códigos de detecção de erro que tenta ultrapassar este problema é a classe dos *códigos de verificação cíclica*, habitualmente designados por CRC (*Cyclic Redundancy Check*). Neste caso, é introduzida uma redundância maior à mensagem, o que permite detectar um maior leque de combinações de erro.

Os *códigos de correcção de erro*, para além da detecção de erros, efectuem a sua correcção desde que o número de erros não exceda um dado valor, designado por *capacidade de correcção* do código. Esta capacidade varia de código para código, sendo geralmente tanto maior, quanto maior for a redundância acrescentada à mensagem original. Existem duas principais classes de códigos de correcção de erro: *códigos de bloco* e *códigos convolucionais*.

O princípio de funcionamento dos *códigos de bloco* consiste em acrescentar à mensagem original um conjunto de símbolos. A mensagem codificada passa então a ser constituída por dois blocos distintos: um bloco correspondente à mensagem original e um bloco correspondente à redundância introduzida pelo código. Estes códigos podem ser divididos duas classes, consoante o número de níveis de representação utilizados: *códigos de bloco binários*, se o alfabeto da mensagem é constituído por apenas 2 símbolos, e *códigos de bloco não-binários*, se o alfabeto contém mais do que 2 símbolos. Os códigos de bloco binários BCH (*Bose-Chaudhuri-Hocquenhem*) e os códigos de bloco não binários RS (*Reed-Solomon*) constituem duas das famílias de códigos mais utilizados em telecomunicações.

O princípio de funcionamento dos *códigos convolucionais* difere do dos códigos de bloco. Neste caso, o codificador opera sobre a mensagem original utilizando uma janela deslizante, produzindo à sua saída e de forma contínua, um conjunto de símbolos codificados. Cada símbolo do sinal a codificar afecta um certo número de símbolos consecutivos na mensagem codificada, sendo este número tanto maior quanto maior for o tamanho da janela deslizante. Dentro da categoria dos códigos convolucionais, existem três principais famílias de códigos: *códigos convolucionais binários*, *codificação-modulação Trellis* e *códigos turbo*.

Os *códigos convolucionais binários* constituem a classe mais simples de códigos convolucionais, operando sobre um alfabeto constituído por dois símbolos (bits). A grande

vantagem destes códigos face aos códigos de bloco binários é a sua descodificação, que pode ser realizada segundo o algoritmo de *Viterbi* [46]. Neste caso, a decisão sobre os símbolos transmitidos, pode ter em consideração quer as características do canal quer os níveis de sinal à saída do detector, correspondentes aos símbolos codificados – *decisão suave* (*soft decision*). Nos códigos de bloco, a implementação de um esquema de decisão suave é complexa, optando-se geralmente por esquemas de *decisão dura* (*hard decision*) ou seja, decisão baseada apenas nos símbolos codificados recebidos.

A *codificação-modulação Trellis* ou *TCM* (*Trellis Coded Modulation*) é uma técnica de codificação que trata a modulação e a codificação em conjunto, i.e., em vez de a codificação e a modulação serem efectuadas sequencialmente, existe um único bloco que realiza as duas operações em simultâneo. O princípio por detrás desta técnica é a utilização de um código convolucional com características tais que a distância Euclidiana entre cada símbolo gerado à sua saída é maximizada.

Os *códigos turbo*, numa abordagem simplista, resultam da conjunção de dois códigos convolucionais distintos – a codificação é realizada utilizando dois codificadores convolucionais em série, separados por um bloco com a função de desordenar as sequências codificadas resultantes do primeiro codificador, evitando-se deste modo erros de rajada. Na descodificação, são realizadas as operações inversas das efectuadas na codificação. Estes códigos apresentam um elevado desempenho, mas a sua complexidade é elevada, pelo que a sua utilização deve ser bem ponderada em sistemas que requeiram baixo custo e rapidez de processamento.

4.2.2 Códigos de bloco

Como já foi referido, os códigos de bloco acrescentam um conjunto de símbolos à mensagem original para formar uma palavra de código. Na figura 4.2 apresenta-se um esquema simplificado da estrutura de uma palavra de um código de bloco. Seja k o comprimento (em símbolos) da mensagem e n o comprimento (em símbolos) da palavra de código correspondente, após acrescentados os símbolos de paridade. Nesta situação, o código é designado por (n,k) e define-se a taxa de codificação do código – c_r – como:

$$c_r = \frac{k}{n}. \quad (4.1)$$

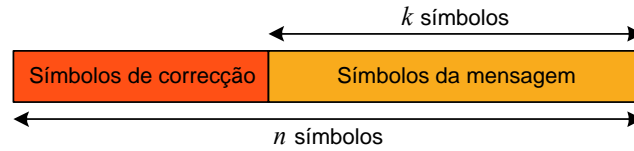


Figura 4.2 – Estrutura de palavra de um código de bloco linear.

Antes de prosseguir, é conveniente definir conceito de distância de *Hamming*. A distância de *Hamming* entre duas palavras de código corresponde ao número de posições para as quais os símbolos da palavra diferem. Como exemplo, a distância de *Hamming* entre as palavras “110101” e “111001” é 2. A partir da distância de *Hamming* define-se a *distância mínima* – d_{\min} – de um código de bloco, como a mínima distância de *Hamming* existente entre qualquer par de palavras desse código. O parâmetro d_{\min} é de particular interesse, pois relaciona-se directamente com a capacidade de correcção do código.

A capacidade de correcção – t – de um código de bloco define-se como o número máximo de erros de símbolo que podem ocorrer numa palavra de código, para o qual o código consegue decodificar correctamente a palavra original. A capacidade de correcção relaciona-se com a distância mínima de acordo com:

$$t = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor, \quad (4.2)$$

em que $\lfloor \cdot \rfloor$ designa a função *floor*(.)²⁶. Geralmente, um aumento da distância mínima entre as palavras do código conduz a uma maior capacidade de correcção de erros, à custa de uma diminuição da taxa de codificação (maior redundância).

4.2.3 Códigos convolucionais binários

A figura 4.3 ilustra o princípio de funcionamento de um codificador convolucional. A mensagem original é uma sequência de bits com comprimento arbitrário – $\{b_1, b_2, b_3, \dots\}$ – que entra sequencialmente no codificador, constituído por um registo de deslocamento com 2 memórias e por um somador lógico (ou-exclusivo). Em cada ciclo de relógio do codificador, o conteúdo do registo é deslocado uma posição para a direita, entra no codificador um novo bit da mensagem e são gerados dois bits codificados: um correspondente ao próprio bit que entra no codificador – b_i – e outro correspondente à soma lógica – l_i – de b_i com o bit anterior da mensagem – b_{i-1} . Como são gerados dois bits à saída do codificador por cada bit que entra, a taxa de codificação deste código é $1/2$.

²⁶ A função *floor*(x) devolve o maior inteiro não superior a x .

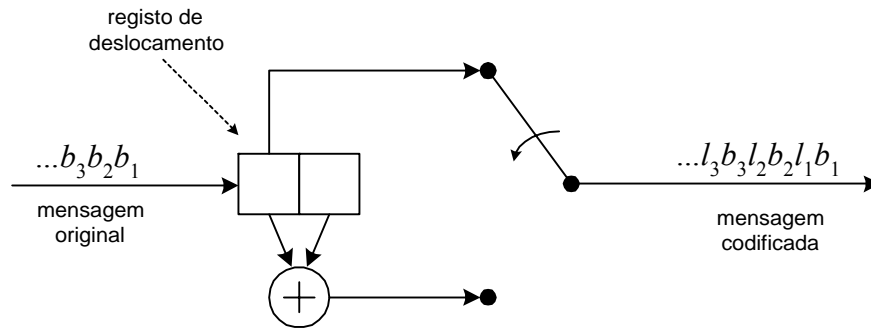


Figura 4.3 – Codificador convolucional simples.

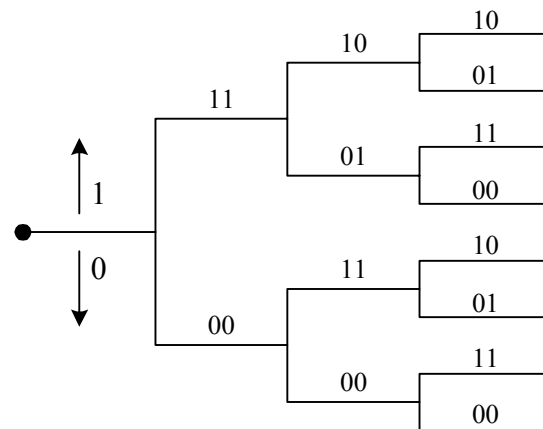


Figura 4.4 – Árvore do código convolucional referente à figura 4.3.

Uma maneira de visualizar o resultado da codificação convolucional é a representação em *diagrama em árvore*, que se apresenta na figura 4.4: ao ser codificada uma sequência de bits, na codificação de um bit com valor lógico “1” deverá ser seguido o ramo superior; na codificação de um bit com valor lógico “0” deverá ser seguido o ramo inferior. Como exemplo, uma mensagem representada pela sequência binária “100” origina a sequência codificada “110100”.

De uma maneira mais geral, num código convolucional binário a mensagem original é codificada com um registo de deslocamento que contém $b \times k$ memórias binárias, em que k é designado por *comprimento restritivo*²⁷. Em cada ciclo de relógio do codificador, o conteúdo dos registos é deslocado b posições para a direita, um grupo de b bits entra no codificador e são gerados n bits codificados. Nestas condições, cada grupo de n bits codificados resulta de uma combinação linear dos $b \times k$ bits que entraram mais recentemente no codificador e o código convolucional designa-se por $(1/c_r, k)$, em que c_r é a taxa de codificação do código, definida por:

²⁷ Na literatura Inglesa, *constraint length*.

$$c_r = \frac{b}{n}. \quad (4.3)$$

De notar que cada bit da mensagem original que entra no codificador convolucional pode influenciar um conjunto de $k \times n$ bits codificados. Devido a este facto, a distância mínima – d_{\min} – característica de um código convolucional, é a menor distância de *Hamming* entre pares de sequências codificadas, com comprimento $k \times n$, contabilizadas a partir do mesmo nó da árvore do código e seguindo ramos diferentes a partir desse nó. Podem também ser medidas distâncias mínimas sobre sequências codificadas de comprimento $L \times n$, desde que $L > k$. Designando o parâmetro L como *profundidade de decodificação*, a distância mínima – d_L – entre sequências codificadas de comprimento $L \times n$, designa-se por *distância mínima de decodificação*. Quando L tende para infinito, a distância mínima resultante – d_f – designa-se por *distância mínima livre*, sendo este um conceito muito importante em códigos convolucionais, pois relaciona-se directamente com o seu desempenho. A relação entre as várias distâncias mínimas definidas é dada por:

$$d_f \geq d_L \geq d_{\min}. \quad (4.4)$$

Os códigos convolucionais podem também ser caracterizados pela sua *função de transferência*, uma função polinomial que contém informação relativa às distâncias do código [31,46], e que é particularmente útil na análise do desempenho da codificação convolucional, como será visto mais adiante.

Em canais corrompidos por ruído aditivo, branco e gaussiano, o algoritmo de *Viterbi* é o decodificador óptimo para os códigos convolucionais [46]. Este algoritmo é usualmente implementado com decisão suave (*soft decision*), já que esta solução conduz a um considerável aumento de desempenho na decodificação face à decisão dura (*hard decision*), justificando o aumento na complexidade do decodificador.

4.3 Aplicação de codificação de canal a marcas-de-água

As figuras 4.5 e 4.6 apresentam, respectivamente, os novos esquemas gerais para inserção e extracção da marca. Estes resultam da introdução, nos esquemas apresentados no capítulo 3, dos blocos que efectuem a codificação e a decodificação de canal.

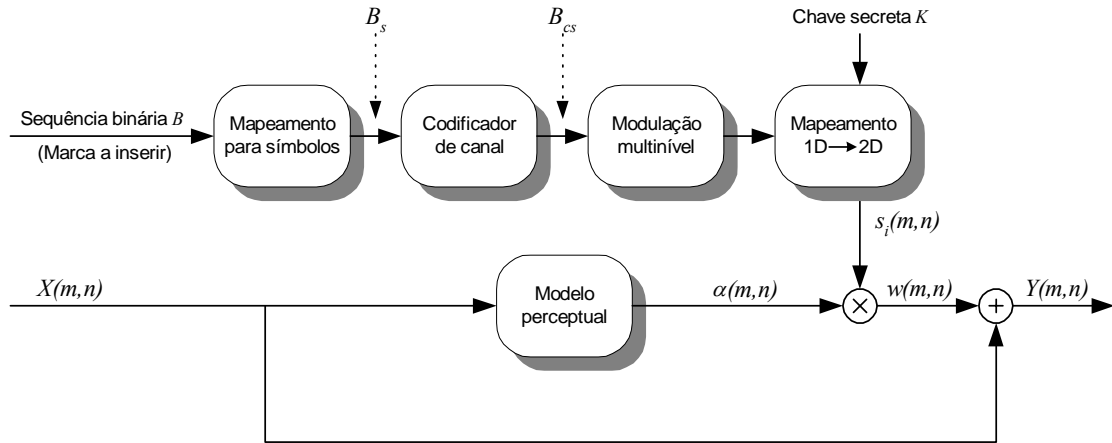


Figura 4.5 – Esquema geral de inserção da marca-de-água (com codificador).

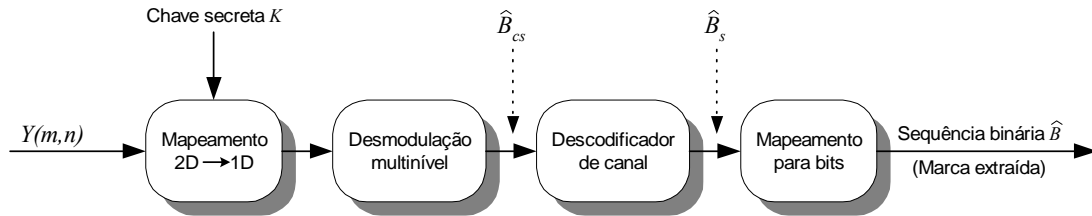


Figura 4.6 – Esquema geral de extração da marca-de-água (com decodificador).

Quando é utilizada codificação de canal, a sequência a inserir na imagem – B_{cs} – é composta por N_{cs} símbolos codificados. A relação entre N_{cs} e o número de símbolos não codificados da mensagem – N_s – é dada por:

$$N_{cs} = N_s / c_r, \quad (4.5)$$

em que c_r é a taxa de codificação, definida em (4.1) e (4.3).

No que se segue, derivam-se as expressões analíticas da probabilidade de erro de bit da marca-de-água extraída, em função dos parâmetros que modelizam o canal – μ e σ – para os códigos estudados nesta tese: códigos de bloco (BCH, RS) e código binário convolucional.

4.3.1 Códigos de bloco binários

Considere-se o caso em que sinalização binária antipodal é utilizada em conjunto com um código de bloco binário (n, k) , com capacidade de correção t .

Admitindo que os erros de bit são acontecimentos independentes e ocorrem com probabilidade P_b , a probabilidade $P(i,n)$ de ocorrerem i erros numa palavra de código composta por n bits é dada por:

$$P(i,n) = \binom{n}{i} P_b^i (1 - P_b)^{n-i}. \quad (4.6)$$

A probabilidade de erro de bit – P_b – obtém-se de (3.39), com μ e σ calculados para o número de bits resultantes da codificação de canal, i.e., $N_b = n = N_b / c_r$. A probabilidade – P_{de} – de se ter à entrada do decodificador uma palavra de código contendo mais do que t erros, é majorada por:

$$P_{de} \leq \sum_{i=t+1}^n P(i,n). \quad (4.7)$$

Substituindo (4.6) em (4.7) resulta:

$$P_{de} \leq \sum_{i=t+1}^n \binom{n}{i} P_b^i (1 - P_b)^{n-i}. \quad (4.8)$$

Para determinar um limite superior para a probabilidade de erro de bit à saída do decodificador de canal – P_{db} – admita-se que a descodificação de uma palavra de código com i erros ($i > t$), pode originar até t erros adicionais na palavra descodificada ($i+t \leq n$). Neste caso, uma fracção $(i+t)/n$ dos k bits de informação poderá ser afectada com erro, vindo:

$$P_{db} \leq \frac{1}{n} \sum_{i=t+1}^n \min(i+t, n) \cdot \binom{n}{i} P_b^i (1 - P_b)^{n-i}. \quad (4.9)$$

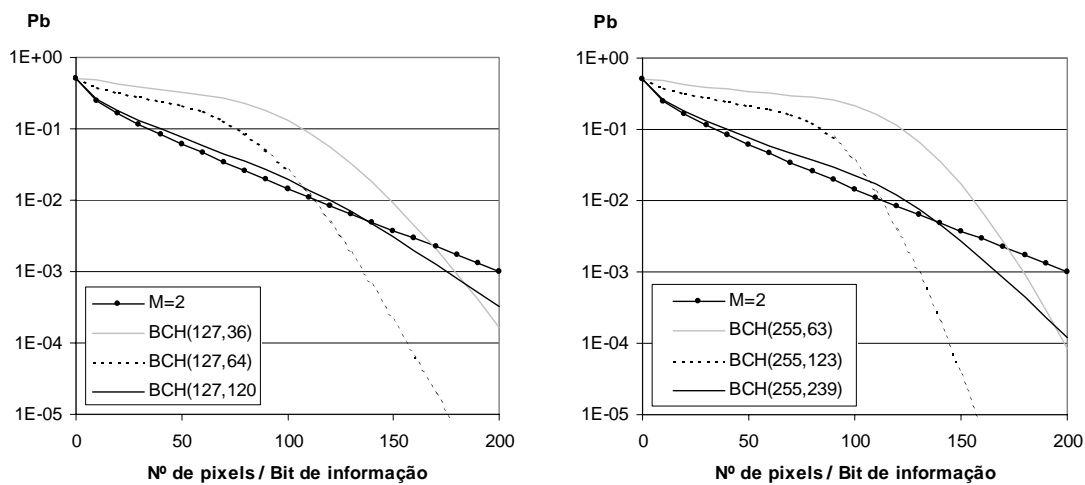
O termo $\min(i+t, n)$ em (4.9) garante que ocorrências (impossíveis) de mais de n erros por palavra não são contabilizadas.

Resultados teóricos

Na figura 4.7 encontram-se representadas curvas da probabilidade de erro de bit após descodificação, em função do número de pixels utilizados para inserir um bit da marca, obtidas para diversos códigos de bloco binários BCH, cujas características estão indicadas na tabela 4.1. Os parâmetros μ e σ necessários para o cálculo de P_b foram obtidos de acordo com (3.30) e (3.31). Todos os resultados são referentes à inserção no domínio espacial da imagem *Lena*, com uma força de inserção de 0.4.

Código BCH (n,k)	Capacidade de correcção (t)	Taxa de codificação (c_r)
(127,36)	15	0.283
(127,64)	10	0.504
(127,120)	1	0.945
(255,63)	30	0.247
(255,123)	19	0.482
(255,239)	2	0.937

Tabela 4.1 – Códigos de bloco binários BCH analisados.

Figura 4.7 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos BCH (Imagem *Lena*).

Observando os gráficos pode-se concluir que a utilização de códigos de bloco binários é vantajosa relativamente ao caso de sinalização binária não codificada. Em todos os casos e a partir de um certo valor do *número de pixels por bit de informação útil*, as curvas correspondentes à utilização de correcção de erro apresentam probabilidades de erro de bit mais baixas e decrescem mais rapidamente do que as curvas correspondentes à utilização de sinalização binária sem codificação de canal. A probabilidade P_b para a qual as curvas referentes a $M=2$, sem codificação, apresentam melhores resultados que as curvas relativas a codificação, assume um valor demasiado elevado ($P_b > 10^{-3}$) para aplicações com interesse prático.

Comparando os diversos códigos apresentados verifica-se que, para a gama de valores de probabilidade de erro apresentadas, o melhor desempenho é conseguido pelos códigos cuja taxa de codificação se situa em valores próximos de $\frac{1}{2}$. Quando a taxa de codificação se aproxima do valor 1, verifica-se que as curvas referentes ao uso de codificação se aproximam da curva

correspondente à ausência de codificação. Isto deve-se ao facto da capacidade de correcção do código ser muito baixa para este valor da taxa de codificação.

Quando as taxas de codificação são próximas, o melhor desempenho é obtido pelo código cujo comprimento de palavra é maior (n maior). Assim sendo, na decisão sobre o código de correcção de erro a utilizar deve-se ter em conta o comprimento em bits da marca.

Resultados experimentais

Para confirmar os resultados obtidos teoricamente, foi também realizada uma simulação experimental utilizando códigos BCH. Avaliar experimentalmente o desempenho de todos os de códigos apresentados anteriormente seria uma tarefa exaustiva, dado o tempo despendido nas simulações. Por esta razão, optou-se por utilizar apenas o código BCH(127,64) que demonstrou teoricamente um bom desempenho face aos restantes códigos analisados.

Na figura 4.8 apresentam-se os resultados experimentais para inserção no domínio espacial. Cada ponto destes gráficos foi obtido após efectuadas 1000 inserções e extracções de marcas-de-água geradas aleatoriamente. Tal como para a análise teórica, a força de inserção foi ajustada para o valor de 0.4. Apresentam-se também as curvas obtidas teoricamente, de forma a que se possam validar directamente as evoluções previstas para P_b .

Como se pode observar, o desempenho demonstrado pelo uso do código BCH(127,64) é superior ao demonstrado na ausência de codificação verificando-se, a partir de um certo valor do número de pixels por bit de informação, probabilidades de erro de bit mais baixas e curvas com declive negativo mais acentuado. Na zona em que a ausência de codificação de canal apresenta melhores resultados (valores baixos do nº de pixels / bit de informação) os valores de probabilidade de erro de bit são demasiado elevados ($P_b > 10^{-2}$), pelo que nunca seriam aceitáveis numa implementação real de um sistema de marcas-de-água.

Os resultados obtidos experimentalmente encontram-se próximos dos previstos teoricamente. O maior desvio verifica-se nos resultados referentes à imagem 02, provavelmente devido às aproximações feitas no cálculo das relações sinal-ruído (μ / σ) utilizadas para traçar as curvas teóricas.

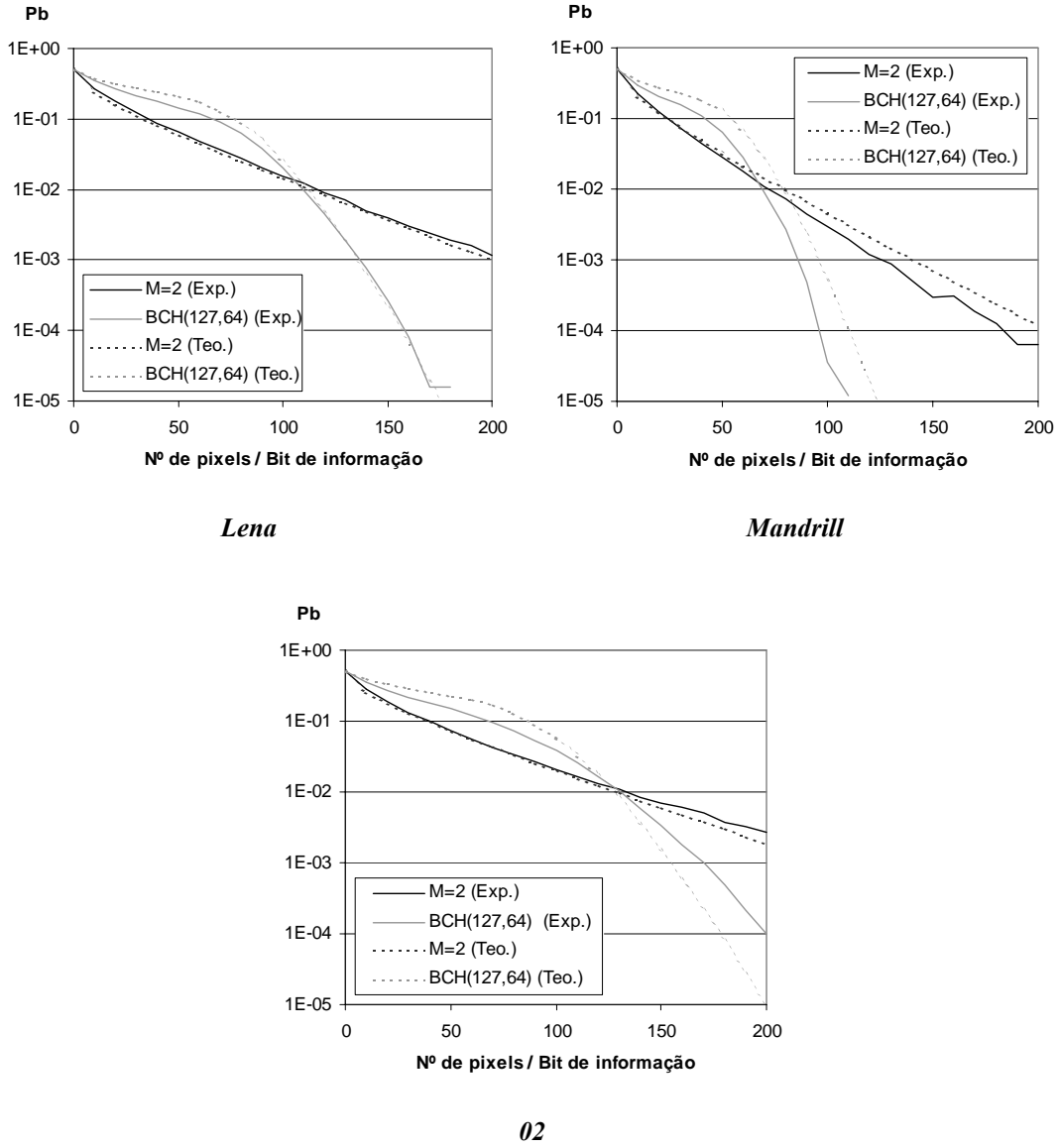


Figura 4.8 – Resultados teóricos e experimentais de P_b vs. N° de pixels / Bit de informação para o código BCH(127,64).

4.3.2 Códigos de bloco não binários

Designando por P_M a probabilidade de erro de símbolo à saída do canal, e admitindo que a existência de i erros de símbolo na palavra recebida pode originar até $i+t$ símbolos errados na palavra decodificada, a probabilidade de erro de símbolo – P_{DM} – é majorada por:

$$P_{DM} \leq \frac{1}{N} \sum_{i=t+1}^N \min(i+t, N) \binom{N}{i} P_M^i (1-P_M)^{N-i}, \quad (4.10)$$

Código RS (N,K)	Capacidade de correcção (t)	Taxa de codificação (c_r)
(10,8)	1	0.800
(14,8)	3	0.571
(22,16)	3	0.727
(30,16)	7	0.533

Tabela 4.2 – Códigos de bloco não-binários (RS) analisados.

em que N designa o comprimento (em símbolos M -ários) da palavra de código e K o comprimento (em símbolos M -ários) da palavra a codificar.

Supondo erros de símbolo equiprováveis, cada erro ocorrerá com probabilidade $P_{DM}/(2^l-1)$, em que l é o número de bits por símbolo. Tendo em conta que existem C_j^l maneiras diferentes de ocorrerem j erros em l bits, o número de médio de bits errados por símbolo é dado por:

$$\sum_{j=1}^l j \binom{l}{j} \frac{P_{DM}}{2^l-1}. \quad (4.11)$$

A probabilidade de erro de bit média – P_{db} – na palavra descodificada obtém-se dividindo (4.11) pelo número de bits por símbolo – l :

$$P_{db} = \frac{\sum_{j=1}^l j \binom{l}{j} \frac{P_{DM}}{2^l-1}}{l} = \frac{2^{l-1}}{2^l-1} P_{DM}. \quad (4.12)$$

Resultados teóricos

Para avaliar o desempenho dos códigos de bloco não-binários (*Reed Solomon*), obtiveram-se as curvas de P_{db} vs. nº de pixels por bit de informação correspondentes à equação (4.12), e para o conjunto de códigos com as características sintetizadas na tabela 4.2.

Na figura 4.9 apresentam-se os resultados obtidos quando estes códigos são utilizados em conjunto com 16 níveis de sinalização. As curvas foram traçadas utilizando o método descrito na secção 4.3.1, com inserção no domínio espacial e para a imagem *Lena*.

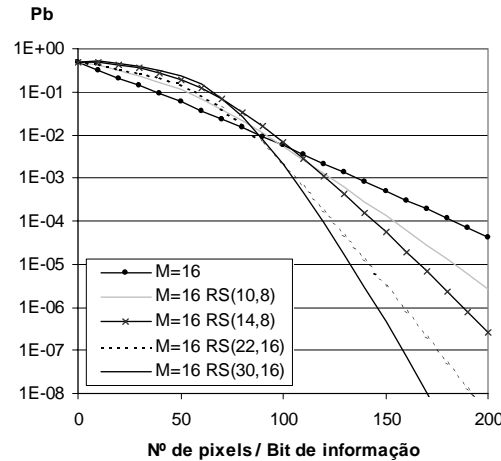


Figura 4.9 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos RS utilizando 16 níveis de sinalização (Imagem *Lena*).

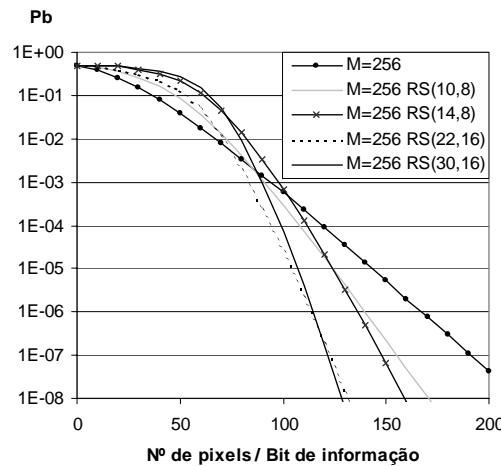


Figura 4.10 – Resultados teóricos de P_b vs. N° de pixels / Bit de informação para diversos códigos RS utilizando 256 níveis de sinalização (*Lena*).

A melhoria de desempenho resultante da utilização dos códigos RS é notável. Observa-se também que com o mesmo valor de K , e dentro da gama de valores de probabilidade de erro de bit presentes no gráfico, é mais vantajoso utilizar o código que apresenta uma taxa de codificação mais baixa. À semelhança do que foi observado para os códigos binários BCH, para taxas de codificação semelhantes, os melhores resultados obtêm-se para os códigos com maior comprimento de palavra (N maior).

A figura 4.10 representa os resultados obtidos utilizando os códigos RS presentes na tabela 4.2, para 256 níveis de sinalização. As conclusões que se podem tirar a partir deste gráfico são equivalentes às do caso $M=16$. A principal diferença reside no facto de as curvas agora representadas caírem com declives mais acentuados, e os pontos de cruzamento ocorrerem para valores mais baixos da probabilidade de erro de bit.

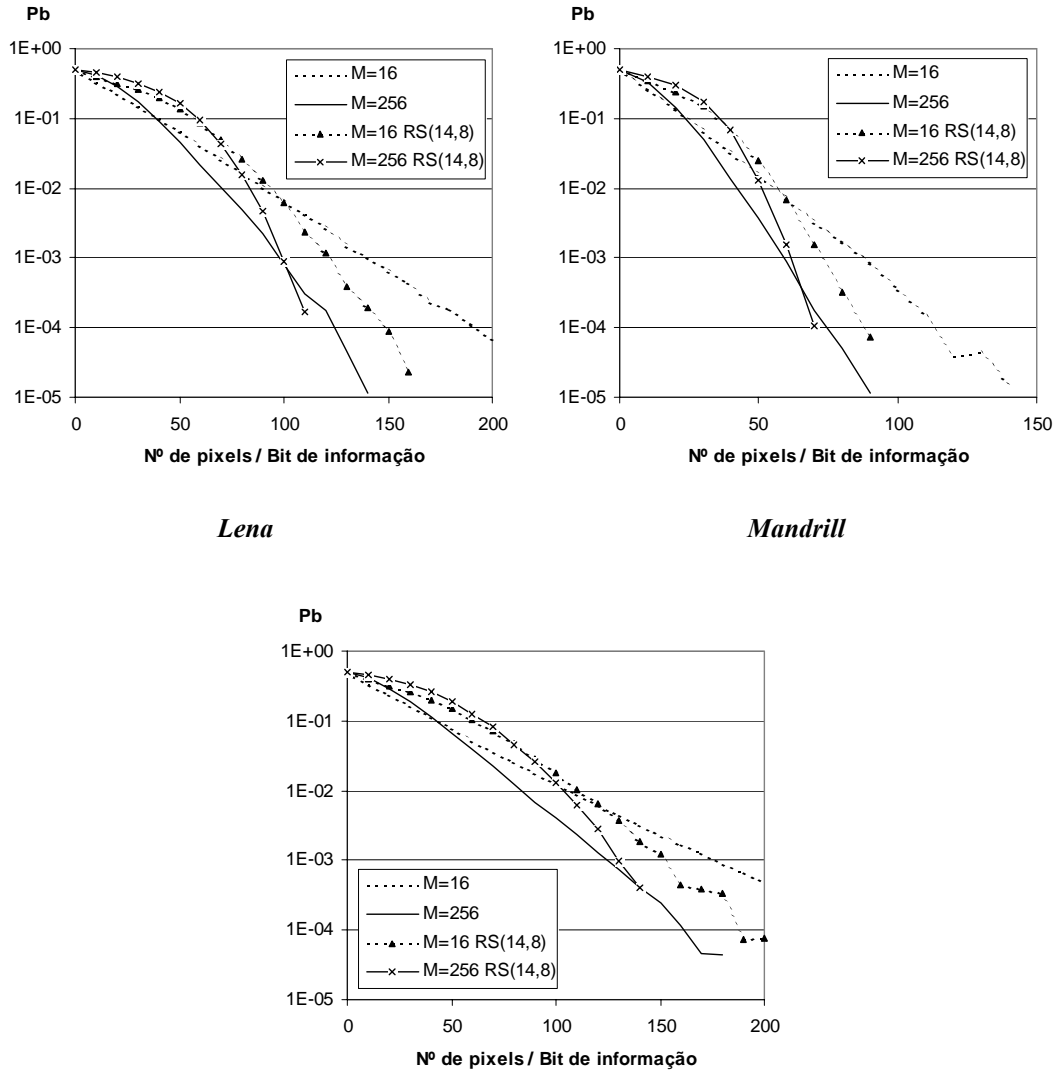


Figura 4.11 – Resultados experimentais de P_b vs. N° de pixels / Bit de informação para o código RS(14,8) utilizando 16 e 256 níveis de sinalização.

Resultados experimentais

A figura 4.11 ilustra os resultados experimentais obtidos para o código RS(14,8), utilizando 16 ou 256 níveis na sinalização. O número de testes realizados por imagem e o valor da força de inserção foram idênticos aos descritos na secção 4.3.1.

O desempenho dos códigos RS aplicados ao sistema é superior ao demonstrado pela sinalização *M-ária* na ausência de codificação. Para qualquer uma das imagens e número de níveis de sinalização usados, as curvas de P_{db} correspondentes à codificação RS ultrapassam e decrescem mais rapidamente que as relativas ao uso de sinalização *M-ária* sem codificação.

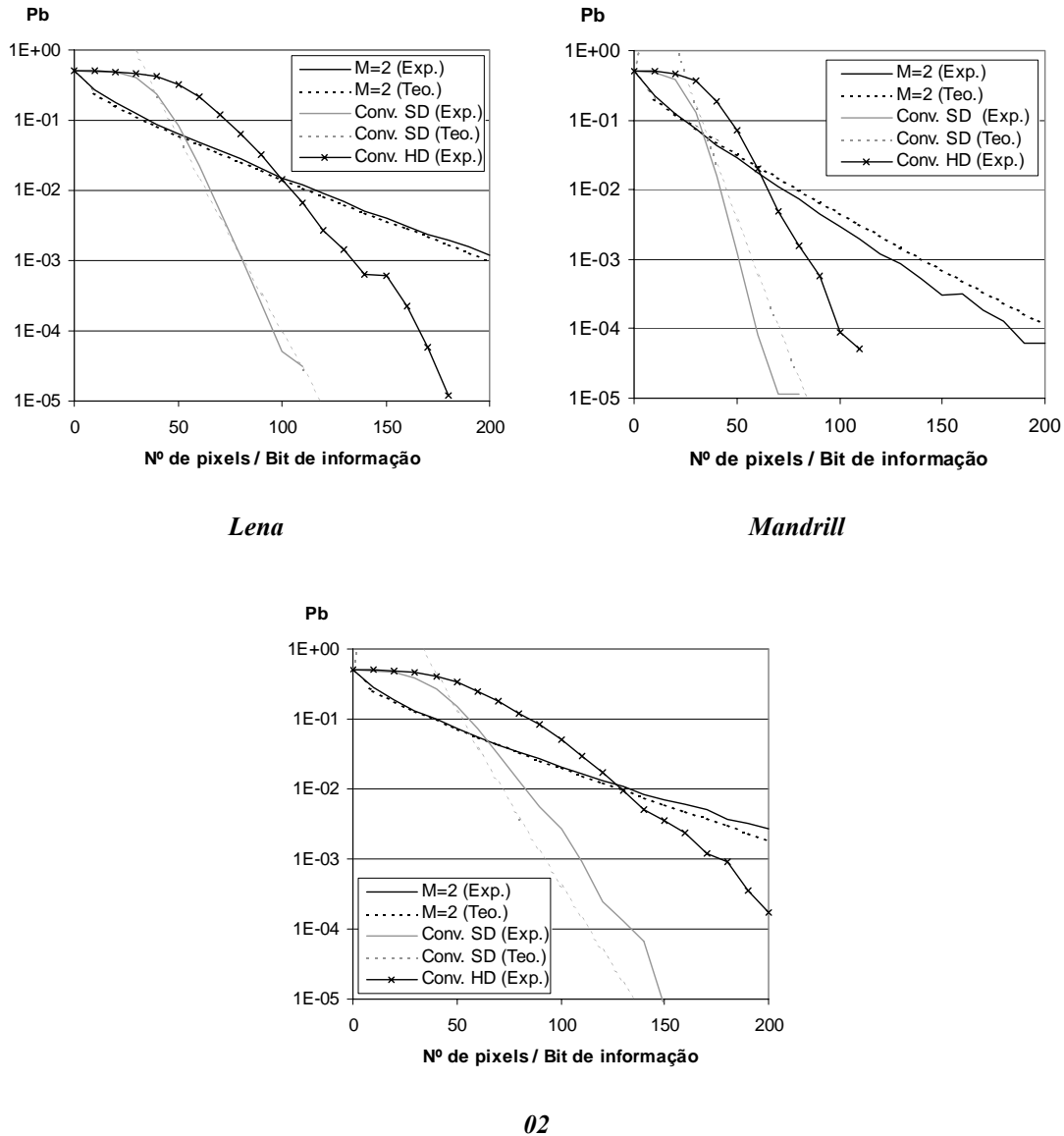


Figura 4.12 – Resultados teóricos e experimentais de P_b vs. N° de pixels / Bit de informação para o código convolucional com $c_r=1/2$ com decisões dura e suave.

Deste modo ficam comprovados os resultados previstos teoricamente e fica demonstrada a eficácia da utilização de códigos de bloco não binários, quando comparada com sinalização multinível sem qualquer tipo de codificação de canal.

4.3.3 Códigos convolucionais binários

Considere-se agora o caso em que a marca-de-água é codificada utilizando um código convolucional. Admita-se ainda que a decodificação do código é realizada utilizando o algoritmo de *Viterbi* com decisão suave. A derivação de um limite para a probabilidade de erro de bit após decodificação – P_{db} – envolve duas características intrínsecas de cada código

convolucional: a sua distância mínima livre – d_f – e a função de transferência do código convolucional – $T(D, N)$ – definidas na secção 4.2.3. Pode-se demonstrar que a probabilidade de erro de bit após decodificação, para boas relações sinal-ruído (i.e., μ/σ suficientemente elevado) à entrada do decodificador, é majorada por [31,46]:

$$P_{db} < Q\left(\sqrt{d_f} \frac{\mu}{\sigma}\right) e^{\frac{d_f \mu^2}{2\sigma^2}} \frac{\partial T(D, N)}{\partial N} \bigg|_{\substack{N=1 \\ D=e^{-\frac{\mu^2}{2\sigma^2}}}} \quad (4.13)$$

em que μ e σ são dados pelas expressões (3.30) e (3.31), e calculados para o comprimento (em bits) da marca codificada.

Resultados

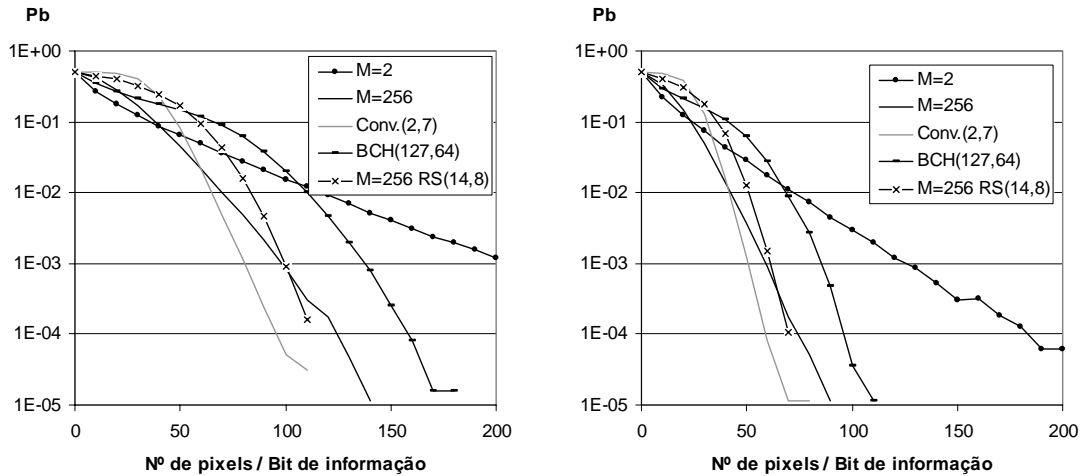
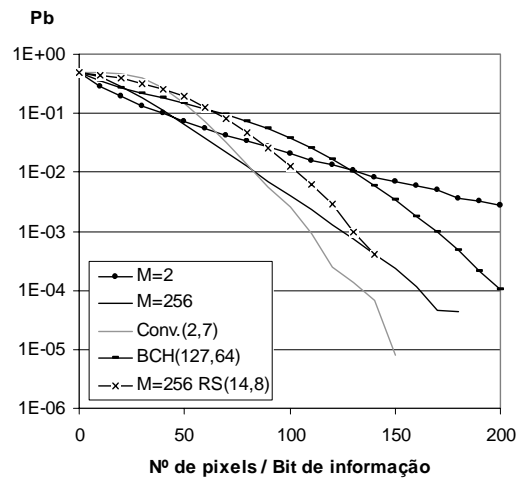
Na figura 4.12 apresentam-se os resultados obtidos para todas as imagens de teste, quando se utiliza um código convolucional binário com taxa de codificação $\frac{1}{2}$, comprimento restritivo 7 e decodificação de *Viterbi*. Os resultados experimentais foram obtidos para os dois tipos de decisão: suave (SD) e dura (HD). Para as curvas teóricas utilizou-se a expressão (4.13), válida para decisão suave.

Como se pode observar, o desempenho é substancialmente melhorado ao ser utilizado um código convolucional, sobretudo se a decodificação for feita com decisão suave.

4.4 Comparação do desempenho

Obtidos os resultados referentes a cada um dos tipos de códigos, interessa agora comparar o desempenho demonstrado por cada um destes. Para facilitar esta comparação, apresentam-se na figura 4.13 as curvas experimentais obtidas para os diversos códigos e para as três imagens de teste.

Com sinalização binária ($M=2$), o código que exhibe melhor desempenho é sem dúvida o código convolucional com decodificação de *Viterbi* e decisão suave. O código BCH apresenta um desempenho intermédio entre os casos binário simples e binário convolucional. De realçar que a utilização de códigos BCH conduz a melhor desempenho que a utilização de sinalização multinível sem codificação, para valores baixos de M (pelo menos para $M \leq 16$). Para se obter um desempenho equivalente ao dos códigos convolucionais, através do uso de sinalização multinível sem codificação, seria necessário um valor de M muito elevado, o que aumentaria bastante a complexidade do desmodulador.

*Lena**Mandrill*

02

Figura 4.13 – Resultados experimentais – P_b vs. N° de Pixels / Bit de informação.

O código RS com 256 níveis de sinalização apresenta um desempenho próximo do demonstrado pelo código convolucional verificando-se até, em alguns dos testes efectuados, que a utilização deste código conduz à inexistência de erros para valores de nº de pixels / bit de informação mais baixos do que no caso da utilização do código convolucional.

Numa implementação real do sistema, a opção pelo tipo de codificação utilizada, bem como pelo número de níveis a utilizar na modulação, deverá ter em conta a complexidade correspondente a cada caso. Em relação aos códigos estudados, a complexidade dos códigos de bloco é inferior à dos códigos convolucionais. Em presença de um sistema multinível, a complexidade do desmodulador cresce linearmente com o número de níveis de sinalização.

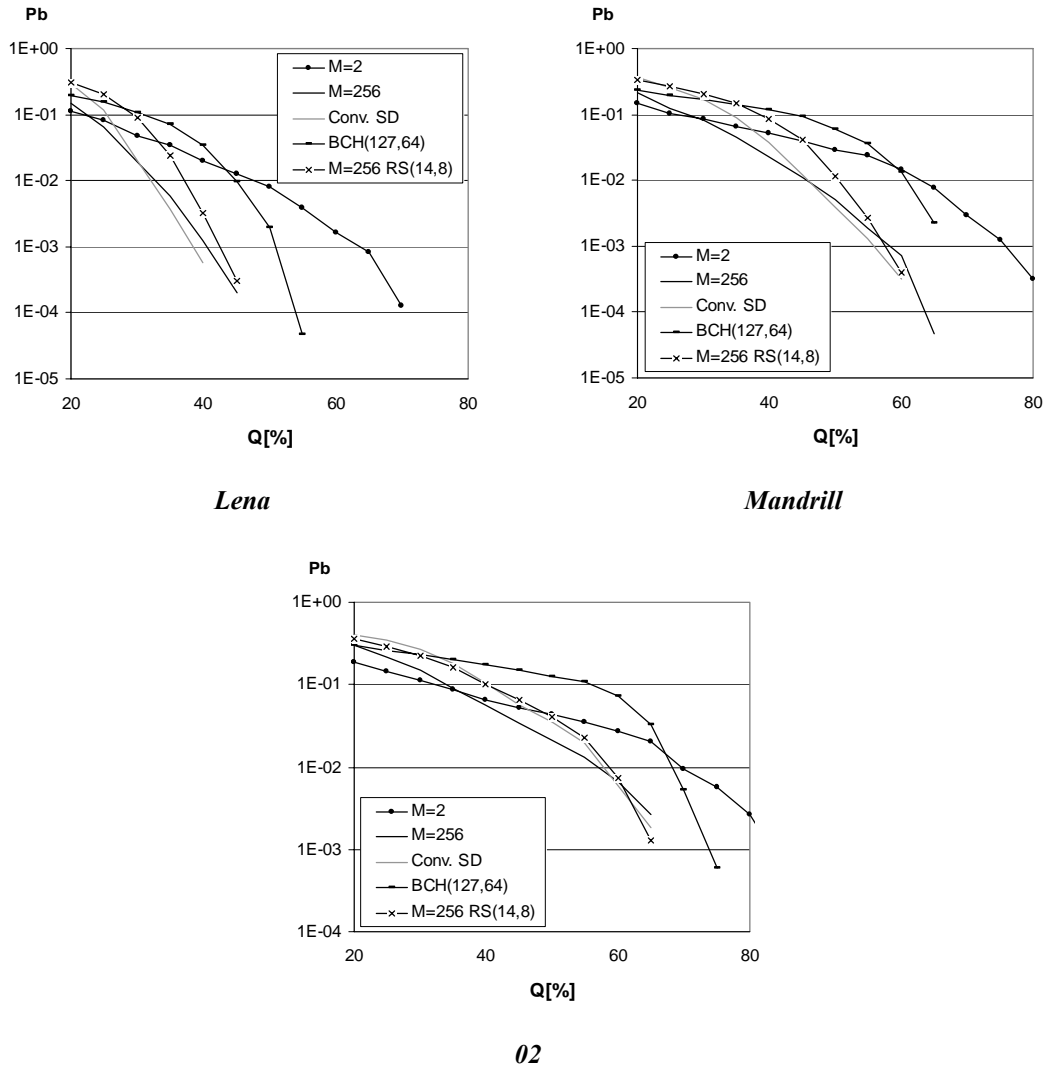


Figura 4.14 – Resultados experimentais em presença de compressão JPEG.

Tendo em conta que os casos que conduziram a melhor desempenho nos testes realizados foram os códigos RS(14,8) com $M=256$ e o código convolucional binário, poder-se-ia optar entre um desmodulador mais complexo e um decodificador mais simples, para o primeiro caso, ou o contrário, para o segundo caso.

4.5 Resultados em presença de compressão JPEG

Na figura 4.14 representam-se as curvas de probabilidade de erro de bit em função do factor de qualidade da compressão JPEG, que visam uma comparação do desempenho dos códigos de correcção de erro estudados em presença de compressão JPEG.

Cada ponto de cada curva resulta de 1000 inserções / extracções de marcas-de-água, excepto para o caso $M=2$, em que esse número é de 250. As marcas-de-água (com comprimento de 64

bits) foram geradas aleatoriamente. As forças de inserção da marca foram ajustadas para cada imagem: 0.4 na imagem *Lena*, 0.2 na imagem *02* e 0.1 na imagem *Mandrill*, à semelhança do que foi efectuado no capítulo 3 e pelas razões aí apontadas.

Por observação dos gráficos verifica-se que, para o caso binário, a curva correspondente à utilização do código convolucional é a que exhibe melhor desempenho. O caso multinível, com $M=256$, é pouco sensível à utilização (ou não) do código RS. No entanto, pode observar-se que o declive da curva correspondente à utilização de códigos tende a acentuar-se no sentido descendente, o que permite especular sobre um melhor desempenho (face ao multinível simples) para as probabilidades de erro mais baixas, onde não existem resultados devido ao número de testes ter sido insuficiente.

4.6 Resultados em presença de ruído branco gaussiano e cortes

Outro teste foi realizado com a finalidade de testar o desempenho da codificação de canal em presença de ruído branco e gaussiano. À semelhança do efectuado no capítulo 3, obteve-se a probabilidade de erro de bit na extracção em função do desvio padrão do ruído – σ_r – e para os diferentes tipos de códigos estudados.

O número de testes realizados, para cada valor de σ_r , foi de 250 para o caso $M=2$ sem codificação e de 500 para os restantes casos. Utilizaram-se marcas-de-água aleatórias com comprimento de 256 bits e a força de inserção foi ajustada para 0.4 em todas as imagens.

Como se pode constatar por observação da figura 4.15, o código convolucional é o que conduz ao melhor desempenho na extracção da marca, apresentando a codificação RS com $M=256$ um desempenho ligeiramente inferior. Na presença de cortes (figura 4.16), as diferenças de desempenho entre os diversos tipos de códigos não são tão nítidas, embora qualquer um deles conduza a uma redução na probabilidade de erro de bit da marca extraída, face à utilização de sinalização multinível simples.

4.7 Resultados no domínio da frequência

O modo como é aplicada a codificação de canal a marcas-de-água com inserção no domínio da frequência é em tudo idêntico ao apresentado para inserção no domínio espacial. Toda a análise teórica referente ao desempenho dos códigos realizada para o domínio espacial, é também válida para o domínio da frequência, desde que os parâmetros que caracterizam o canal (μ e σ) sejam estimados de acordo com o apresentado no capítulo anterior (expressões (3.71) e (3.72)).

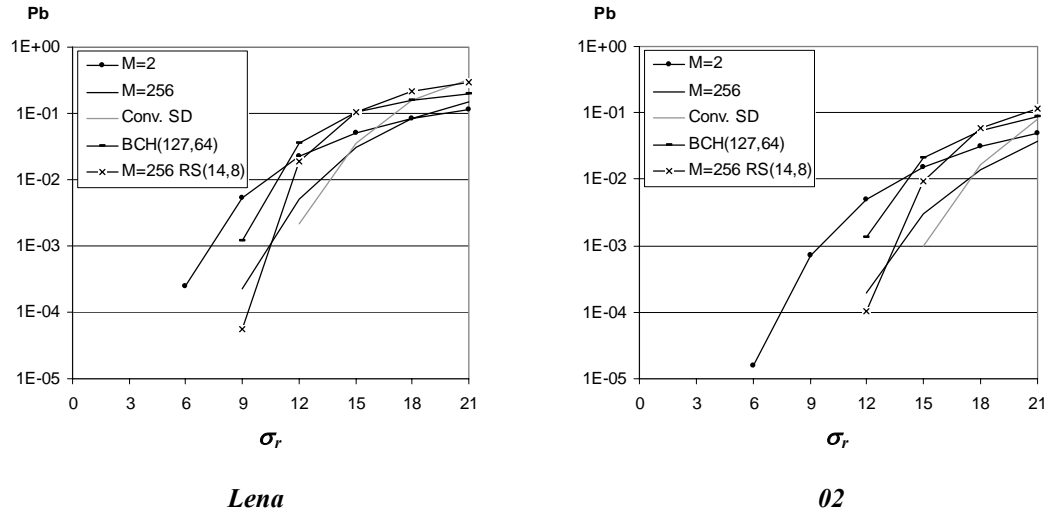


Figura 4.15 – Resultados experimentais em presença de ruído branco gaussiano.

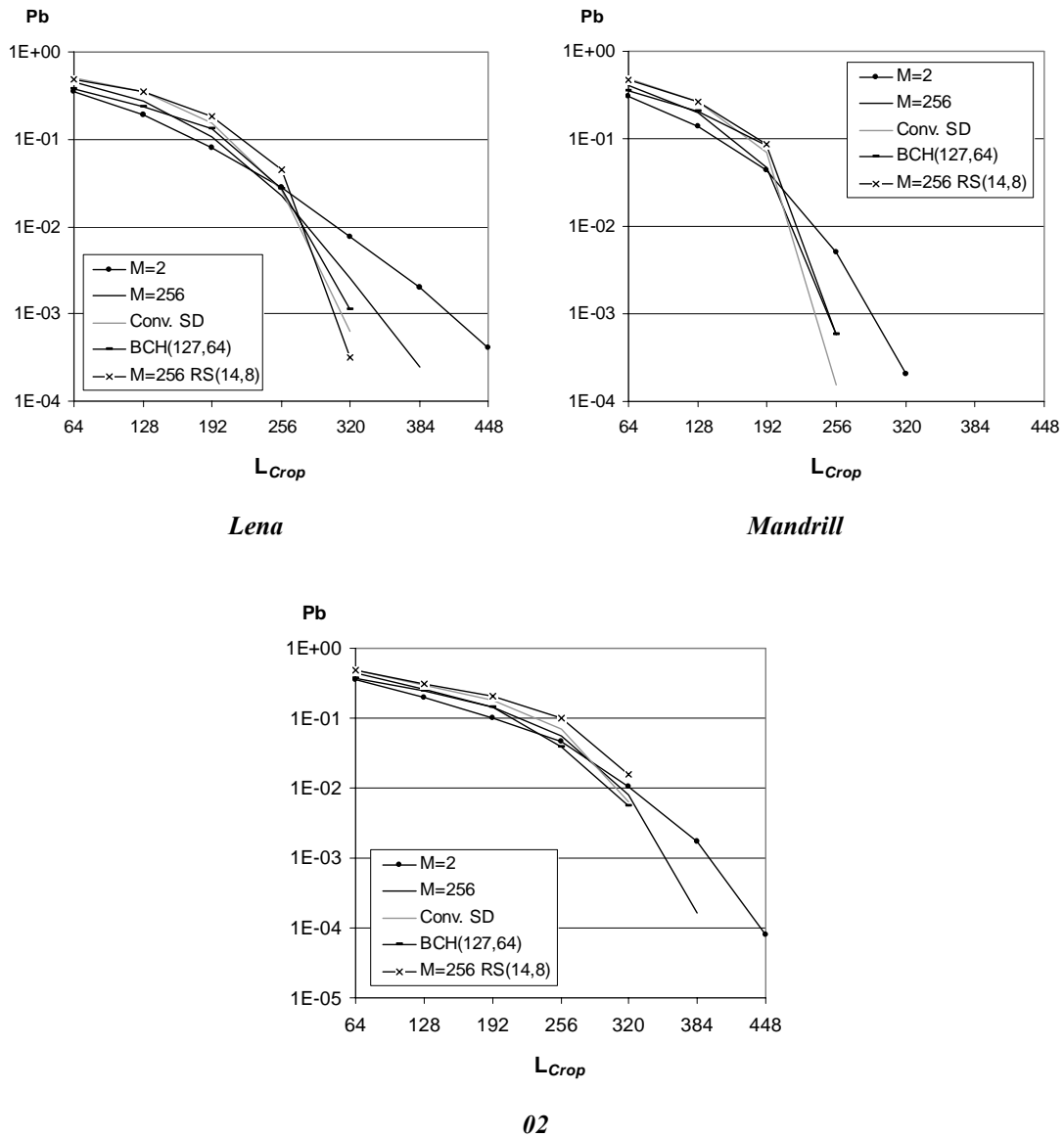


Figura 4.16 – Resultados experimentais em presença de cortes.

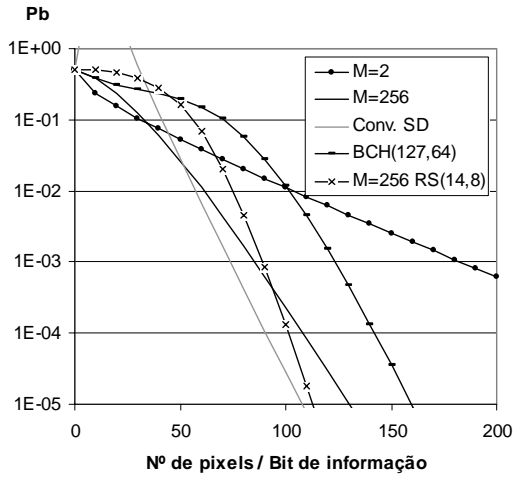
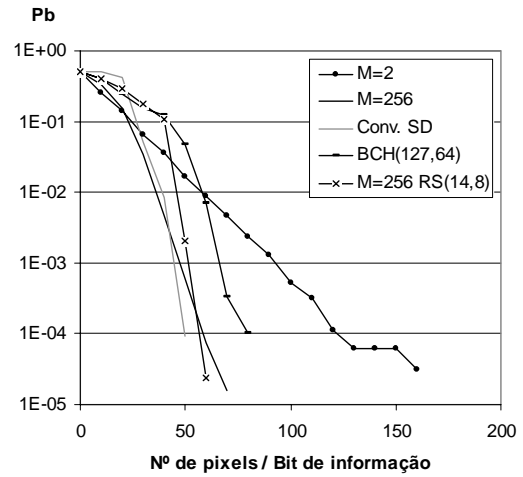
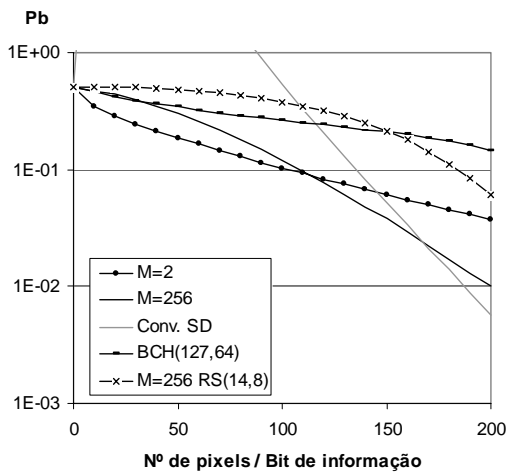
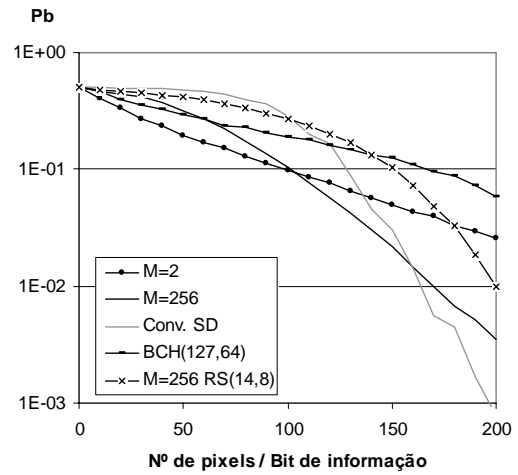
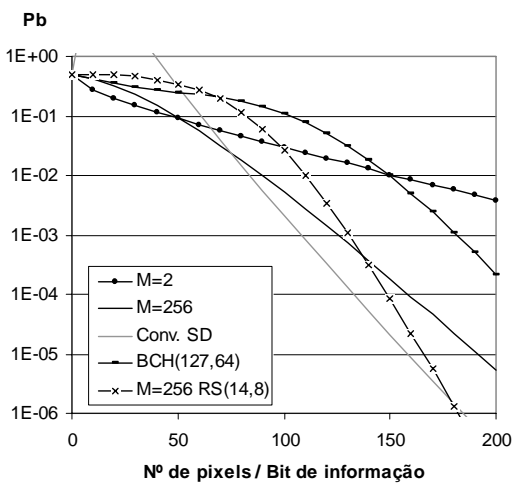
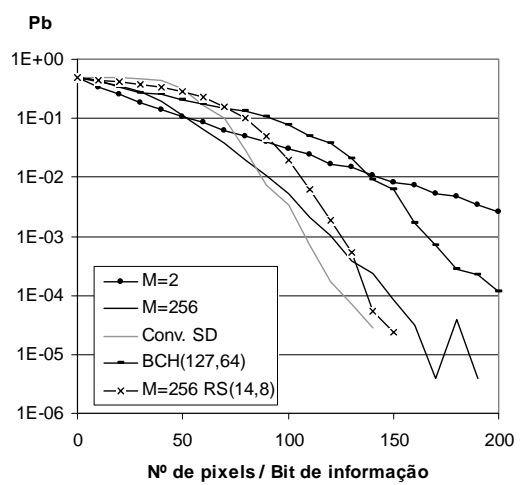
*Lena (Teo.)**Lena (Exp.)**Mandrill (Teo.)**Mandrill (Exp.)**02 (Teo.)**02 (Exp.)*

Figura 4.17 – Resultados teóricos e experimentais (domínio da frequência).

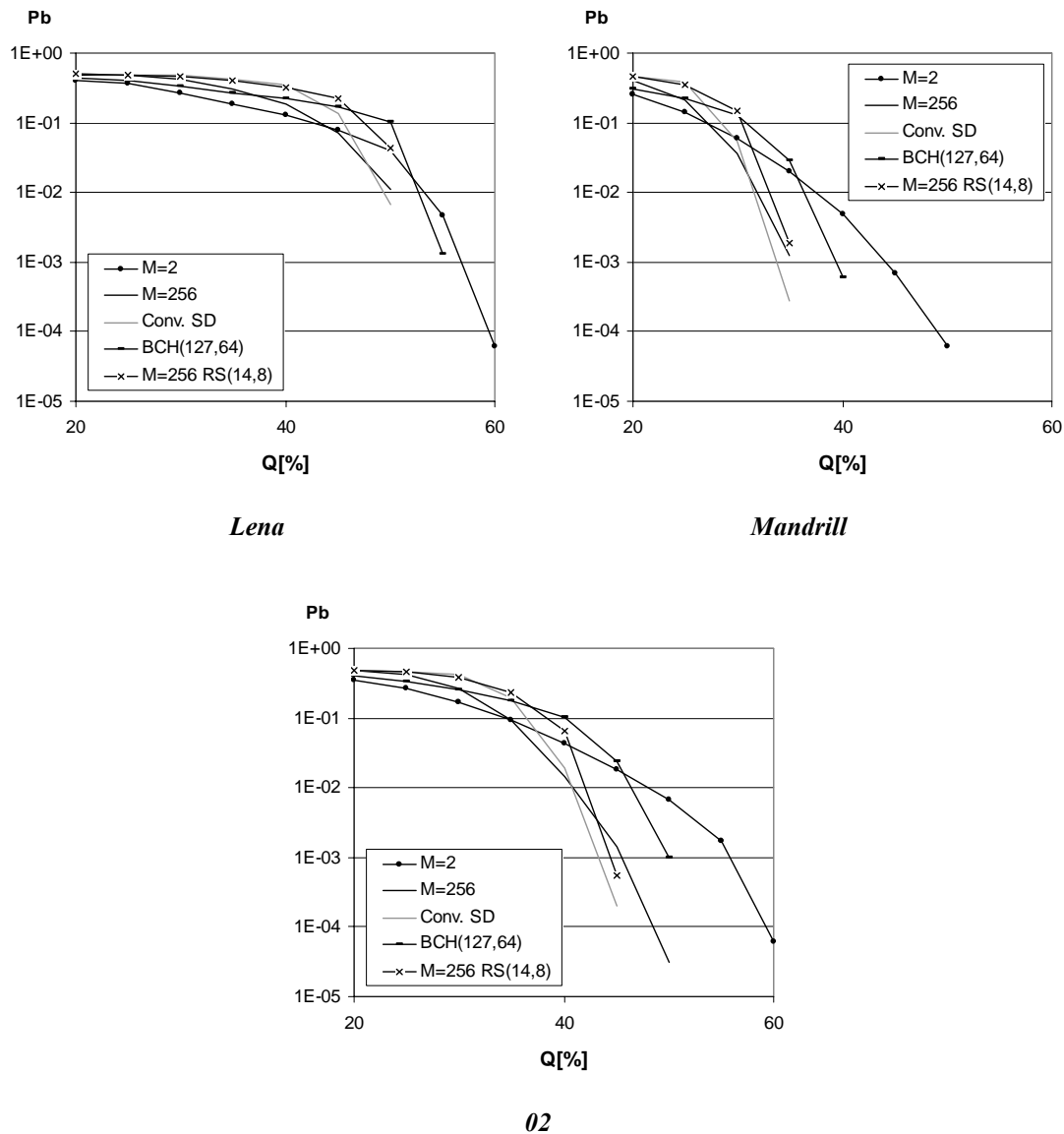


Figura 4.18 – Resultados experimentais em presença de compressão JPEG (domínio da frequência).

À semelhança do que foi feito nas secções anteriores apresenta-se, na figura 4.17, a evolução teórica e experimental das curvas de probabilidade de erro de bit em função do número de pixels por bit de informação útil, para os vários códigos em análise. A força de inserção foi ajustada para o valor 1.0.

Observa-se que, para as imagens *Lena* e *02*, a utilização quer de sinalização binária com codificação convolucional, quer de 256 níveis de sinalização com codificação RS, conduzem claramente a um melhor desempenho do que nos restantes casos. Na imagem *Mandrill* este facto não é tão evidente no caso da codificação RS mas, atendendo ao declive das curvas, é de esperar que o seja para valores mais elevados do número de pixels por bit de informação útil.

Observa-se também que, na imagem *Lena*, os resultados previstos teoricamente encontram-se bastante afastados dos resultados obtidos experimentalmente, embora qualitativamente o resultado da comparação entre os diversos códigos seja o mesmo.

Para concluir esta secção, apresentam-se na figura 4.18 resultados experimentais obtidos quando em presença de compressão JPEG. Neste teste, a força de inserção foi ajustada para 1.5, mantendo-se os restantes parâmetros de teste descritos na secção 4.5. De novo se verifica o melhor desempenho demonstrado pelos códigos convolucional binário e RS com 256 níveis de sinalização.

4.8 Considerações finais

Ao longo deste capítulo ficou demonstrado que as marcas-de-água podem beneficiar com a utilização de codificação de canal, já que esta conduz a probabilidades de erro de bit no processo de extracção da marca-de-água mais baixas do que as verificadas sem essa forma de codificação.

De entre os códigos de correcção de erros estudados para o caso binário, o melhor desempenho foi demonstrado pelo código convolucional com descodificação de *Viterbi* e decisão suave. O código de bloco binário BCH apresenta resultados superiores ao do caso binário simples, sem codificação, mas aquém dos demonstrados pelo código convolucional. De acordo com a literatura sobre códigos de correcção de erros, esta conclusão seria expectável em presença de um canal corrompido por ruído branco e gaussiano, o que se enquadra no modelo analisado para extracção de marcas-de-água em imagens não sujeitas a manipulações. Introduzindo ruído não gaussiano através de compressão JPEG verificou-se que, em termos qualitativos, a relação entre os desempenhos atingidos pelo uso de códigos BCH ou pelo uso de códigos convolucionais binários é semelhante ao demonstrado no caso de ruído gaussiano.

Para sinalização multinível ($M > 2$) que, como se verificou no capítulo 3, conduz a melhor desempenho que o caso binário, é possível diminuir os erros na extracção da marca através da utilização de códigos de bloco não-binários (*Reed-Solomon*). Neste caso, quando a modulação é realizada utilizando 256 níveis de sinalização, o sistema apresenta resultados análogos aos apresentados pela utilização de códigos convolucionais, com descodificação segundo o algoritmo de *Viterbi* e decisão suave. De salientar que a utilização de códigos BCH conduz a melhores resultados que a sinalização multinível simples (i.e., sem códigos de correcção) quando o número de níveis de sinalização não é suficientemente elevado – nos testes realizados, e na ausência de codificação de canal, apenas a utilização de 256 níveis de sinalização permitiu

atingir melhor desempenho. No entanto e para os códigos estudados, se se associar a codificação RS à modulação *M-ária* atinge-se um desempenho superior ao exibido pelo caso binário com códigos BCH. As conclusões retiradas desta análise aplicam-se a ambos os espaços utilizados para inserção da marca-de-água.

Por último, estas conclusões são válidas em presença de compressão JPEG, adição de ruído branco e gaussiano, ou cortes sobre a imagem.

Capítulo 5

Técnicas de combinação de sinal

5.1 Introdução

A tecnologia das marcas-de-água em vídeo pode ser analisada como um sistema multi-canal, se em cada trama for inserida a mesma marca-de-água e se cada trama puder ser considerada como um canal independente. Nesta situação, a detecção da marca-de-água pode ser melhorada considerando simultaneamente um grupo de tramas consecutivas, aplicando técnicas de combinação de sinal, amplamente utilizadas no domínio das comunicações rádio com diversidade.

A figura 5.1 esquematiza a combinação de sinal. Admitindo que cada conjunto de tramas é composto por J tramas e que x_1, x_2, \dots, x_J são os sinais (símbolos ou saídas dos detectores) extraídos em cada uma das tramas do conjunto, a marca-de-água detectada será o resultado de uma combinação desses sinais. Colocam-se desde logo algumas questões fundamentais – que sinais $x_i, i \in \{1, \dots, J\}$ combinar e como os combinar de forma a minimizar a probabilidade de erro de bit da marca extraída.

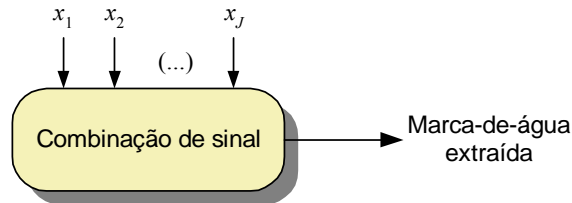


Figura 5.1 – Esquema de combinação de sinal.

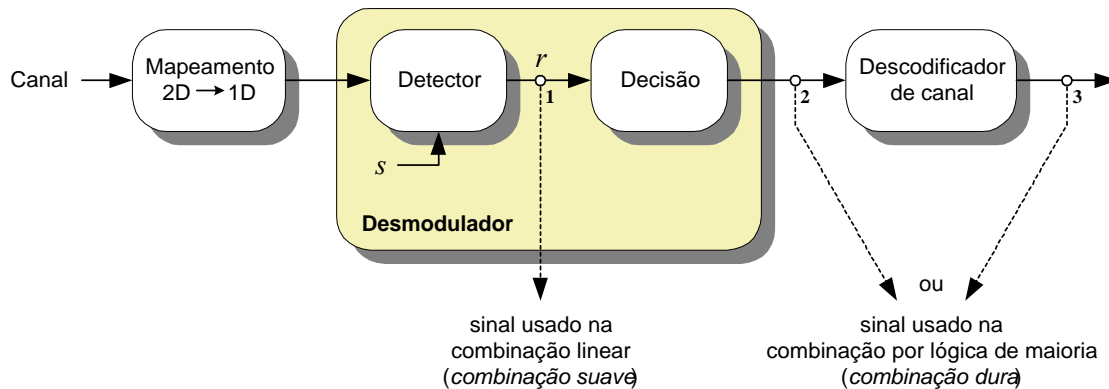


Figura 5.2 – Esquema de extração da marca com saídas utilizadas na combinação de sinal.

Tendo em conta o sinal a utilizar, as técnicas de combinação podem ser divididas em dois grupos distintos:

- *Combinação dura* – se a marca final resultar da combinação dos diversos símbolos, extraídos trama a trama;
- *Combinação suave* – se a marca final resultar da combinação das saídas dos detectores, numa etapa prévia à desmodulação multinível.

O primeiro grupo pode ainda ser dividido em dois subgrupos, consoante os símbolos considerados sejam os detectados à saída do descodificador multinível – combinação com símbolos codificados – ou os detectados à saída do descodificador de canal – combinação com símbolos descodificados.

Tendo em conta a forma de combinação do sinal, duas estratégias foram estudadas:

- *Lógica de maioria* – se os símbolos recebidos trama a trama são combinados de modo a que o símbolo final extraído seja o que ocorre mais vezes;
- *Combinação linear* – se o sinal utilizado para decidir a marca inserida resultar de uma soma pesada dos sinais extraídos em cada canal.

No esquema apresentado na figura 5.2 destacam-se os pontos do sistema de extração da marca nos quais poderá ser feita a combinação de sinal. Na estratégia *combinação linear*, combinam-se os sinais obtidos à saída dos detectores – r – (nó 1). Na estratégia *lógica de maioria* são possíveis duas soluções: combinação dos símbolos codificados (nó 2), ou combinação dos símbolos obtidos após descodificação (nó 3).

A estratégia *lógica de maioria* foi utilizada em [8] e [19], embora em esquemas de assinatura distintos do considerado neste trabalho. Com efeito, tanto em [8] como em [19] a inserção da marca é orientada ao bloco de imagem (de dimensão 8×8 pixels) e efectuada no domínio da DCT [8], ou pixel [19]. A sinalização é, em ambos os casos, binária e apenas [19] utiliza codificação de canal (códigos BCH e convolucionais). Em [8] analisam-se valores de J de 25 ou 50 tramas, sobre vídeo comprimido a 6 Mbit/s, enquanto que em [19] J é, em cada momento, igual ao número de tramas já processadas. A estratégia *combinação linear* foi também utilizada em [19], num esquema muito simples em que é dado o mesmo peso (constante e unitário) a cada canal.

Neste capítulo, para além de se avaliar o desempenho das técnicas de combinação de sinal propostas em [8] e [19], quando usadas em conjunto com um sistema de marcas-de-água baseado em espalhamento de espectro, efectua-se o estudo analítico de ambas as técnicas. No caso da estratégia *lógica de maioria*, esta análise permite determinar qual o ponto, no sistema de recepção, em que deverá ter lugar a combinação. Para a *combinação linear*, a análise teórica é efectuada no sentido de obter os pesos que maximizam a relação sinal-ruído do sinal combinado – *pesos óptimos* – e de avaliar o efeito da utilização de pesos não óptimos, de que é exemplo a utilização de *pesos constantes e unitários*. O estudo analítico é complementado com uma avaliação experimental dos vários métodos, utilizando três sequências de vídeo CCIR-601, codificadas em MPEG-2 a 2, 4 e 6 Mbit/s.

5.2 Combinação com lógica de maioria

A estratégia de combinação utilizando lógica de maioria inclui-se no grupo de técnicas *combinação dura* pois os sinais a combinar são símbolos. O símbolo considerado como correcto é o símbolo que ocorrer mais vezes dentro do conjunto de canais considerado.

Quando esta estratégia é aplicada a uma sequência de vídeo, cada trama é considerada um canal onde é enviado um conjunto de símbolos, resultantes da modulação multinível e, eventualmente, codificação de canal, da marca-de-água. O conjunto de símbolos final, a partir do qual é

recuperada a sequência de bits correspondente à marca-de-água, é obtido utilizando lógica de maioria sobre os símbolos detectados em cada trama.

Considerando que a combinação é feita utilizando um conjunto de J tramas, o i -ésimo símbolo – X_i – é obtido de acordo com:

$$X_i = \text{moda}(x_{i1}, x_{i2}, \dots, x_{iJ}), i \in \{1..N\}, \quad (5.1)$$

em que N é o número de símbolos inseridos (e extraídos) em cada trama, e x_{ij} é o i -ésimo símbolo extraído na trama j .

Na figura 5.3 apresenta-se um exemplo da aplicação desta técnica, para $J=3$ e $N=5$, onde $\mathbf{x}_i = \{x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}\}$, $i \in \{1..3\}$, designa o conjunto de símbolos extraídas na trama i e $\mathbf{X} = \{X_1, X_2, X_3, X_4, X_5\}$ designa o conjunto de símbolos obtidos por lógica de maioria. Admite-se estar em presença de um sistema multinível com $M=4$ e com símbolos A^1 a A^4 .

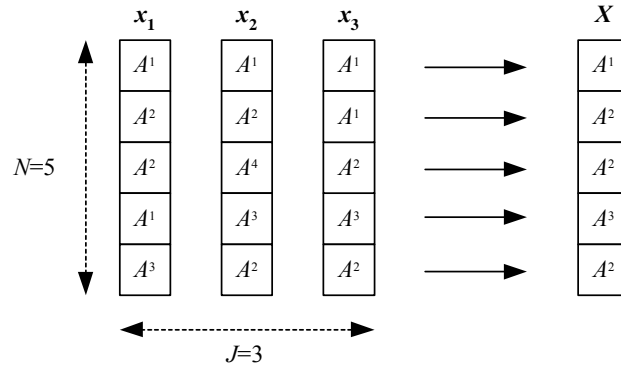


Figura 5.3 – Exemplo de aplicação de lógica de maioria.

Estabelecidos os princípios de funcionamento do sistema de extracção utilizando lógica de maioria é necessário verificar, caso sejam utilizados códigos correctores de erro, qual das seguintes hipóteses conduzirá a melhor desempenho:

1. Extracção dos símbolos inseridos e aplicação da lógica de maioria sobre os símbolos codificados; decodificação do conjunto de símbolos obtidos por combinação. Neste caso, a lógica de maioria é aplicada no nó **2** da figura 5.2;
2. Extracção dos símbolos inseridos e sua decodificação, trama a trama; utilização de lógica de maioria sobre os símbolos decodificados. Neste caso, a lógica de maioria é aplicada no nó **3** da figura 5.2.

Para tal, considere-se o caso mais simples em que $M=2$ e suponha-se que é utilizado um código de bloco binário BCH(n, k), com capacidade de correcção t . Seja P_b a probabilidade de erro de bit da marca à saída do canal (i.e., antes da descodificação de canal); admita-se que o seu valor é aproximadamente o mesmo em todas as tramas e que existe independência entre erros. Supondo que se adopta a hipótese 1, a probabilidade de erro de bit – P_{bf} – após aplicação de lógica de maioria sobre um grupo de J tramas é:

$$P_{bf} = \sum_{i=\frac{J+1}{2}}^J C_i^J P_b^i (1-P_b)^{J-i}, \quad (5.2)$$

ou, admitindo que $P_b \ll 1$:

$$P_{bf} \approx C_{\frac{J+1}{2}}^J P_b^{\frac{J+1}{2}}. \quad (5.3)$$

Como se viu anteriormente (expressão (4.9)) a probabilidade de erro de bit – P_{db} – à saída do descodificador de canal é majorada por:

$$P_{db} \leq \frac{1}{n} \sum_{i=t+1}^n \min(i+t, n) C_i^n P^n (1-P)^{n-i}, \quad (5.4)$$

em que P representa a probabilidade de erro de bit à entrada do descodificador. Substituindo P na equação (5.4) pelo resultado obtido em (5.3) obtém-se, para a probabilidade de erro de bit – p_1 – da marca extraída:

$$p_1 \leq \frac{1}{n} \sum_{i=t+1}^n \min(i+t, n) C_i^n P_{bf}^i (1-P_{bf})^{n-i}, \quad (5.5)$$

ou,

$$\begin{aligned} p_1 &\leq \frac{1}{n} \sum_{i=t+1}^n \min(i+t, n) C_i^n \left(C_{\frac{J+1}{2}}^J P_b^{\frac{J+1}{2}} \right)^i \left(1 - C_{\frac{J+1}{2}}^J P_b^{\frac{J+1}{2}} \right)^{n-i} \\ &\approx \frac{1}{n} (2t+1) C_{t+1}^n \left(C_{\frac{J+1}{2}}^J \right)^{t+1} P_b^{\frac{(J+1)(t+1)}{2}}. \end{aligned} \quad (5.6)$$

Na hipótese 2, a descodificação é feita trama a trama e a combinação utiliza os símbolos descodificados. A probabilidade de erro de bit – p_2 – do sinal combinado é, neste caso:

$$p_2 \approx C_{\frac{J+1}{2}}^J P_{db}^{\frac{J+1}{2}}, \quad (5.7)$$

em que P_{db} é a probabilidade de erro de bit em cada trama, à saída do decodificador de canal. Substituindo P_{db} pelo resultado (5.4) e tendo em conta que neste caso $P=P_b$, obtém-se:

$$p_2 \approx C_{\frac{J+1}{2}}^J \left[\frac{1}{n} \sum_{i=t+1}^n \min(i+t, n) C_i^n P_b^i (1-P_b)^{n-i} \right]^{\frac{J+1}{2}}, \quad (5.8)$$

ou, admitindo novamente que $P_b \ll 1$:

$$p_2 \approx C_{\frac{J+1}{2}}^J \left(\frac{2t+1}{n} C_{t+1}^n \right)^{\frac{J+1}{2}} P_b^{\frac{(t+1)(J+1)}{2}}. \quad (5.9)$$

Para concluir sobre qual das hipóteses é mais vantajosa, considere-se o quociente p_1/p_2 :

$$\frac{p_1}{p_2} = \frac{\frac{2t+1}{n} C_{t+1}^n \left(C_{\frac{J+1}{2}}^J \right)^{t+1} P_b^{\frac{(t+1)(J+1)}{2}}}{\left(\frac{2t+1}{n} C_{t+1}^n \right)^{\frac{J+1}{2}} C_{\frac{J+1}{2}}^J P_b^{\frac{(t+1)(J+1)}{2}}} = \left(C_{\frac{J+1}{2}}^J \right)^t \left(\frac{2t+1}{n} C_{t+1}^n \right)^{\frac{1-J}{2}}. \quad (5.10)$$

Facilmente se verificará que, para valores típicos de J , n e t se tem $p_1 \ll p_2$, concluindo-se ser mais vantajoso utilizar a hipótese 1.

5.3 Combinação linear

O combinador linear combina as saídas dos detectores relativos a cada trama – *combinação suave* – realizando uma soma pesada das diferentes saídas, com o intuito de dar maior ênfase aos canais (tramas) com melhor relação sinal-ruído.

Na figura 5.4 apresenta-se o diagrama de blocos do combinador linear para o caso binário ($M=2$)²⁸, onde se utilizou a seguinte simbologia:

- N : número de símbolos inseridos em cada trama;
- J : número de canais (tramas) combinados;
- r_i^j : saída do detector relativo ao canal i , $i \in [1..J]$, e símbolo j , $j \in \{1..N\}$;
- c_i : peso associado ao canal i .

²⁸ Se $M > 2$, devem ser considerados $M/2$ combinadores lineares, correspondentes às $M/2$ saídas dos detectores existentes no desmodulador multinível

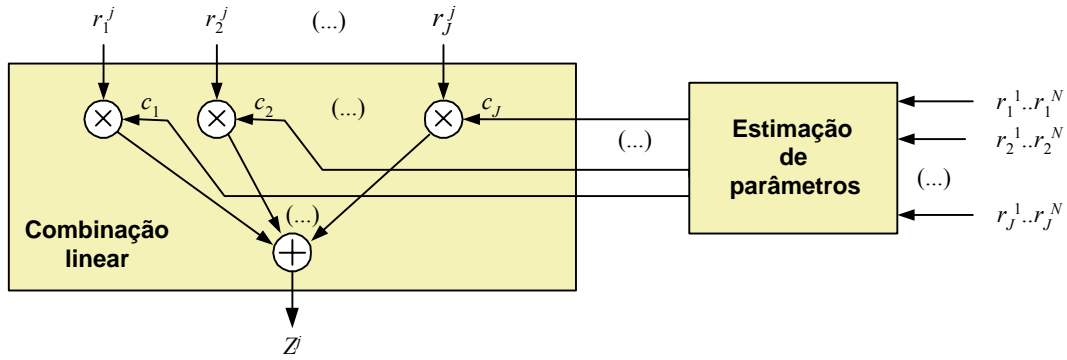


Figura 5.4 – Diagrama de blocos do sistema de combinação linear de sinal (para $M=2$).

Como referido anteriormente, os coeficientes r_i^j resultam da soma de um número elevado de variáveis aleatórias com a mesma distribuição pelo que, aplicando o teorema do limite central, podem ser considerados como variáveis aleatórias com distribuição Gaussiana. Sendo assim, os valores r_i^j podem ser modelizados como a saída de um canal com ruído aditivo e Gaussiano (AWGN), $r_i^j = a_i^j b^j + n_i^j$, $j \in \{1..N\}$, $i \in \{1..J\}$, onde os coeficientes n_i^j são variáveis aleatórias Gaussianas independentes, de média nula. Admite-se o caso mais simples de modulação binária antípodal, pelo que $b^j = -1$, se $bit_j = 0$; $b^j = 1$, se $bit_j = 1$, onde bit_j representa o j -ésimo bit da marca codificada (i.e., à saída do codificador de canal).

O sistema de combinação de sinal é constituído por dois sub-sistemas: o estimador de parâmetros e o combinador de canal. No que se segue, irão ser analisados três tipos de combinação: i) *combinação óptima*; ii) *combinação quase óptima*; iii) *combinação com pesos iguais e constantes*.

5.3.1 Combinação óptima

O modelo matemático do combinador linear pode ser escrito como:

$$Z^j = \sum_{i=1}^J c_i r_i^j = \sum_{i=1}^J c_i a_i^j b^j + \sum_{i=1}^J c_i n_i^j. \quad (5.11)$$

O objectivo do algoritmo de combinação é maximizar a relação sinal-ruído do sinal à saída do combinador – SNR_Z – definida por

$$SNR_Z = \frac{E[Z^j]^2}{\text{var}[Z^j]} = \frac{(\sum_{i=1}^J c_i E[a_i^j])^2}{\sum_{i=1}^J c_i^2 \text{var}(n_i^j)}. \quad (5.12)$$

Considerando que para o mesmo canal (mesma trama), a média e variância da saída dos detectores é idêntica para todos os símbolos (processo estacionário), tem-se

$$\begin{aligned} \text{var}(n_i^j) &= \text{var}(n_i) = \sigma_i^2 \\ E[a_i^j] &= a_i, \end{aligned} \quad (5.13)$$

pelo que

$$SNR_z = \frac{(\sum_{i=1}^J c_i a_i)^2}{\sum_{i=1}^J c_i^2 \sigma_i^2}. \quad (5.14)$$

Utilizando a desigualdade de *Shwarz*

$$(\sum_{i=1}^J (c_i \sigma_i) (\frac{a_i}{\sigma_i}))^2 \leq \sum_{i=1}^J c_i^2 \sigma_i^2 \cdot \sum_{i=1}^J (\frac{a_i}{\sigma_i})^2, \quad (5.15)$$

e dividindo ambos os termos de (5.14) por $\sum_{i=1}^J c_i^2 \sigma_i^2$ obtém-se, como limite superior de SNR_z

$$SNR_z \leq \sum_{i=1}^J (\frac{a_i}{\sigma_i})^2. \quad (5.16)$$

Os pesos para os quais se atinge o valor máximo da relação SNR_z – pesos ótimos – são:

$$c_i = \frac{a_i}{\sigma_i^2}, \quad (5.17)$$

como se pode provar por substituição de (5.17) em (5.14).

Uma vez que os parâmetros que caracterizam quer o sinal, quer o ruído, não são geralmente conhecidos, o limite superior da SNR não é normalmente atingido na prática. Quando se trabalha com parâmetros desconhecidos, é usual efectuar estimativas destes parâmetros a partir dos dados disponíveis. Nesta situação, pode-se obter uma estimativa dos pesos utilizando (5.17) mas com as estimativas \hat{a}_i e $\hat{\sigma}_i^2$ dos parâmetros envolvidos. Esta questão será tratada em detalhe na secção 5.3.5.

Utilizando modulação binária ($M=2$) e antipodal, a probabilidade de erro de bit à saída do combinador – P_b^{LO} – é dada por:

$$P_b^{LO} = Q(\sqrt{SNR_Z}) = Q\left(\sqrt{\sum_{i=1}^J \left(\frac{a_i}{\sigma_i}\right)^2}\right) \leq Q(\max(\frac{a_i}{\sigma_i})) \leq P_b. \quad (5.18)$$

Admitindo que a estatística (média e variância) do sinal à saída dos detectores é a mesma para todos os canais – $a_i = a$ e $\sigma_i = \sigma$, \forall_i – tem-se, para $J \geq 2$ e $a/\sigma \geq 1$:

$$P_b^{LO} = Q\left(\sqrt{J} \frac{a}{\sigma}\right) \ll P_b = Q\left(\frac{a}{\sigma}\right), \quad (5.19)$$

o que prova o interesse em se utilizar esta forma de combinação de sinal.

5.3.2 Combinação quase óptima

Numa abordagem intuitiva (e que foi inicialmente seguida neste trabalho) à questão da combinação de vários canais, poder-se-á ser “tentado” a utilizar pesos proporcionais à relação sinal-ruído (ou à sua raiz quadrada) de cada canal. Neste caso, $c_i = a_i / \sigma_i$ relação que, após substituição em (5.14), conduz a:

$$SNR_Z = \frac{\left(\sum_{i=1}^J \frac{a_i^2}{\sigma_i}\right)^2}{\sum_{i=1}^J a_i^2}, \quad (5.20)$$

e a uma probabilidade de erro de bit – P_b^{LQO} – para modulação binária antípodal de

$$P_b^{LQO} = Q\left(\frac{\sum_{i=1}^J \frac{a_i^2}{\sigma_i}}{\sqrt{\sum_{i=1}^J a_i^2}}\right). \quad (5.21)$$

Admitindo, mais uma vez, que a estatística (média e variância) do sinal à saída dos detectores é a mesma para todos os canais e que $J \geq 2$ e $a/\sigma \geq 1$, a relação (5.19) será também verificada para este caso.

5.3.3 Combinação com pesos iguais e constantes

As técnicas de combinação atrás descritas requerem o cálculo, para cada trama, dos parâmetros estatísticos (média e variância) das saídas dos detectores. Tal poderá não ser desejável, quer por razões computacionais, quer por não ser possível efectuar uma boa estimativa dos parâmetros

necessários, correndo-se o risco de estar a penalizar bons canais e a valorizar canais com baixa SNR. Como alternativa mais simples, poder-se-ão utilizar pesos iguais e unitários, i.e., $c_i = 1, \forall_i$. Substituindo esta relação em (5.14), obtém-se

$$SNR_z = \frac{(\sum_{i=1}^J a_i)^2}{\sum_{i=1}^J \sigma_i^2} \quad (5.22)$$

e uma probabilidade de erro de bit – P_b^{LIC} – para modulação binária antipodal de

$$P_b^{LIC} = Q\left(\frac{\sum_{i=1}^J a_i}{\sqrt{\sum_{i=1}^J \sigma_i^2}}\right). \quad (5.23)$$

A relação expressa em (5.19) é mais uma vez verificada, se se admitirem canais estatisticamente idênticos.

5.3.4 Comparação das várias estratégias de combinação linear

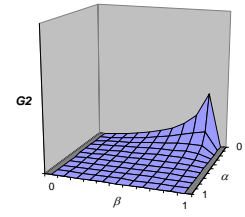
As três técnicas de combinação descritas conduzem a desempenhos idênticos na hipótese de canais com a mesma caracterização estatística (Gaussianos, com média e variância idênticas). Esta hipótese não é verificada na presença de compressão, já que esta conduz uma degradação das tramas (e da marca nelas inseridas) que varia ao longo da sequência vídeo. Tal facto é comprovado pelas figuras 5.9 a 5.11 onde se encontra representada a evolução de a_i e σ_i ao longo de algumas tramas da sequência de vídeo *Stefan*, após compressão MPEG-2. De forma a adquirir uma melhor percepção do desempenho de cada combinador numa situação mais próxima da realidade considere-se que, para uma fracção β dos canais combinados:

$$\begin{aligned} a_i &= \alpha a, \quad \text{com } \alpha \in [0,1] \\ \sigma_i &= \gamma \sigma, \quad \text{com } \gamma \geq 1, \end{aligned} \quad (5.24)$$

e que, para os restantes canais, $a_i = a$ e $\sigma_i = \sigma$. Com os parâmetros α e γ em (5.25) pretende-se contabilizar, respectivamente, uma redução no valor médio do sinal recebido (*fading*) e um aumento na potência de ruído, resultantes do processamento da sequência (e.g., compressão). Tem-se neste caso, para as três estratégias estudadas:

α	$\beta=0.1$			$\beta=0.5$			$\beta=0.9$		
	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
G1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
G2	1.043	1.012	1.000	1.292	1.054	1.001	1.738	1.037	1.001
G3	1.043	1.012	1.000	1.292	1.054	1.001	1.738	1.037	1.001

a)

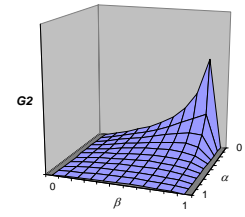


b)

Figura 5.5 – a) Valores de G_1 , G_2 e G_3 para algumas combinações de α, β ;
b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=1$.

α	$\beta=0.1$			$\beta=0.5$			$\beta=0.9$		
	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
G1	1.000	1.001	1.004	1.001	1.010	1.019	1.004	1.020	1.012
G2	1.106	1.066	1.037	1.643	1.267	1.107	2.474	1.185	1.043
G3	1.106	1.064	1.032	1.642	1.254	1.086	2.463	1.162	1.031

a)

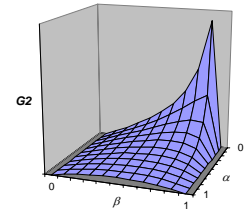


b)

Figura 5.6 – a) Valores de G_1 , G_2 e G_3 para algumas combinações de α, β ;
b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=1.5$.

α	$\beta=0.1$			$\beta=0.5$			$\beta=0.9$		
	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
G1	1.000	1.003	1.010	1.001	1.024	1.050	1.010	1.060	1.041
G2	1.189	1.143	1.105	2.035	1.537	1.291	3.237	1.382	1.123
G3	1.189	1.139	1.094	2.033	1.500	1.229	3.204	1.304	1.078

a)



b)

Figura 5.7 – a) Valores de G_1 , G_2 e G_3 , para algumas combinações de α, β ;
b) Evolução de G_2 com α, β . Em ambos os casos, $\gamma=2$.

$$P_b^{LO} = Q\left(\sqrt{J} \frac{a}{\sigma} \sqrt{(1-\beta) + \beta \left(\frac{\alpha}{\gamma}\right)^2}\right) \quad (5.25)$$

$$P_b^{LQO} = Q\left(\sqrt{J} \frac{a}{\sigma} \frac{(1-\beta) + \beta \frac{\alpha^2}{\gamma}}{\sqrt{(1-\beta) + \beta \alpha^2}}\right) \quad (5.26)$$

$$P_b^{LIC} = Q\left(\sqrt{J} \frac{a}{\sigma} \frac{(1-\beta) + \beta \alpha}{\sqrt{(1-\beta) + \beta \gamma^2}}\right) \quad (5.27)$$

As figuras 5.5 a 5.7 quantificam o desempenho relativo das três estratégias, para alguns valores de α , β e γ . Nesta figuras, utilizaram-se as seguintes designações:

- G_1 – ganho do combinador *ótimo* relativamente ao *quase ótimo*;
- G_2 – ganho do combinador *ótimo* relativamente ao de *pesos iguais e constantes*;
- G_3 – ganho do combinador *quase ótimo* relativamente ao de *pesos iguais e constantes*,

designando-se por *ganho* a redução no valor de \sqrt{SNR} (argumento da função $Q(\cdot)$ nas expressões (5.25) a (5.27)) para o mesmo desempenho.

Estes resultados permitem concluir que as estratégias *pesos ótimos* e *pesos quase ótimos*, apresentam desempenhos semelhantes e, relativamente à estratégia *pesos iguais*, tanto mais eficazes quanto maior a percentagem de tramas severamente degradadas ($\alpha \rightarrow 0$, $\beta \rightarrow 1$), como se espera acontecer para taxas de compressão elevadas.

5.3.5 Estimador de máxima verosimilhança (ML)

Na ausência de conhecimento *a priori* das estatísticas relativas a a_i e σ_i^2 , necessárias para os combinadores *ótimo* e *quase ótimo*, o estimador ML (*Maximum-likelihood*) conduz à melhor estimativa para estes parâmetros [44].

Se em cada um dos J canais se efectuarem N medidas, supostas independentes, a f.d.p. conjunta é dada por:

$$p(\mathbf{r} | \vec{a}, \vec{b}, \vec{\sigma}) = \prod_{i=1}^J \prod_{j=1}^N (2\pi\sigma_i^2)^{-1/2} e^{-|r_i^j - a_i b^j|/2\sigma_i^2}, \quad (5.28)$$

onde \mathbf{r} é uma matriz de dimensão $J \times N$, \vec{a} e $\vec{\sigma}$ são vectores de dimensão J e \vec{b} é um vector de dimensão N . Aplicando logaritmo natural a (5.28), obtém-se:

$$\ln(p(\mathbf{r} | \vec{a}, \vec{b}, \vec{\sigma})) = -\frac{N}{2} \sum_{i=1}^J \ln(2\pi\sigma_i^2) - \sum_{i=1}^J \frac{1}{2\sigma_i^2} \sum_{j=1}^N (r_i^j - a_i b^j)^2. \quad (5.29)$$

As estimativas ML de a_i e σ_i^2 são as que maximizam (5.29). Derivando (5.29) relativamente a estes parâmetros e igualando a zero as derivadas, resulta:

$$\hat{a}_i = \frac{1}{N} \sum_{j=1}^N (r_i^j b^j); \quad (5.30)$$

$$\hat{\sigma}_i^2 = \frac{1}{N} \sum_{j=1}^N (r_i^j b^j - \hat{a}_i)^2, \quad i = 1, 2, \dots, J. \quad (5.31)$$

Os valores b^j necessários (5.30) e em (5.31), são a representação em símbolos binários antípodais (i.e., $b^j \in \{-1; 1\}$), dos bits que constituem a marca-de-água e, como tal, não são conhecidos *a priori*. Neste trabalho foram testados dois tipos de estimativas para estes valores:

- Utilizar o sinal de r_i^j (se $r_i^j > 0 \Rightarrow b^j=1$; se $r_i^j < 0 \Rightarrow b^j=-1$) – **método 1**;
- Utilizar a marca mais provável até ao momento, baseando o seu cálculo nas marcas detectadas nas janelas anteriores; na primeira janela utilizam-se ganhos unitários para a combinação de sinal – **método 2**.

Utilizando o segundo método espera-se que à medida que a sequência de vídeo vai decorrendo, a estimativa da marca-de-água transmitida vai-se aproximando da verdadeira marca-de-água, como será comprovado na secção 5.4.1.

Um problema inerente ao cálculo da estatística do sinal recebido em cada trama, reside no facto de o número de símbolos que compõem a marca poder ser insuficiente para o cálculo de uma boa estimativa. Sendo assim, é de esperar melhores resultados quando o número de níveis utilizados na sinalização é baixo, já que a sequência de símbolos que compõe a marca terá, neste caso, um comprimento maior.

5.4 Resultados

Com a finalidade de avaliar o desempenho, em vídeo, das técnicas de codificação de canal e de combinação de sinal analisadas, realizaram-se vários testes sobre sequências de vídeo, sujeitas a compressão MPEG-2. As sequências utilizadas nestes testes encontram-se representadas na figura 5.8.



(a)



(b)



(c)

Figura 5.8 – Sequências de vídeo CCIR-601 com 300 tramas cada:

a) Stefan; b) Mobile & Calendar; c) Table-Tennis.

A estrutura de tramas utilizada na compressão foi do tipo *GOP 12* (*Group of Pictures 12*)²⁹, com distância entre tramas *P* (tramas predictas) de 3, o que conduz à sequência de tramas codificadas – *IBBPBBPBBPBB...*. Em relação aos ritmos binários, foram realizados testes para compressão a 2, 4 e 6 Mbit/s.

5.4.1 Resultados com inserção no domínio espacial

Análise da influência da compressão nos parâmetros a_i e σ_i

Nesta secção avalia-se a degradação causada no sinal – a marca – pela utilização de compressão MPEG-2. Nas figuras 5.9 a 5.11 encontra-se representada a evolução de a (valor médio do sinal recebido em cada trama) e σ (desvio padrão do ruído) ao longo de três *GOPs* da sequência de vídeo *Stefan*, quando comprimida a 2, 4 e 6 Mbit/s. Para medição destes parâmetros, utilizou-se uma marca-de-água com 64 bits, codificada utilizando códigos convolucionais. A força de inserção foi ajustada para 0.2. As estimativas de a e σ foram realizadas recorrendo à equação (5.31) e com conhecimento *a priori* dos valores b^j .

²⁹ Consultar o anexo D para uma descrição dos vários tipos de tramas usadas na norma MPEG-2.

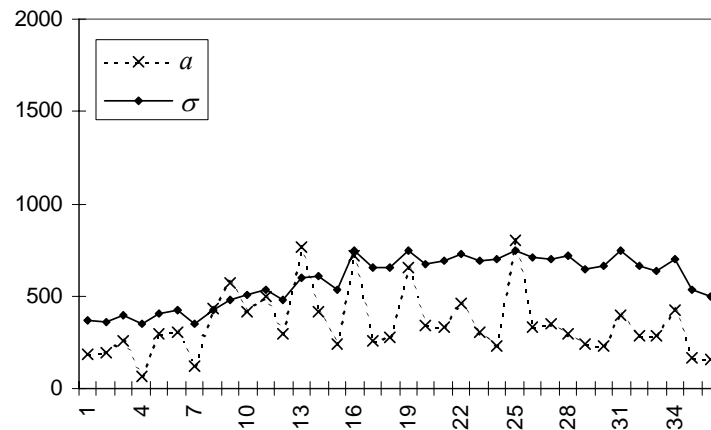


Figura 5.9 – Evolução de a e σ em 3 GOPs da sequência *Stefan* MPEG-2 @ 2 Mbit/s.

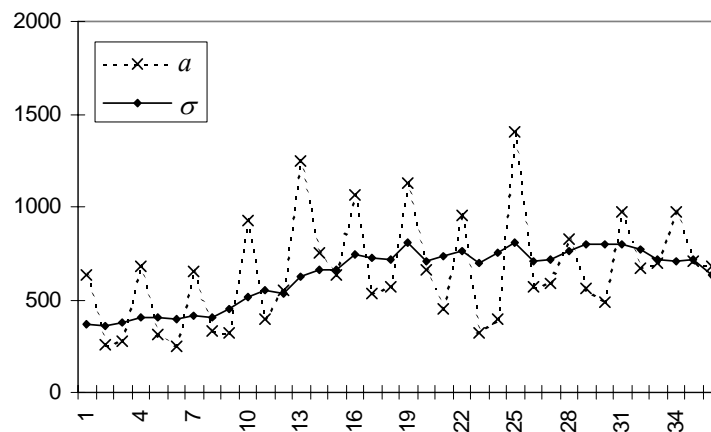


Figura 5.10 – Evolução de a e σ em 3 GOPs da sequência *Stefan* MPEG-2 @ 4 Mbit/s.

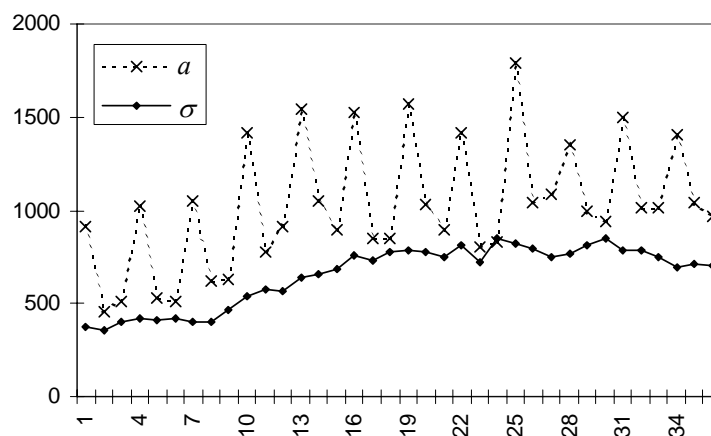


Figura 5.11 – Evolução de a e σ em 3 GOPs da sequência *Stefan* MPEG-2 @ 6 Mbit/s.

Por observação dos gráficos apresentados, pode-se constatar que:

- Para qualquer dos débitos binários considerados, existe uma grande variação de a ao longo da sequência comprimida, o que comprova que a degradação do sinal varia fortemente de trama para trama. Para 4 e 6 Mbit/s a degradação é mais acentuada em tramas do tipo B do que em tramas do tipo I e P ;
- A redução no valor de a (desvanecimento da marca), é tanto maior quanto maior for a taxa de compressão;
- A variação de σ é baixa, quer ao longo das tramas, quer variando a taxa de compressão. A este facto corresponderá um valor de $\gamma \approx 1$ na análise teórica efectuada na secção 5.3.4, sendo de prever resultados semelhantes para as estratégias combinação *ótima* e combinação *quase-ótima*.

Análise da influência da compressão nos valores de c_i

Nas figuras 5.12 a 5.20 apresenta-se a evolução, ao longo de três *GOPs* da sequência de vídeo *Stefan*, dos pesos c_i correspondentes à combinação linear *ótima* ($c_i = a_i / \sigma_i^2$), estimados com base nos seguintes métodos:

- Utilizando apenas os sinais r_i obtidos em cada trama, sendo os valores b^j em (5.31) obtidos de acordo com o **método 1** (descrito em 5.3.5);
- Utilizando os sinais r_i recebidos e os valores b^j obtidos a partir de uma estimativa da marca codificada em cada trama segundo o **método 2** (descrito em 5.3.5);
- Utilizando os sinais r_i recebidos e o conhecimento *a priori* da marca codificada inserida em cada trama (i.e., os valores b^j são conhecidos). Este método seria irrealizável na prática, já que se pretende extracção cega da marca, encontrando-se aqui apresentado para efeitos de comparação com os outros métodos.

Nos mesmos gráficos onde se encontram representados os pesos c_i referentes a cada trama, é também apresentada a probabilidade de erro de bit obtida experimentalmente na extracção da marca-de-água e considerando $J=1$. Em qualquer dos casos, considerou-se uma marca com 64 bits de comprimento, codificada utilizando um código convolucional (com *decisão suave*). O valor da força de inserção foi ajustado para 0.2.

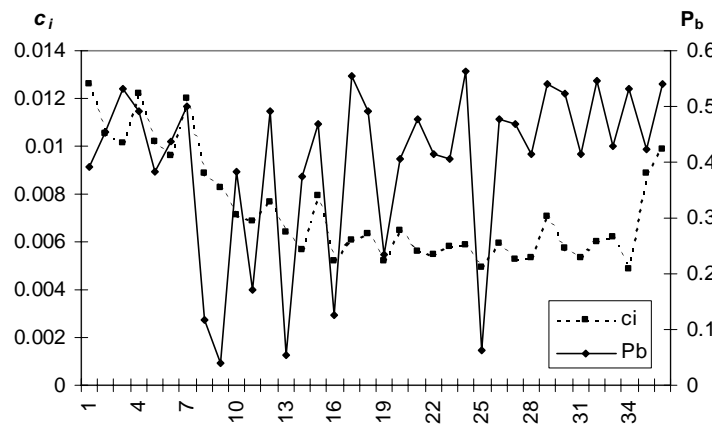


Figura 5.12 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 2 Mbit/s, utilizando o método 1.

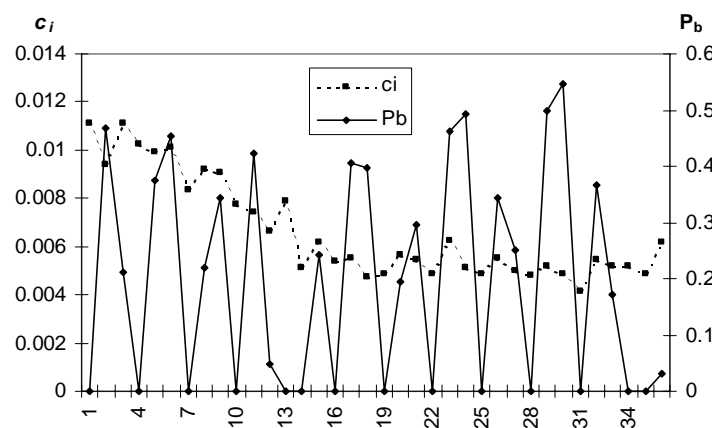


Figura 5.13 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 4 Mbit/s, utilizando o método 1.

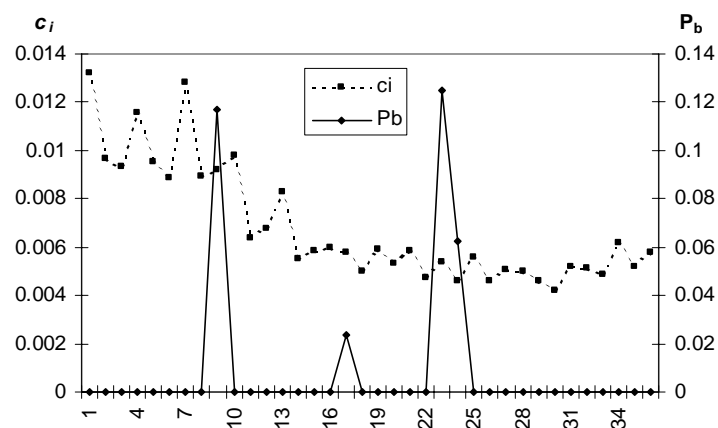


Figura 5.14 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 6 Mbit/s, utilizando o método 1.

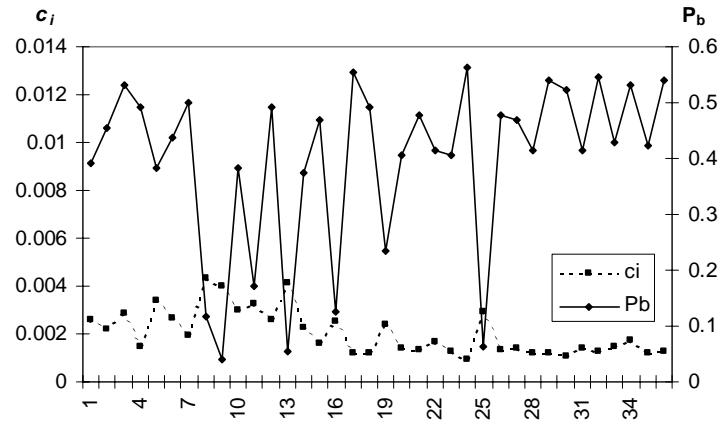


Figura 5.15 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 2 Mbit/s, utilizando o método 2.

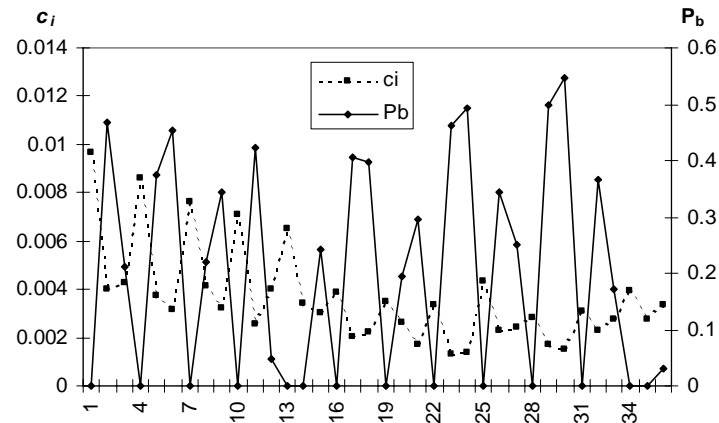


Figura 5.16– Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 4 Mbit/s, utilizando o método 2.

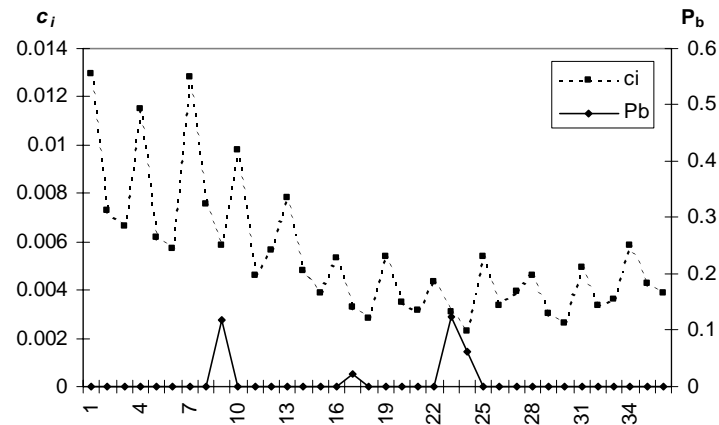


Figura 5.17 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 6 Mbit/s, utilizando o método 2.

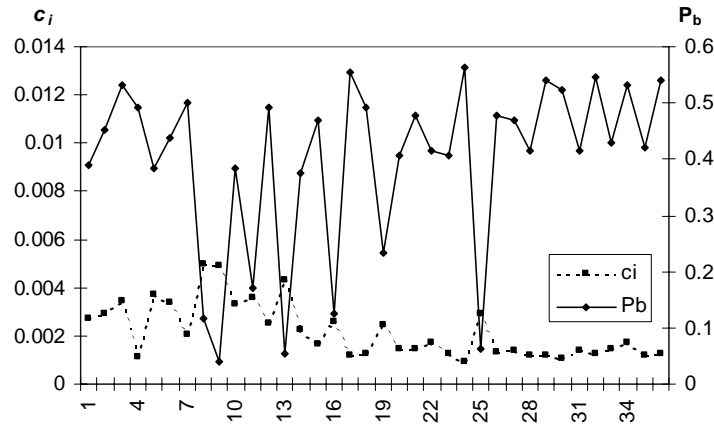


Figura 5.18 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 2 Mbit/s, utilizando conhecimento *a priori* dos símbolos inseridos no cálculo de a_i e σ_i .

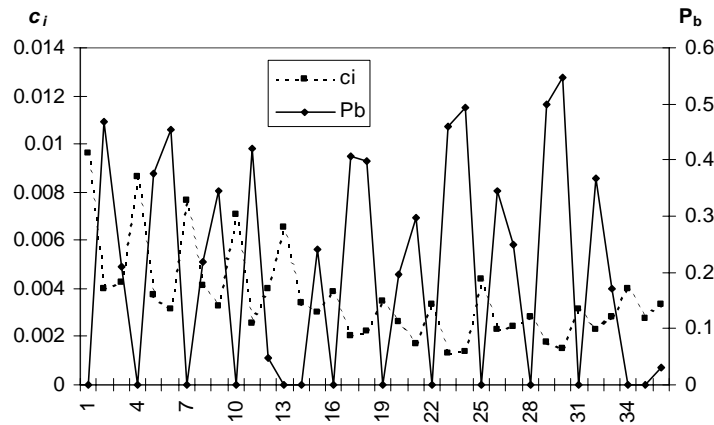


Figura 5.19 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 4 Mbit/s, utilizando conhecimento *a priori* dos símbolos inseridos no cálculo de a_i e σ_i .

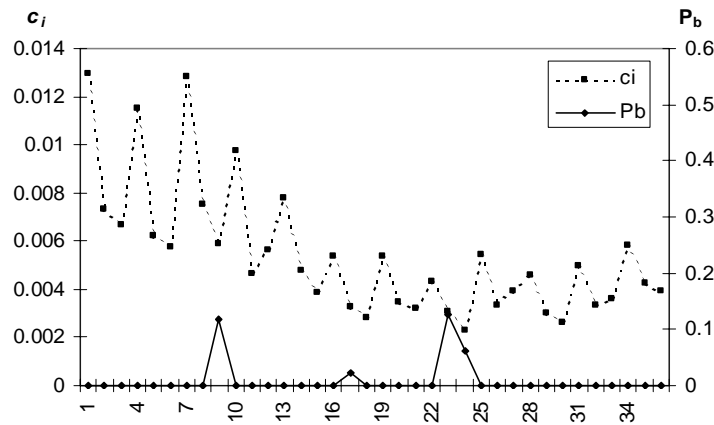


Figura 5.20 – Evolução de c_i em 3 GOPs da sequência *Stefan* MPEG-2 @ 6 Mbit/s, utilizando conhecimento *a priori* dos símbolos inseridos no cálculo de a_i e σ_i .

Como se pode concluir a partir das figuras apresentadas, o método que produz estimativas mais precisas para os pesos é o **método 2**, que apresenta resultados idênticos aos obtidos com conhecimento *a priori* da marca-de-água codificada. Este facto torna-se também claro ao constatar-se que o **método 1** apresenta muitas vezes valores de pesos elevados, para tramas onde a probabilidade de erro de bit é alta, como se pode observar nas figuras 5.12 e 5.13. Este facto é indesejável, já que o objectivo das técnicas de combinação de sinal é precisamente dar ênfase, em termos de detecção, aos sinais recebidos com melhor qualidade.

Avaliação experimental das técnicas de combinação de sinal

Os resultados obtidos com inserção espacial podem ser observados nas tabelas 5.2 a 5.7, que representam a percentagem de janelas de combinação para as quais todos os bits da marca-de-água foram extraídos correctamente. As taxas apresentadas são resultados médios obtidos a partir dos resultados correspondentes a cada uma das sequências de vídeo representadas na figura 5.8. J representa o número tramas (ou canais) consideradas em cada janela. No caso em que $J=1$, a extracção da marca é feita trama a trama.

Nas tabelas apresentadas, as designações LM, LO, LQO e LIC representam, respectivamente:

- LM – Lógica de maioria;
- LO – Combinação linear *ótima*;
- LQO – Combinação linear *quase ótima*;
- LIC – Combinação linear com pesos iguais e constantes.

Foram utilizados dois valores para a força de inserção β – 0.2 e 0.3 – conduzindo ambos a sequências marcadas indistintas das originais. O tipo de codificação de canal e número de níveis utilizados na sinalização foram também variados de forma a avaliar-se o seu impacto no desempenho de cada estratégia de combinação. Considerou-se sinalização multinível sem codificação de canal com $M=2$, 16 e 256, sinalização binária com codificação BCH(127,64), códigos convolucionais binários (2,7) com *decisão dura* e *suave*, e sinalização multinível com $M=16$ e 256, combinada com codificação RS(14,8).

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.01	0.00	0.03	0.08	0.07	0.02	0.17	0.24	0.24	0.10	0.60	0.67	0.67
M=16	0.05	0.00	0.11	0.23	0.22	0.04	0.41	0.51	0.49	0.33	0.82	0.86	0.87
M=256	0.11	0.01	0.36	0.51	0.47	0.21	0.66	0.72	0.67	0.58	0.94	0.94	0.92
BCH(127,64)	0.04	0.02	0.09	0.24	0.23	0.06	0.41	0.52	0.51	0.47	0.80	0.83	0.83
M=16 RS(14,8)	0.07	0.00	0.15	0.27	0.27	0.12	0.51	0.60	0.59	0.52	0.89	0.90	0.88
M=256 RS(14,8)	0.09	0.00	0.25	0.44	0.40	0.15	0.59	0.67	0.62	0.57	0.92	0.92	0.93
Conv. Dec. Dura	0.05	0.02	0.12	0.26	0.27	0.12	0.43	0.56	0.55	0.51	0.80	0.88	0.87
Conv. Dec. Suave	0.13	-	0.40	0.55	0.55	-	0.67	0.75	0.76	-	0.92	0.92	0.93

Nº Níveis / Código	J=15				J=20				J=25			
	LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.50	0.77	0.80	0.80	0.76	0.89	0.84	0.87	0.83	0.92	0.94	0.92
M=16	0.58	0.95	0.95	0.97	0.78	0.98	0.98	0.96	0.86	1.00	1.00	1.00
M=256	0.72	0.95	0.95	0.95	0.73	1.00	1.00	0.98	0.86	1.00	1.00	1.00
BCH(127,64)	0.78	0.92	0.95	0.95	0.89	0.96	0.98	0.98	0.97	1.00	1.00	1.00
M=16 RS(14,8)	0.70	0.95	0.95	0.95	0.78	1.00	0.98	1.00	0.94	1.00	1.00	1.00
M=256 RS(14,8)	0.67	0.95	0.95	0.95	0.73	1.00	0.98	0.98	0.81	1.00	1.00	1.00
Conv. Dec. Dura	0.83	0.93	0.93	0.93	0.89	0.96	0.98	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	-	0.97	0.98	0.98	-	0.98	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.3 – Taxa de sucesso na extracção com compressão MPEG-2 @ 2Mbit/s – $\beta = 0.2$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.12	0.07	0.43	0.49	0.52	0.29	0.67	0.74	0.73	0.66	0.90	0.92	0.93
M=16	0.19	0.07	0.66	0.73	0.71	0.44	0.88	0.89	0.88	0.83	0.94	0.96	0.96
M=256	0.24	0.12	0.86	0.86	0.83	0.58	0.92	0.92	0.92	0.87	0.98	0.98	0.98
BCH(127,64)	0.19	0.32	0.65	0.70	0.71	0.67	0.86	0.88	0.87	0.90	0.96	0.96	0.96
M=16 RS(14,8)	0.21	0.12	0.73	0.76	0.75	0.54	0.91	0.91	0.90	0.88	0.96	0.96	0.97
M=256 RS(14,8)	0.23	0.12	0.81	0.83	0.80	0.51	0.92	0.93	0.92	0.82	0.96	0.96	0.96
Conv. Dec. Dura	0.21	0.40	0.72	0.72	0.74	0.69	0.89	0.90	0.91	0.90	0.97	0.96	0.96
Conv. Dec. Suave	0.25	-	0.87	0.88	0.88	-	0.93	0.93	0.94	-	0.99	0.99	0.99

Nº Níveis / Código	J=15				J=20				J=25			
	LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.87	0.95	0.95	0.95	0.91	1.00	1.00	1.00	0.97	1.00	1.00	1.00
M=16	0.92	1.00	0.98	0.97	0.91	1.00	1.00	1.00	0.97	1.00	1.00	1.00
M=256	0.92	1.00	1.00	1.00	0.98	1.00	1.00	1.00	0.97	1.00	1.00	1.00
BCH(127,64)	0.93	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16 RS(14,8)	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	0.92	1.00	1.00	1.00	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	0.95	1.00	0.98	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.4 – Taxa de sucesso na extracção com compressão MPEG-2 @ 2Mbit/s – $\beta = 0.3$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.21	0.29	0.77	0.81	0.82	0.67	0.92	0.93	0.93	0.96	1.00	1.00	1.00
M=16	0.28	0.33	0.94	0.96	0.95	0.85	0.99	0.99	0.99	1.00	1.00	1.00	1.00
M=256	0.36	0.60	0.99	1.00	0.99	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BCH(127,64)	0.27	0.79	0.92	0.94	0.94	0.94	0.99	0.99	0.99	1.00	1.00	1.00	1.00
M=16 RS(14,8)	0.29	0.61	0.95	0.97	0.97	0.89	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	0.31	0.62	0.99	0.99	0.99	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	0.28	0.85	0.95	0.96	0.96	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	0.39	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.5 – Taxa de sucesso na extracção com compressão MPEG-2 @ 4Mbit/s – $\beta = 0.2$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.46	0.88	0.98	0.98	0.98	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16	0.74	0.91	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256	0.90	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BCH(127,64)	0.74	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16 RS(14,8)	0.83	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	0.89	0.97	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	0.77	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	0.92	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.6 – Taxa de sucesso na extracção com compressão MPEG-2 @ 4Mbit/s – $\beta = 0.3$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.37	0.92	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16	0.66	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256	0.90	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BCH(127,64)	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16 RS(14,8)	0.80	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	0.89	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	0.73	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	0.95	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.7 – Taxa de sucesso na extracção com compressão MPEG-2 @ 6Mbit/s – $\beta = 0.2$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BCH(127,64)	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16 RS(14,8)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.8 – Taxa de sucesso na extracção com compressão MPEG-2 @ 6Mbit/s – $\beta = 0.3$.

Os resultados experimentais confirmam o esperado teoricamente:

- Independentemente do valor de J , verifica-se um aumento na taxa de sucesso na extracção da marca, se:
 - aumentar a força de inserção da marca, ou diminuir a taxa de compressão (aumento do ritmo do vídeo comprimido);
 - aumentar o número de níveis utilizados na sinalização;
 - forem utilizados códigos correctores de erro, especialmente codificação binária convolucional com *decisão suave* na decodificação, ou codificação RS combinada com sinalização multinível.
- A extracção da marca com base numa janela de J tramas aumenta a taxa de detecção da marca, sendo este aumento tanto mais significativo quanto maior for J ;
- Os métodos de combinação de sinal que apresentam melhores resultados são os correspondentes às estratégias *pesos óptimos* e *pesos quase óptimos*. A estratégia *pesos iguais e constantes*, apesar de mais simples que as anteriores, apresenta resultados bastante próximos para débitos binários iguais ou superiores a 4 Mbits/s. Na estratégia *lógica de maioria*, os resultados vão melhorando à medida que se aumenta J , mas exibem um desempenho bastante inferior ao da combinação linear.
- Para os testes efectuados a 2 Mbit/s, o valor de 100% de sucesso na extracção só se atinge ou com sinalização multinível ou com códigos de correcção de erros, em conjunto com combinação de sinal.

5.4.2 Resultados com inserção no domínio da frequência

De forma equivalente ao que foi realizado com inserção no domínio espacial, apresentam-se nesta secção resultados ilustrativos do desempenho das várias técnicas de combinação de sinal com inserção no domínio da frequência. Os parâmetros de teste são idênticos aos descritos para o domínio espacial, à excepção da força de inserção, que foi ajustada para $\beta=1.5$. Com este ajuste, a PSNR após inserção é semelhante à resultante no domínio espacial com $\beta=0.2$. Os resultados obtidos e que se encontram sintetizados nas tabelas 5.8 a 5.10, permitem afirmar que as conclusões obtidas para a inserção no domínio espacial podem ser generalizadas para a inserção no domínio da frequência.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.20	0.24	0.28	0.30	0.30	0.27	0.32	0.34	0.36	0.33	0.47	0.57	0.58
M=16	0.25	0.26	0.33	0.34	0.35	0.32	0.41	0.45	0.50	0.41	0.59	0.68	0.69
M=256	0.28	0.27	0.36	0.44	0.42	0.33	0.53	0.60	0.58	0.50	0.70	0.74	0.72
BCH(127,64)	0.25	0.29	0.33	0.35	0.34	0.32	0.39	0.47	0.51	0.44	0.66	0.71	0.72
M=16 RS(14,8)	0.26	0.26	0.33	0.37	0.36	0.32	0.44	0.53	0.53	0.43	0.68	0.71	0.71
M=256 RS(14,8)	0.26	0.27	0.36	0.39	0.39	0.32	0.48	0.59	0.58	0.44	0.71	0.72	0.74
Conv. Dec. Dura	0.25	0.29	0.32	0.35	0.35	0.34	0.41	0.47	0.50	0.47	0.60	0.68	0.72
Conv. Dec. Suave	0.28	-	0.40	0.48	0.48	-	0.54	0.65	0.64	-	0.74	0.77	0.80

Nº Níveis / Código	J=15				J=20				J=25			
	LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.40	0.65	0.73	0.73	0.56	0.71	0.73	0.73	0.75	0.78	0.81	0.83
M=16	0.52	0.75	0.82	0.78	0.69	0.78	0.89	0.84	0.72	0.92	0.92	0.94
M=256	0.70	0.85	0.85	0.82	0.76	0.91	0.96	0.93	0.81	0.94	1.00	0.94
BCH(127,64)	0.70	0.77	0.77	0.78	0.76	0.78	0.87	0.89	0.78	0.89	1.00	1.00
M=16 RS(14,8)	0.67	0.75	0.77	0.77	0.73	0.82	0.84	0.84	0.78	0.97	0.97	0.94
M=256 RS(14,8)	0.67	0.80	0.83	0.82	0.76	0.93	0.98	0.96	0.78	0.97	0.97	0.97
Conv. Dec. Dura	0.72	0.75	0.80	0.78	0.71	0.80	0.87	0.89	0.78	0.86	0.92	0.92
Conv. Dec. Suave	-	0.87	0.90	0.90	-	0.98	0.98	0.98	-	1.00	1.00	1.00

Tabela 5.9 – Taxa de sucesso na extracção com compressão MPEG-2 @ 2 Mbit/s – $\beta = 1.5$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.62	0.78	0.85	0.87	0.87	0.87	0.90	0.91	0.91	0.89	0.96	0.97	0.97
M=16	0.73	0.82	0.90	0.91	0.89	0.87	0.94	0.95	0.94	0.94	0.99	0.99	0.99
M=256	0.80	0.84	0.93	0.94	0.92	0.87	0.97	0.97	0.97	0.94	1.00	1.00	1.00
BCH(127,64)	0.75	0.89	0.90	0.91	0.91	0.91	0.95	0.95	0.96	0.96	0.98	0.98	0.98
M=16 RS(14,8)	0.77	0.84	0.91	0.92	0.92	0.90	0.95	0.95	0.95	0.93	0.98	0.99	1.00
M=256 RS(14,8)	0.82	0.87	0.92	0.93	0.92	0.90	0.98	0.98	0.98	0.93	1.00	1.00	1.00
Conv. Dec. Dura	0.74	0.89	0.90	0.91	0.91	0.91	0.95	0.95	0.96	0.97	0.99	0.99	1.00
Conv. Dec. Suave	0.82	-	0.93	0.93	0.94	-	0.98	0.98	0.98	-	1.00	1.00	1.00

Tabela 5.10 – Taxa de sucesso na extracção com compressão MPEG-2 @ 4 Mbit/s – $\beta = 1.5$.

Nº Níveis / Código	J=1	J=3				J=5				J=10			
		LM	LIC	LQO	LO	LM	LIC	LQO	LO	LM	LIC	LQO	LO
M=2	0.79	0.97	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16	0.91	0.99	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BCH(127,64)	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=16 RS(14,8)	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M=256 RS(14,8)	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Dura	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Conv. Dec. Suave	0.98	-	1.00	1.00	1.00	-	1.00	1.00	1.00	-	1.00	1.00	1.00

Tabela 5.11 – Taxa de sucesso na extracção com compressão MPEG-2 @ 6 Mbit/s – $\beta = 1.5$.

Para vídeo comprimido a 2 Mbit/s (taxa de compressão mais elevada) e para valores baixos de J , os resultados obtidos com inserção no domínio da frequência revelam um melhor desempenho face aos obtidos para o domínio espacial. Nas restantes situações, a inserção no domínio espacial conduz a maiores taxas de sucesso na extracção.

5.5 Considerações finais

Neste capítulo verificou-se experimentalmente que, no caso de vídeo e à semelhança do que acontece para imagens fixas, é possível melhorar o desempenho na extracção da marca-de-água, quer aumentando o número de níveis utilizados na modulação da marca, quer utilizando códigos de correcção de erros. Neste último caso, os melhores resultados foram obtidos pelos códigos convolucionais e pelos códigos de bloco não binários (RS) com 256 níveis de sinalização.

A extracção da marca com base num conjunto de J tramas aumenta a probabilidade de a marca ser correctamente extraída, sendo este aumento tanto mais significativo quanto maior for J . Verifica-se também que a utilização de combinação linear de sinal conduz a taxas de sucesso na extracção da marca superiores às obtidas com um esquema de decisão por lógica de maioria. No conjunto das técnicas de combinação de sinal estudadas, as estratégias combinação linear *ótima* e combinação linear *quase-ótima* conduzem ao melhor desempenho, sobretudo para a taxa de compressão mais elevada (correspondente a um débito binário de 2 Mbit/s). Este resultado está de acordo com o previsto teoricamente. A combinação linear com *pesos iguais e constantes* demonstrou uma boa relação desempenho / complexidade. Embora esta estratégia conduza globalmente a resultados inferiores aos da combinação linear *ótima*, a sua simplicidade constitui uma grande vantagem para aplicações que requeiram baixo custo e maior rapidez de processamento.

Estas conclusões aplicam-se para os dois domínios de inserção estudados. De salientar, no entanto, que os resultados obtidos para inserção no domínio da frequência apenas exibem melhor desempenho face ao domínio espacial quando a taxa de compressão é elevada e o número de tramas utilizadas para combinação de sinal é baixo.

Capítulo 6

Conclusões

A tecnologia das marcas-de-água consiste na inserção, de forma imperceptível, de informação adicional sobre produtos multimédia em formato digital. A informação inserida varia com o cenário de aplicação podendo, como exemplo, conter dados descritivos ou de referência, a identificação do detentor dos direitos de autor, informação que permita avaliar a integridade do conteúdo ou dados para controlo de cópias. Embora os requisitos a cumprir variem consoante a finalidade a que se destinam as marcas-de-água, é usual exigir que o algoritmo de marcação seja *seguro*, i.e., a marca deve ser lida apenas por receptores autorizados, e *robusto*, i.e., a marca deve permanecer no produto marcado após este ser sujeito a manipulações que não reduzam o seu valor comercial.

Muitas das técnicas de marca-de-água propostas na literatura são modelizadas como um processo de comunicação, no qual se pretende enviar uma mensagem – a marca – através de um canal ruidoso – o produto multimédia. No entanto, e ao contrário dos sistemas de comunicação tradicionais, deverão ser cumpridos critérios de qualidade relativamente ao suporte de transmissão, i.e., a qualidade do produto marcado deve ser idêntica à do produto original. Este critério, juntamente com os requisitos *robustez* e *segurança*, levou a que vários investigadores

tenham sugerido a utilização dos princípios do espalhamento de espectro em técnicas de marca-de-água.

O principal objectivo desta tese foi avaliar o impacto, em sistemas de marcas-de-água baseados em espalhamento de espectro, de técnicas habitualmente utilizadas em sistemas de comunicação digital. Este estudo foi restringido a imagens fixas e vídeo. Entre as técnicas analisadas, encontram-se:

- Modulação multinível;
- Codificação para correcção de erros;
- Técnicas de combinação de sinal.

Como ficou demonstrado no capítulo 3, a partir de uma modelização estatística adequada para o canal de transmissão – a imagem a ser marcada – é possível prever, teoricamente, desempenhos na extracção da marca bastante próximos dos obtidos experimentalmente. Confirmou-se ainda que a utilização de modulação multinível baseada em espalhamento de espectro conduz a um aumento do sucesso na extracção da marca-de-água, diminuindo a frequência de erros com o aumento do número de níveis de sinalização. Esta conclusão é válida na presença de compressão JPEG, cortes da imagem e ruído aditivo, branco e gaussiano.

No capítulo 4 demonstrou-se que a utilização de códigos para correcção de erros pode ter um impacto significativo no desempenho das técnicas de marca-de-água baseadas em espalhamento de espectro, sobretudo para os valores de probabilidade de erro com interesse prático. Entre os vários códigos analisados, destacaram-se os códigos de bloco não binários (*Reed-Solomon*) com 256 níveis de modulação, e os códigos convolucionais binários, com descodificação segundo o algoritmo de *Viterbi* e decisão suave.

No capítulo 5 analisou-se a aplicação de marcas-de-água a sequências vídeo, sujeitas a compressão MPEG-2, como um processo de comunicação multi-canal. Verificou-se que, para esta caso, a extracção da marca pode ser substancialmente melhorada considerando um grupo de tramas consecutivas, aplicando técnicas de combinação de sinal, amplamente utilizadas no domínio das comunicações rádio com diversidade. De entre as técnicas analisadas, verificou-se que as estratégias *combinação linear óptima* e *combinação linear quase-óptima* conduzem ao melhor desempenho. Para vídeo comprimido a 2 Mbit/s (o menor dos débitos considerados)

verificou-se que, para se atingir 100% de sucesso na extracção da marca, é imperativo utilizar técnicas de combinação de sinal associadas a codificação de canal e/ou modulação multinível.

As técnicas estudadas ao longo da tese foram avaliadas para dois espaços de inserção distintos – *domínio espacial* e *domínio da frequência*. No primeiro caso, a marca é inserida directamente no espaço da imagem, por alteração de uma ou mais componentes de cor, tendo-se neste trabalho optado pela utilização da componente da luminância. No segundo caso, a marca é inserida através da alteração de coeficientes espectrais resultantes de uma transformação em frequência da imagem. Para inserção neste domínio, optou-se pela utilização da transformada DCT, orientada ao bloco de dimensões 8×8 pixels. Esta escolha foi feita tendo em conta as normas de compressão mais utilizadas (JPEG e MPEG-2), que recorrem também a este tipo de transformada. Independentemente do domínio escolhido para inserção, a aplicação das diversas técnicas estudadas ao longo da tese conduziram a melhorias de desempenho face à transmissão binária simples. Na presença de compressão os resultados foram favoráveis à inserção no domínio espacial.

Atendendo aos resultados obtidos nesta tese, podem-se perspectivar os seguintes tópicos de investigação para trabalho futuro:

- Aplicação de outras classes de códigos de correcção de erros, nomeadamente códigos *turbo* e codificação-modulação *Trellis*. Com base na teoria da comunicação, é de esperar um melhor desempenho com a utilização destas técnicas de codificação de canal;
- Aplicação de combinação de sinal em imagem fixa colorida, tirando partido do facto de cada componente da cor poder ser encarada como um canal;
- Implementação de métodos de simulação numérica que possam ser aplicados em testes às marcas-de-água, com a finalidade de se obterem resultados experimentais com maior rapidez. Entre estes métodos encontram-se as técnicas de *importance sampling*;
- Determinar os coeficientes DCT adequados para inserção da marca-de-água e a dependência dessa escolha do tipo de imagem; avaliar o desempenho de modelos perceptuais desenvolvidos para este domínio diferentes da solução considerada nesta tese;
- Aplicar as técnicas estudadas à inserção no domínio da transformada *wavelet* (DWT), uma vez que a nova norma para compressão de imagens fixas – JPEG2000 – recorre a esta transformada.

Bibliografia

1. A. J. Ahumada e H. A. Peterson, "Luminance-model-based DCT quantization for color image compression", *Proc. SPIE on Human Vision, Visual Processing, and Digital Display*, Vol. 1666, S. Jose CA, EUA, 1992.
2. M. Barni, F. Bartolini, V. Cappellini, A. Lippi e A. Piva, "A DWT-based technique for spatio-frequency masking of digital signatures", *Proc. SPIE on Security and Watermarking of Multimedia Contents*, Vol. 3657, S. Jose CA, EUA, Janeiro de 1999.
3. M. Barni, F. Bartolini, V. Cappellini e A. Piva, "A DCT-domain system for robust image watermarking", *Signal Processing: Special Issue on Watermarking*, Vol. 66, Nº 3, Maio de 1998, pp. 357-372.
4. T. Brandão, M.P. Queluz e A. Rodrigues, "Diversity enhancement of coded spread spectrum video watermarking", *Proc. WPCM01*, Aalborg, Dinamarca, Setembro de 2001. Comunicação pré-seleccionada para publicação na revista *Wireless Personal Communications*.
5. T. Brandão, M.P. Queluz e A. Rodrigues, "Improving spread spectrum based image watermarking through non-binary channel coding", *Proc. ConfTele 2001 – 3rd Conference on Telecommunications*, Figueira da Foz, Portugal, Abril de 2001.
6. T. Brandão, M.P. Queluz e A. Rodrigues, "On the use of error correction codes in spread spectrum based image watermarking", *Proc. IEEE Pacific Rim Conference on Multimedia*, Pequim, China, Outubro de 2001.
7. T. Brandão, M.P. Queluz e A. Rodrigues, "Performance improvement of spatial watermarking through efficient non-binary channel coding", *Proc. SPIE on Security and Watermarking of Multimedia Contents III*, Vol. 4314, S. Jose CA, EUA, Janeiro de 2001.
8. C. Busch, W. Funk e S. Wolthusen, "Digital watermarking: from concepts to real-time video applications", *IEEE Computer Graphics and Applications*, Vol. 19, Nº 1, Janeiro / Fevereiro de 1999, pp. 25-35.

9. G. Caronni, "Assuring ownership rights for digital images", *Proc. VIS 95, Session "Reliable IT Systems"*, Vieweg Publishing Company, Germany, 1995, pp.251-263.
10. *Certimark project – Information Society Technologies*, "User's requirements, metrics & parameters", Relatório técnico, 2001.
11. I. J. Cox, J. Kilian, F. T. Leighton e T. Shamon, "Secure spread spectrum watermarking for multimedia" *IEEE Transactions on Image Processing*, Vol. 6, Nº 12, Dezembro de 1997.
12. S. Craver, N. Memon, B.-L. Yeo e M. M. Yeung, "Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks and implications", *IEEE Journal on Selected Areas in Communications*, Vol.16, Nº 4, Maio de 1998.
13. J. F. Delaigle, C. De Vleeschouwer e B. Macq, "Low cost perceptive digital picture watermarking method" – *Proc. ECTAST'97*, Maio de 1997, pp.153-167.
14. G. Depovere, T. Kalker e J.-P. Linnartz, "Improved watermark detection reliability using filtering before correlation", *Proc. IEEE International Conference on Image Processing (ICIP)*, 1998.
15. F. Fonseca, *Assinatura digital de imagens II*, Trabalho Final de Curso, Instituto Superior Técnico, Maio de 2001.
16. F. Hartung e B. Girod, "Watermarking of uncompressed and compressed video", *Signal Processing: Special Issue on Watermarking*, Vol. 66, Nº 3, Maio de 1998.
17. F. Hartung e M. Kutter, "Multimedia watermarking techniques", *Proc. IEEE, Special Issue on Identification & Protection of Multimedia Information*, Vol. 87, Nº 7, Julho de 1999, pp 1079-1107.
18. J. R. Hernández, M. Amado e F. Pérez-González, "DCT-domain watermarking techniques for still images: detector performance analysis and a new structure", *IEEE Transactions on Image Processing*, Vol. 9, Nº 1, Janeiro de 2000.

19. J. Hernández, J.-F. Delaigle e B. Macq, “Improving data hiding by using convolutional codes and soft-decision decoding”, *Proc. SPIE on Security and Watermarking of Multimedia Contents II*, Vol. 3971, Janeiro de 2000.
20. J. R. Hernández, F. Pérez-González, J. M. Rodríguez e G. Nieto, “Performance analysis of a 2-D multipulse amplitude modulation scheme for data hiding and watermarking of still images”, *IEEE Journal on Selected Areas in Communications*, Vol. 16, Nº 4, Maio de 1998.
21. J. Hernández, J. M. Rodríguez e F. Pérez-González, “Improving the performance of spatial watermarking of images using channel coding”, *IEEE Signal Processing Magazine*, Nº 80, 2000.
22. A. K. Jain, *Fundamentals of digital image processing*, Prentice-Hall International, Inc., 1989.
23. M. C. Jeruchim, P. Balabau e K. S. Shanmugan, *Simulation of communication systems*, Plenum Press, 1992.
24. T. Kalker, G. Depovere, J. Haitsma e M. J. Maes, “Video wartermarking for broadcast monitoring”, *Proc. SPIE on Security and Watermarking of Multimedia Contents*, Vol. 3657, S. Jose CA, EUA, Janeiro de 1999.
25. M. Kutter, *Digital image watermarking: hiding information in images*, Tese de Doutoramento, École Polytechnique Fédérale de Lausanne, Lausanne, Suíça, 1999.
26. M. Kutter e F. Petitcolas, “A fair benchmark for watermarking systems”, *Proc. SPIE on Security and Watermarking of Multimedia Contents*, Vol. 3657, Janeiro de 1999.
27. P. Lamy Bandeira, J. M. Martinho e T. Rosa-Limpo, *Assinatura digital de imagens*, Trabalho Final de Curso, Instituto Superior Técnico, Julho de 1999.
28. L. Litwin e K. Ramaswamy, “Error control coding – an overview of modern coding techniques in a digital communications system”, *IEEE Potentials*, Vol. 20, Nº 1, Fevereiro / Março de 2001.

29. C.-Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller e Y. M. Lui, "Rotation, scale, and translation resilient public watermarking for images", *Proc. SPIE on Security and Watermarking of Multimedia Contents*, Vol. 3971, S. Jose CA, EUA, Janeiro de 2000.
30. P. Meerwald e A. Uhl, "A survey of Wavelet-domain watermarking algorithms", *Proc. SPIE on Security and Watermarking of Multimedia Contents III*, Vol. 4314, S. Jose CA, EUA, Janeiro de 2001.
31. A. M. Michelson e A. M. Levesque, *Error-control techniques for digital communication*, Wiley-Interscience, 1985.
32. A. Papoulis, *Probability, random variables and stochastic processes*, McGraw-Hill, 3rd edition, 1991.
33. C. I. Podilchuck e W. Zeng, "Image-adaptive watermarking using visual models", *IEEE Journal on Selected Areas in Communications*, Vol. 16, N° 4, Maio de 1998.
34. J. G. Proakis, *Digital communications*, McGraw-Hill International Editions, 2nd edition, 1989.
35. S. Sangwine e R. Horne, *The colour image processing handbook*, Chapman & Hall, 1998.
36. B. Schneider, *Applied cryptography*, Wiley, 1996.
37. R. G. van Schyndel, A. Z. Tirkel e C. F. Osborne, "A digital watermarking", *Proc. IEEE International Conference on Image Processing (ICIP)*, 1994.
38. C. E. Shannon, "A mathematical theory of communication", *Bell Systems Technical Journal*, Vol. 27, N° 4, 1948.
39. P. J. Smith, M. Shafi e H. Gao, "Quick simulation: a review of importance sampling techniques in communications systems", *IEEE Journal on Selected Areas in Communication*, Vol. 15, N° 4, Maio de 1997, pp 597-613.
40. M. D. Swanson, B. Zhu e A. H. Tewfik, "Robust data hiding for images", *Proc. of the IEEE Digital Signal Processing Workshop*, Loen, Noruega, Maio de 1996, pp.37-40.

-
41. K. Tanaka, Y. Nakamura e K. Matsui, "Embedding secret information into a dithered multilevel image", *Proc. of the 1990 IEEE Military Communications Conference*, Setembro de 1990.
 42. K. Tanaka, Y. Nakamura e K. Matsui, "Embedding the attribute information into a dithered image", *Systems and Computers in Japan*, Vol. 21, Nº 7, 1990.
 43. A. Tirkel, G. Rankin, R. van Schyndel, W. Ho, N. Mee e C. Osborne, "Electric watermark", *Proc. DICTA 1993*, Dezembro de 1993, pp.666-672.
 44. M. van Trees, *Detection, estimation, and modulation theory*, J. Wiley and Sons, Inc., 1967.
 45. V. A. Vlnrotter, E. R. Rodemich e S. J. Dolinar, Jr., "Real-time combining of carrier array signals using ML weight estimates", *IEEE Transactions on Communications*, Vol. 40, Nº 3, Março de 1992, pp 604-615.
 46. A. J. Viterbi e J. K. Omura, *Principles of digital communication and coding*, McGraw-Hill, 1979.
 47. R. B. Wolfgang e E. J. Delp, "Overview of image security techniques with applications in multimedia systems", *Proc. SPIE International Conference on Voice, Video and Data Communications*, Dallas, Texas, EUA, Novembro de 1997.

ANEXOS

Anexo A

Limite de *Shannon*

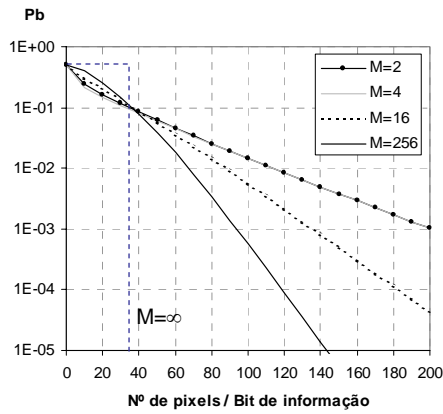
Neste anexo mostra-se que o ponto de cruzamento das curvas teóricas da probabilidade de erro de bit (P_b), obtidas no capítulo 3 para inserção no domínio espacial e modulação *M*-ária sem codificação, ocorre para um valor da relação sinal-ruído por bit (SNR) correspondente ao limite de Shannon.

Pelo teorema de *Shannon-Hartley* [38], a capacidade (em bit/s) de um canal contínuo corrompido por ruído aditivo, branco e gaussiano, é dada por:

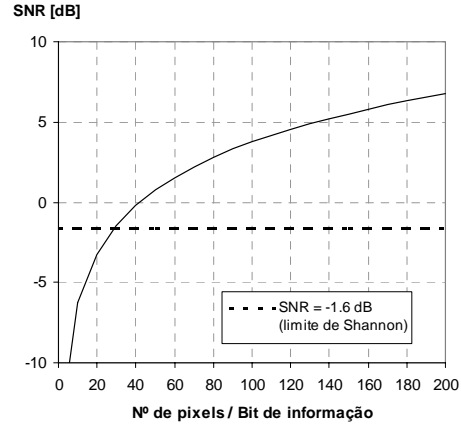
$$C = B \log_2 \left(1 + \frac{S}{N} \right), \quad (\text{A.1})$$

onde B é a largura de banda do canal (Hz) e S/N é a relação sinal-ruído na recepção. A relação (A.1) indica ser possível, para uma dada capacidade, reduzir a potência emitida à custa de um aumento na largura de banda. No entanto, pode-se demonstrar que no caso ideal de largura de banda infinita, a transmissão fiável (i.e., com probabilidade de erro na recepção arbitrariamente pequena) de informação a um ritmo igual à capacidade do canal, implica um valor mínimo na relação sinal ruído por bit de $\ln 2$ (-1.6 dB). Este limite teórico é designado por limite de *Shannon*. Se for utilizada modulação *M*-ária bi-ortogonal (ou ortogonal), pode-se demonstrar que fazendo $M \rightarrow \infty$, a probabilidade de erro de bit tende para zero desde que S/N seja superior a -1.6 dB [34].

A figura A.1 ilustra esta situação limite, para o caso da imagem *Lena*. Em A.1-a) encontra-se representada a probabilidade de erro de bit na recepção em função do número de pixels utilizados para transporte de um bit de informação e para diversos valores de M . Verifica-se que ao ponto de cruzamento das curvas corresponde um valor do número de pixels por bit de informação entre 30 e 35, sendo o declive das curvas tanto mais acentuado quanto maior for M . No caso limite, com $M \rightarrow \infty$, o declive tenderá para infinito e, para valores do *número de pixels por bit de informação* superiores ao ponto de cruzamento, a probabilidade de erro de bit tenderá para zero.



(a)



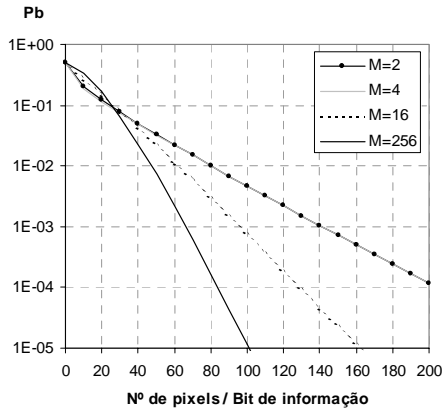
(b)

Figura A.1 – Limite de Shannon – imagem *Lena*:

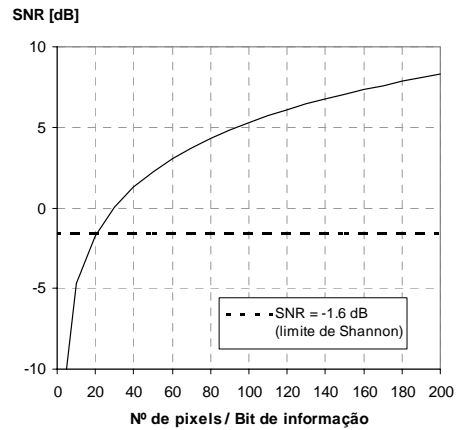
a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.

A figura A.1-b) representa a evolução da SNR por bit em função do número de pixels utilizados para transmissão de um bit de informação útil (o valor da SNR referente ao limite de *Shannon* é também representado a tracejado). Desta figura, confirma-se que o ponto de cruzamento das várias curvas ocorre de facto para um valor de SNR próximo de -1.6 dB.

As figuras A.2 e A.3 ilustram a mesma situação para as imagens *Mandrill* e *02*.



(a)



(b)

Figura A.2 – Limite de Shannon – imagem *Mandrill*:

a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.

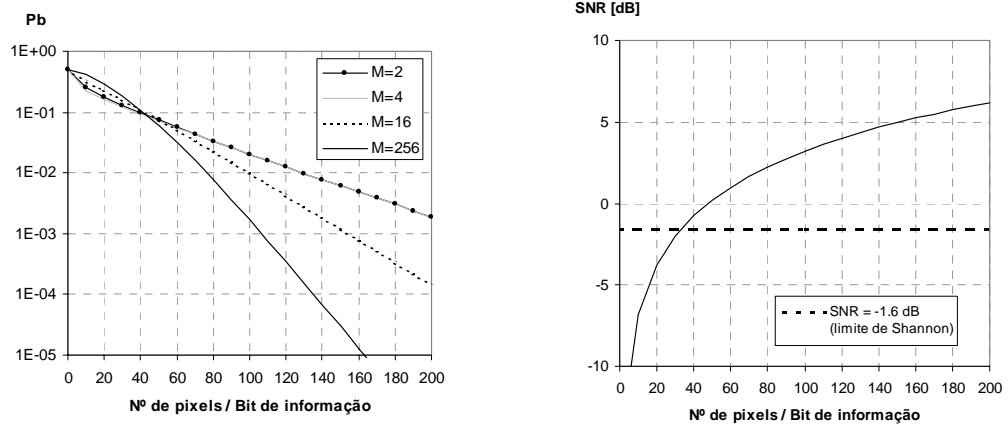


Figura A.3 – Limite de Shannon – imagem 02:

a) P_b vs. N° de pixels/bit de informação; b) SNR vs. N° de pixels/bit de informação.

Anexo B

Frequências espaciais em ciclos por grau

Neste anexo apresenta-se a conversão, de *ciclos por bloco* para *ciclos por grau*, das frequências espaciais correspondentes às funções base da transformada DCT, orientada ao bloco. As frequências espaciais, em ciclos por grau, são utilizadas no modelo perceptual descrito neste trabalho para inserção no domínio da frequência (DCT).

Seja H a altura do écran e d a distância do observador ao écran. Considerando que apenas é visualizada uma fracção α da altura do écran (figura B.1), o ângulo de visualização – θ – é dado por:

$$\theta = 2 \arctg \left(\frac{\alpha H}{2d} \right), \quad (\text{B.1})$$

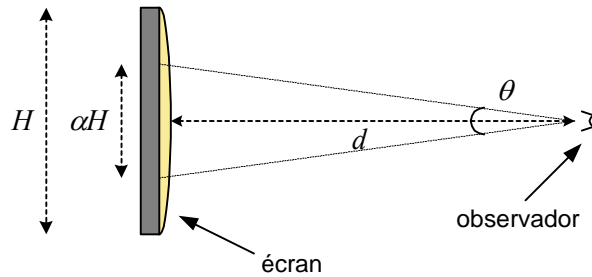


Figura B.1 – Ângulo de observação.

ou, resolvendo em ordem a α :

$$\alpha = 2 \frac{d}{H} \operatorname{tg} \left(\frac{\theta}{2} \right). \quad (\text{B.2})$$

Para ângulos de observação pequenos é válida a aproximação:

$$\operatorname{tg} \left(\frac{\theta}{2} \right) \approx \frac{\theta}{2}, \quad (\text{B.3})$$

obtendo-se, por substituição de (B.3) em (B.2):

$$\alpha \approx \frac{d}{H} \theta. \quad (\text{B.4})$$

i ou j	f_{i0} ou f_{0j} (ciclos / bloco)
0	–
1	0.5
2	1.0
3	1.5
4	2.0
5	2.5
6	3.0
7	3.5

Tabela B.1 – Frequências espaciais (em ciclos/bloco) correspondentes aos coeficientes DCT.

Considerando que cada imagem é constituída por N_V linhas (ou N_V pixels na vertical), o número de pixels contidos na fracção α da altura do écran, segundo o ângulo de visualização θ , é:

$$n = \frac{d}{H} \cdot N_V \cdot \theta . \quad (\text{B.5})$$

As frequências espaciais correspondentes a cada um dos coeficientes DCT são habitualmente expressas em ciclos por bloco (tabela B.1). Como as dimensões de cada bloco são de 8×8 pixels, conclui-se que:

$$1 \text{ ciclo/bloco} = \frac{1}{8} \text{ ciclos/pixel} . \quad (\text{B.6})$$

Segundo (B.5), dN_V / H corresponde ao número de pixels por radiano do ângulo de visualização, o que permite rescrever a relação (B.6) em termos de ciclos por radiano:

$$1 \text{ ciclo/bloco} = \frac{dN_V}{8H} \text{ ciclos/radiano} , \quad (\text{B.7})$$

ou, atendendo a que π radianos correspondem a 180 graus:

$$1 \text{ ciclo/bloco} = \frac{d\pi N_V}{1440H} \text{ ciclos/grau} . \quad (\text{B.8})$$

A expressão anterior também é válida para a direcção horizontal. Designando por N_H o número de pixels por linha, obtém-se:

$$1 \text{ ciclo/bloco} = \frac{d\pi N_H}{1440H} \text{ ciclos/grau} . \quad (\text{B.9})$$

j	f_{0j} (ciclos / grau)		
	$d=4H$	$d=5H$	$d=6H$
1	3.142	3.927	4.712
2	6.283	7.854	9.425
3	9.425	11.781	14.137
4	12.566	15.708	18.850
5	15.708	19.635	23.562
6	18.850	23.562	28.274
7	21.991	27.489	32.987

Tabela B.2 – Frequências espaciais horizontais com $N_H = 720$ (em ciclos/grau).

i	f_{i0} (ciclos / grau)		
	$d=4H$	$d=5H$	$d=6H$
1	2.513	3.142	3.770
2	5.027	6.283	7.540
3	7.540	9.425	11.310
4	10.053	12.566	15.080
5	12.566	15.708	18.850
6	15.080	18.850	22.619
7	17.593	21.991	26.389

Tabela B.3 – Frequências espaciais verticais com $N_V = 576$ (em ciclos/grau).

A título de exemplo, foram calculadas as frequências espaciais correspondentes ao modelo perceptual, em ciclos por grau, para resoluções de $N_H = 720$ pixels, $N_V = 576$ pixels e para diferentes valores da distância de observação – d . Os resultados obtidos encontram-se sintetizados nas tabelas B.2 e B.3.

Anexo C

Estrutura do desmodulador

Neste anexo apresenta-se a dedução detalhada da estrutura do desmodulador utilizado no esquema de extracção da marca-de-água no domínio da frequência.

Atendendo ao esquema de marcas-de-água apresentado ao longo deste trabalho, o valor do espaço Y na posição (m,n) , pode ser obtido através de:

$$Y(m,n) = X(m,n) + w(m,n), \quad (C.1)$$

ou, designando o par (m,n) por $[m]$,

$$Y[m] = X[m] + w[m] \Leftrightarrow X[m] = Y[m] - w[m]. \quad (C.2)$$

No espaço X , a f.d.p.³⁰ conjunta dos valores $X[m]$ (coeficientes DCT), assumindo independência estatística entre coeficientes, pode ser expressa por:

$$f(X[m]) = \prod_m A[m] e^{-|B[m]X[m]|^{c[m]}}, \quad (C.3)$$

em que $A[m]$, $B[m]$ e $c[m]$ são os parâmetros que caracterizam a distribuição gaussiana generalizada característica dos coeficientes DCT que compõem o espaço X . Após inserção da marca e por substituição de (C.2) em (C.3), obtém-se:

$$f(Y[m]|w[m]) = \prod_m A[m] e^{-|B[m](Y[m]-w[m])|^{c[m]}}. \quad (C.4)$$

Considere-se agora que foi inserida a marca-de-água w_l . Designando por w_m todas as marcas-de-água diferentes de w_l e assumindo marcas equiprováveis, a decisão sobre a marca inserida que minimiza a probabilidade de erro de bit na extracção é dada pelo teste de máxima verosimilhança [44], no qual a marca estimada é a que satisfaz:

$$f(Y[m]|w_l[m]) > f(Y[m]|w_m[m]), \quad \forall_{m \neq l}, \quad (C.5)$$

³⁰ f.d.p. – função densidade de probabilidade.

ou

$$\frac{f(Y[\mathbf{m}]|w_l[\mathbf{m}])}{f(Y[\mathbf{m}]|w_m[\mathbf{m}])} > 1, \forall_{m \neq l}. \quad (\text{C.6})$$

Aplicando a função logaritmo a (C.6) tem-se:

$$\log \frac{f(Y[\mathbf{m}]|w_l[\mathbf{m}])}{f(Y[\mathbf{m}]|w_m[\mathbf{m}])} > 0, \forall_{m \neq l}. \quad (\text{C.7})$$

Substituindo (A.4) em (A.7) e após algumas manipulações algébricas simples, obtém-se:

$$\sum_m B[\mathbf{m}]^{c[m]} \left(|Y[\mathbf{m}] - w_m[\mathbf{m}]|^{c[m]} - |Y[\mathbf{m}] - w_l[\mathbf{m}]|^{c[m]} \right) > 0, \forall_{m \neq l}. \quad (\text{C.8})$$

Considere-se agora a extracção de um bit b_i da marca, previamente inserido no conjunto de posições S_i . Admitindo que é utilizada sinalização binária antipodal ($b_i \in \{-1, 1\}$), três situações distintas podem ocorrer:

$$b_{il} = b_{im} \quad (1)$$

$$b_{il} = 1 \text{ e } b_{im} = -1 \quad (2)$$

$$b_{il} = -1 \text{ e } b_{im} = 1 \quad (3)$$

Considere-se ainda que é utilizada a sequência de espalhamento s para representar o bit 1. Utilizando sequências bi-ortogonais, a sequência que representa o bit 0 corresponde a $-s$.

No caso (1), quer se decida ou não pela marca correcta, é extraído o bit correcto.

Para o caso (2) e tendo em conta (C.8), resulta:

$$\sum_{m \in S_i} B[\mathbf{m}]^{c[m]} \left(|Y[\mathbf{m}] + \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[m]} - |Y[\mathbf{m}] - \alpha[\mathbf{m}]s[\mathbf{m}]|^{c[m]} \right) > 0. \quad (\text{C.9})$$

Para o caso (3) vem:

$$\begin{aligned}
& \sum_{m \in S_i} B[m]^{c[m]} \left(|Y[m] - \alpha[m]s[m]|^{c[m]} - |Y[m] + \alpha[m]s[m]|^{c[m]} \right) > 0 \Leftrightarrow \\
& \Leftrightarrow \sum_{m \in S_i} B[m]^{c[m]} \left(|Y[m] + \alpha[m]s[m]|^{c[m]} - |Y[m] - \alpha[m]s[m]|^{c[m]} \right) < 0.
\end{aligned} \tag{C.10}$$

O bit b_i da marca pode então ser obtido considerando o sinal do valor:

$$r_i = \sum_{m \in S_i} B[m]^{c[m]} \left(|Y[m] + \alpha[m]s[m]|^{c[m]} - |Y[m] - \alpha[m]s[m]|^{c[m]} \right). \tag{C.11}$$

Para modulação multinível ir-se-á mostrar, por indução, que o esquema da estrutura geral do desmodulador (ver secção 3.4 do cap. 3, figura 3.9) é também o adequado para este caso. De acordo com esta estrutura, na desmodulação são utilizadas $M/2$ sequências de espalhamento ortogonais e o desmodulador é constituído por $M/2$ detectores, aos quais se segue a decisão.

Considerando que a saída do j -ésimo detector, na extracção do i -ésimo símbolo da marca-de-água é dada por:

$$r_i^j = \sum_{m \in S_i} B[m]^{c[m]} \left(|Y[m] + \alpha[m]s^j[m]|^{c[m]} - |Y[m] - \alpha[m]s^j[m]|^{c[m]} \right), \tag{C.12}$$

será necessário demonstrar que apenas um dos detectores apresenta à sua saída um valor esperado não nulo. Para tal, suponha-se que é introduzido um símbolo com sequência de espalhamento s^1 no conjunto de posições S_i . Tem-se neste caso:

$$Y[m] = X[m] + \alpha[m]s^1[m], m \in S_i. \tag{C.13}$$

Substituindo (C.13) em (C.12), vem:

$$\begin{aligned}
r_i^j = \sum_{m \in S_i} B[m]^{c[m]} & \left(|X[m] + \alpha[m]s^1[m] + \alpha[m]s^j[m]|^{c[m]} - \right. \\
& \left. - |X[m] + \alpha[m]s^1[m] - \alpha[m]s^j[m]|^{c[m]} \right).
\end{aligned} \tag{C.14}$$

O conjunto de sequências de espalhamento utilizadas na modulação – \mathbf{S}_M – é dado por:

$$\mathbf{S}_M = \{s_1, s_2, \dots, s_{M/2}, -s_{M/2}, -s_{M/2-1}, \dots, -s_1\}. \tag{C.15}$$

Na desmodulação, os valores r_i^j à saída dos detectores são obtidos com base nas primeiras $M/2$ sequências de \mathbf{S}_M . Continuando a supor que foi inserido um símbolo com sequência de espalhamento s^1 , duas situações distintas podem ocorrer:

- (a) A combinação é realizada fazendo $j=1$ em (C.14);
- (b) A combinação é realizada fazendo $j \neq 1$, em (C.14). Neste caso, tem-se $s^1 \cdot s^j = 0$ para qualquer uma das $M/2-1$ sequências restantes.

O caso (a) conduz ao estudado anteriormente para o caso binário. Em relação ao caso (b) o valor de r_i^j será dado por:

$$\begin{aligned}
 r_i^j &= \sum_{m \in S_i} B[m]^{c[m]} \left(\left| X[m] + \alpha[m] s^1[m] + \alpha[m] s^j[m] \right|^{c[m]} - \right. \\
 &\quad \left. - \left| X[m] + \alpha[m] s^1[m] - \alpha[m] s^j[m] \right|^{c[m]} \right) \\
 &= \sum_{m \in S_i} B[m]^{c[m]} \left(\left| X[m] + \alpha[m] (s^1[m] + s^j[m]) \right|^{c[m]} - \right. \\
 &\quad \left. - \left| X[m] + \alpha[m] (s^1[m] - s^j[m]) \right|^{c[m]} \right).
 \end{aligned} \tag{C.16}$$

O valor esperado de r_i^j , para $j \neq 1$, é dado por:

$$\begin{aligned}
 E[r_i^j] &= \sum_{m \in S_i} B[m]^{c[m]} E \left[\left| X[m] + \alpha[m] (s^1[m] + s^j[m]) \right|^{c[m]} - \right. \\
 &\quad \left. - \left| X[m] + \alpha[m] (s^1[m] - s^j[m]) \right|^{c[m]} \right].
 \end{aligned} \tag{C.17}$$

Atendendo às características das sequências de espalhamento, tem-se:

$$P(s^1(m) = s^j(m)) = P(s^1(m) = -s^j(m)) = 1/2, \tag{C.18}$$

resultando, por substituição de (C.18) em (C.17):

$$\begin{aligned}
 E[r_i^j] &= \sum_{m \in S_i} B[m]^{c[m]} \left[\frac{1}{2} \left(\left| X[m] + 2\alpha[m] s^1[m] \right|^{c[m]} - |X[m]|^{c[m]} \right) + \right. \\
 &\quad \left. + \frac{1}{2} \left(|X[m]|^{c[m]} - \left| X[m] + 2\alpha[m] s^1[m] \right|^{c[m]} \right) \right] \\
 &= 0.
 \end{aligned} \tag{C.19}$$

Fica então demonstrado que, para os detectores em que $j \neq 1$, o valor esperado para r_i será nulo. Deste modo, a decisão sobre um determinado símbolo deve ser feita tendo em conta a saída do detector com maior valor absoluto, ficando a decisão limitada a dado símbolo, ou ao seu antipodal. Esta segunda decisão é realizada tendo em conta o sinal de r_i , à semelhança do caso binário (em que existem apenas dois símbolos antipodais).

Anexo D

Classificação de tramas na norma MPEG-2 vídeo

Na norma MPEG-2 vídeo existem três tipos básicos de tramas comprimidas – tramas *intra* (ou tramas **I**), tramas *predictas* (ou tramas **P**) e tramas *bidireccionais* (ou tramas **B**).

As tramas **I** são codificadas sem qualquer tipo de referência a outras tramas. Cada trama **I** é tratada como uma imagem e as suas componentes de cor são codificadas de forma independente. Deste modo, a taxa de compressão obtida pelas tramas **I** é relativamente pequena face aos restantes tipos de tramas. O número – N – de tramas predictas (**P** ou **B**) contidas entre duas tramas **I** consecutivas varia tipicamente entre 3 e 12, sendo a estrutura resultante designada por *GOP N* (*Group of Pictures N*).

A codificação de uma trama **P** é feita com base numa trama **I** ou numa trama **P** precedente. Nesta codificação utiliza-se estimação e compensação do movimento, pelo que a taxa de compressão obtida em tramas **P** é significativamente maior do que a obtida em tramas **I**. O número de tramas contidas entre uma trama **P** e a trama **P** (ou **I**) precedente correspondente é designado por *passo de predição* (PP). O valor de PP varia tipicamente entre 1 e 3. Na figura D.1 ilustra-se uma estrutura de tramas *GOP 3*, contendo apenas tramas dos tipos **I** e **P**, e com passo de predição unitário.

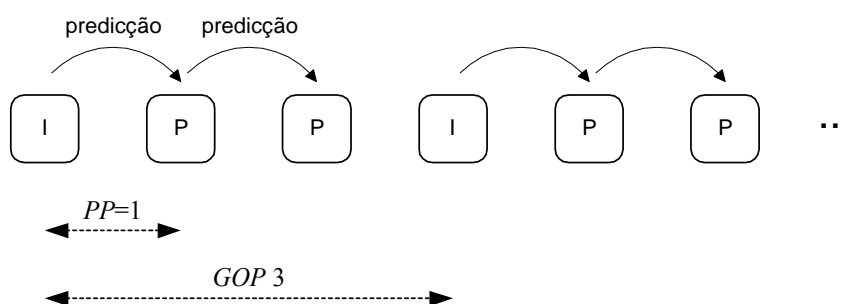


Figura D.1 – Estrutura de tramas *GOP 3*, com passo de predição 1.

O conteúdo das tramas **B** é predicto a partir não só do passado, mas também do futuro, i.e., a predição é bidireccional. Deste modo, cada trama **B** é obtida a partir da trama **P** (ou **I**) precedente e da trama **P** (ou **I**) subsequente. As tramas **B** são as que atingem uma maior taxa de

compressão. A figura D.2 ilustra uma sequência de codificação contendo tramas de todos os tipos referidos, segundo uma estrutura *GOP* 6, com passo de predição 3.

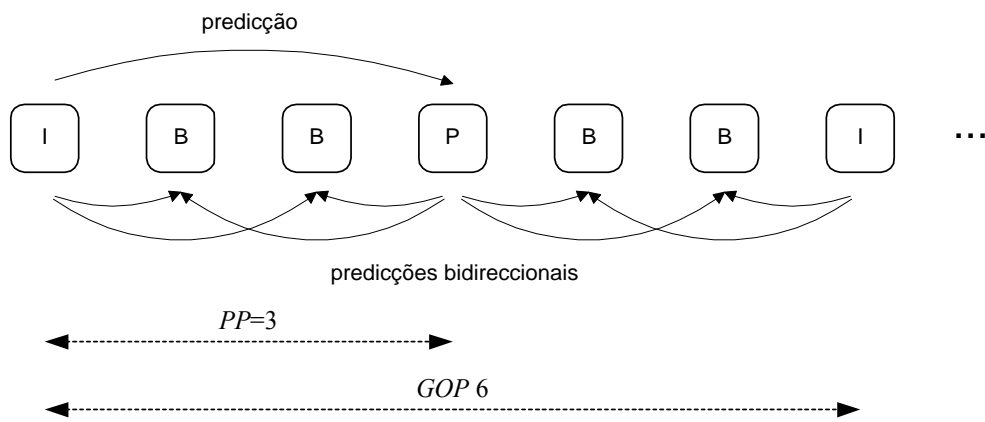


Figura D.2 – Estrutura de tramas *GOP* 6, com passo de predição 3.

