Proteção dos direitos autorais de imagem estática utilizando Criptografia Visual e Marca d'Água

Eduardo Almeida Feijó

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Ciência da Computação
Orientador: Prof. Dr. Routo Terada

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, maio de 2016

Proteção dos direitos autorais de imagem estática utilizando Criptografia Visual e Marca d'Água

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 18/05/2016. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Routo Terada (orientador) - IME-USP
- Prof. Dr. Marco Dimas Gubitoso - IME-USP
- Prof^{a}. Dr^{a}. Denise Hideko Goya - UFABC

Agradecimentos

Agradeço em primeiro lugar à minha esposa Teca, que compreendeu a importância deste meu desafio e me apoiou em todos os momentos.

Agradeço ao meu orientador Prof. Dr. Routo Terada pelo incentivo permanente e por acreditar em meu potencial.

Agradeço ao IME-USP e aos seus integrantes (professores, alunos e funcionários) por me receberem de braços abertos, novamente, após um intervalo de quase 30 anos.

Agradeço aos criadores de imagens estáticas, em especial aos fotógrafos - grupo do qual orgulhosamente faço parte, pela inspiração do tema desta dissertação.

Resumo

A tecnologia atual não oferece prevenção contra *cópia*, *adulteração* ou *plágio* de uma imagem estática em meio digital sem autorização do verdadeiro autor. Dado que tais malfeitos não podem ser evitados, resta ao criador da obra original lutar *a posteriori* por seus direitos nos fóruns adequados (no tribunal, por exemplo). Na época da fotografia analógica com filme, o negativo poderia ser utilizado como prova. Hoje este recurso raramente está disponível e se faz necessária uma solução alternativa. A técnica de Marca d'Água é uma das possibilidades criptográficas existentes para apoiar o autor em sua defesa. O princípio da Marca d'Água é o encapsulamento de informações relevantes, preferencialmente de forma imperceptível, na imagem a ser protegida. Tais informações, quando extraídas da imagem marcada, devem revelar o verdadeiro autor num processo de disputa.

Soluções de Marca d'Água combinada com Criptografia Visual são encontradas na literatura. A principal vantagem deste caminho é a propriedade *Imperceptível por segurança perfeita* que a Marca d'Água assume quando tratada por Criptografia Visual. O segredo (neste caso, a Marca d'Água) é segmentado via Criptografia Visual em 2 *transparências*: uma delas é encapsulada na imagem a ser protegida e a outra é mantida pelo verdadeiro autor. Basta a sobreposição de tais transparências para que a Marca d'Água seja revelada.

Nesta pesquisa propomos um novo método, denominado MACV, que combina Marca d'Água, Criptografia Visual e um algoritmo de hashing. O MACV apresenta, entre outras, as seguintes propriedades desejáveis de Marca d'Água: *imperceptível por segurança perfeita*, *alta entropia*, *armazenamento na própria imagem* e *sem ambiguidade*. Veremos em nossa pesquisa bibliográfica que há uma lacuna de soluções que apresentem, *simultaneamente*, todas estas propriedades. Esta lacuna torna o MACV único em sua categoria.

Palavras-chave: marca d'água, criptografia visual, proteção de direitos autorais.

Abstract

Current technology does not prevent *copy*, *tamper* or *plagiarism* of a still image in digital media without authorization of the true author. Since such misdeeds can not be avoided, the creator of the original work has to defend his rights *a posteriori* in the appropriate forum (in court, for example). A negative could be used as evidence in the era of analog photography. Currently this feature is rarely available and an alternative solution is needed. Watermark is one of the existing cryptographic possibilities to support the author in his defense. A Watermark is embedded into the image to be protected, preferably imperceptibly. Such information when extracted from the watermarked image should reveal the true author in a dispute process.

Solutions using Watermark with Visual Cryptography are found in the literature. The main advantage of this way is the property *Imperceptible by perfect secrecy* that the Watermark assumes when treated by Visual Cryptography. The key (in this case, the Watermark) is partitioned by Visual Cryptography in 2 *transparencies*: one is embedded in the image to be protected and the other is kept by the true author. The Watermark is revealed by overlapping such transparencies.

In this research we propose a new method, called MACV, combining Watermark, Visual Cryptography and a hashing algorithm. The MACV has, among others, the following desirable properties of Watermark: *invisible by perfect secrecy*, *high entropy*, *stored into the image itself* and *unambiguously*. We will see in our research that there is no solution that presents, *simultaneously*, all these properties. This fact makes the MACV the first one to show all these properties

Keywords: digital image watermarking, visual cryptography, copyright protection.

Sumário

Lista de Abreviaturas **ix**

Lista de Símbolos **xi**

Lista de Figuras **xiii**

Lista de Tabelas **xvii**

1 Introdução **1**

1.1 Considerações Preliminares 1

1.2 Objetivos 1

1.3 Contribuições originais 2

1.4 Organização do Trabalho 3

2 Criptografia Visual **5**

2.1 Objetivo do capítulo 5

2.2 Introdução à Criptografia Visual 5

2.3 Modelo Básico: $(2,2) - VCS$ 6

2.3.1 Segurança perfeita 10

2.4 Shamir's Secret Sharing: $(k,n) - TS$ 11

2.5 Modelo Geral: $(k,n) - VCS$ 12

2.5.1 Caso particular $(2,n) - VCS$ 16

2.5.2 Caso particular $(3,3) - VCS$ 16

2.6 Vulnerabilidade em Criptografia Visual 17

2.7 Conclusões do capítulo 20

3 Marca d'Água **21**

3.1 Objetivo do capítulo 21

3.2 Introdução à Marca d'Água 21

3.2.1 Classificação de Marcas d'Água 23

3.2.2 Definições 25

3.2.3 Valor comercial de uma imagem 26

3.3 Propriedades ideais da Marca d'Água 28

3.4 Marca d'Água de Cox 30
3.4.1 Camada *conteúdo* de Cox 31
3.4.2 Camada *transporte* de Cox 31
3.5 Ataques à Marca d'Água 35
3.6 Conclusões do capítulo 36

4 Método MACV **37**
4.1 Objetivo do capítulo 37
4.2 ideias principais do MACV 37
4.3 Algoritmos do MACV 42
4.3.1 MACV.g - Gerador de Marca d'Água do MACV 42
4.3.2 MACV.e - Encapsulador de Marca d'Água do MACV 43
4.3.3 MACV.v - Verificador de Marca d'Água do MACV 44
4.4 Resultados do MACV 46
4.4.1 Aplicação dos algoritmos MACV a uma imagem em tons de cinza . . 47
4.4.2 Aplicação dos algoritmos MACV a uma imagem colorida 51
4.5 Análise de segurança do MACV 54
4.5.1 Segurança da camada de *transporte* do MACV 55
4.5.2 Segurança da camada de *conteúdo* do MACV 64
4.6 Análise das propriedades do MACV 65
4.7 Soluções baseadas em Criptografia Visual e Marca d'Água 67
4.8 Conclusões do capítulo 68

5 Conclusões **71**
5.1 Considerações Finais 71
5.2 Sugestões para Pesquisas Futuras 72

Referências Bibliográficas **75**

Lista de Abreviaturas

$(2,2)-VCS$	Modelo básico de Criptografia Visual, esquema 2×2
$(k,n)-VCS$	Modelo geral de Criptografia Visual, esquema $k \times n$
$(k,n)-TS$	*Threshold Scheme* do SSS
SSS	*Shamir's Secret Sharing*
$H(V)$	Distância de *Hamming*
DCT	*Discrete Cosine Transform*
$IDCT$	Inversa da *Discrete Cosine Transform*
DWT	*Discrete Wavelet Transform*
LSB	*Least Significant Bit*
EIC	Entidade Independente de Certificação
$I_{Original}$	Imagem original a ser protegida
$I_{Marcada}$	Imagem hospedeira de uma Marca d'Água
$I_{Litigio}$	Imagem em disputa de autoria
M_{Autor}	Marca d'Água de autoria
$M_{Secreta}$	Transparência gerada pelo $(2,2)-VCS$, mantida em segurança pelo autor
$M_{Publica}$	Transparência gerada pelo $(2,2)-VCS$, encapsulada em $I_{Marcada}$
$M_{Extraida}$	Marca d'Água extraída de $I_{Litigio}$
SHA	*Secure Hash Algorithm*
$ASCII$	*American Standard Code for Information Interchange*

Lista de Símbolos

α Em Criptografia Visual, indica o contraste entre brancos e pretos

α Em Marca d'Água, indica o fator de escala no método de Cox

Lista de Figuras

2.1 Exemplo do $(2,2) - VCS$. As transparências A e B são geradas a partir de X. A sobreposição de A por B (ou B por A) revela X. 6
2.2 Imagem secreta X com 1181×295 pixels 7
2.3 Matriz M contendo 2×2 sub-pixels: há 6 combinações possíveis de dois 0 e dois 1 7
2.4 Codificação de um pixel P (branco ou preto) nas transparências A e B 8
2.5 Ampliação e a codificação de alguns pixels da imagem secreta X 9
2.6 Transparências A e B têm 2362×590 pixels cada 9
2.7 Imagem obtida pela sobreposição das transparências A e B 9
2.8 Sobreposição das transparências A e B após filtro computacional 10
2.9 Exemplo da matriz booleana $S = [s_{i,j}]$, onde $1 \leq i \leq n$ e $1 \leq j \leq m$ 14
2.10 Exemplo de vetor V com $H(V) = m - 5$ 14
2.11 Fraude contra o $(2,3) - VCS$ 18

3.1 Fluxo de uso de Marca d'Água em 3 momentos: em ambiente protegido, em ambiente desprotegido (neste fluxo, na Internet) e em caso de litígio 23
3.2 Exemplo de Marca d'Água visível 24
3.3 Exemplo de Marca d'Água invisível: a imagem original e sua respectiva imagem marcada com 33,5% dos bits sutilmente alterados sem aparentes diferenças 24
3.4 Exemplo de eliminação de Marca d'Água com evidente perda de *valor comercial* 27
3.5 Fluxo de inserção de Cox, considerando uma Marca d'Água na distribuição $\{0,1\}$ 33
3.6 Fluxo de extração de Cox, obtendo uma Marca d'Água na distribuição $\{0,1\}$ 34
3.7 Fluxo com categorias de ataque (intencional ou não) 35

4.1 Perda de *valor comercial* de uma imagem marcada ao ser atacada para remoção da Marca d'Água robusta 38
4.2 Construção do novo método - Etapa 1/ 5: adotar a camada *transporte* de Cox e excluir a camada *conteúdo* 39
4.3 Construção do novo método - Etapa 2/ 5: não usar informações de baixa entropia na camada *conteúdo* 40

4.4 Construção do novo método - Etapa 3/ 5: usar informação com significado e segurança perfeita na camada *conteúdo* 40
4.5 Construção do novo método - Etapa 4/ 5: usar informação com significado criptografada pelo $(2,2)-VCS$ na camada *conteúdo* 41
4.6 Construção do novo método - Etapa 5/ 5: aplicar hash e $(2,2)-VCS$ para produzir a camada *conteúdo* 41
4.7 Fluxo de execução do MACV em diferentes ambientes: protegido, Internet e no tribunal 42
4.8 Fluxo de execução do MACV.g 43
4.9 Fluxo de execução do MACV.e 44
4.10 Fluxo de execução do MACV.v. A imagem original $I_{Original}$ e a respectiva transparência $M_{Secreta}$ são apresentadas pelo autor verdadeiro. A imagem em disputa $I_{Litigio}$ é apresentada pela outra parte litigante 45
4.11 $I_{Original}$: imagem em tons de cinza *Pescador Tailandês* com 1.000×666 pixels 48
4.12 Aplicação de MACV.g à imagem original *Pescador Tailandês*. As resultantes $M_{Publica}$ e $M_{Secreta}$ têm 200×40 pixels 48
4.13 Fluxo de encapsulamento da Marca d'Água $M_{Publica}$ pelo MACV.e na imagem original *Pescador Tailandês* 49
4.14 A imagem $I_{Marcada}$ tem 1.000×666 pixels e é o resultado da aplicação do MACV.e à imagem original *Pescador Tailandês*. As imagens $I_{Original}$ e a respectiva $I_{Marcada}$ são visualmente muito parecidas, embora tenham 38,4% pixels diferentes entre si 49
4.15 A imagem $I_{Marcada}$ é gerada com $\alpha = 0,1$ ao inves de $\alpha = 0,01$ 50
4.16 Fluxo de verificação da Marca d'Água pelo MACV.v aplicado à imagem *Pescador Tailandês* 50
4.17 Imagem colorida *Senhora Indiana* com 1.000×1.000 pixels 51
4.18 Aplicação de MACV.g à imagem original *Senhora Indiana*. As resultantes $M_{Publica}$ e $M_{Secreta}$ têm 200×40 pixels 52
4.19 Fluxo de encapsulamento da Marca d'Água $M_{Publica}$ pelo MACV.e na imagem original *Senhora Indiana* 52
4.20 A imagem $I_{Marcada}$ tem 1.000×1.000 pixels e é o resultado da aplicação do MACV.e à imagem original *Senhora Indiana*. As imagens $I_{Original}$ e a respectiva $I_{Marcada}$ são visualmente muito parecidas, embora tenham 25,6% pixels diferentes entre si 53
4.21 Fluxo de verificação da Marca d'Água pelo MACV.v aplicado à imagem *Senhora Indiana* 54
4.22 Ataque: redução da resolução de $I_{Marcada}$ em 36%. Mostramos a $I_{Litigio}$ com 800×800 pixels e a respectiva $M_{Extraida}$ 55
4.23 Ataque: redução da resolução de $I_{Marcada}$ em 75%. Mostramos a $I_{Litigio}$ com 500×500 pixels e a respectiva $M_{Extraida}$ 56

4.24 Ataque: redução da resolução de $I_{Marcada}$ em 99%. Mostramos a $I_{Litigio}$ com 100×100 pixels e a respectiva $M_{Extraida}$ 57
4.25 Ataque: interpolação de $I_{Marcada}$ em 400%. Mostramos a $I_{Litigio}$ com 2.000×2.000 pixels e a respectiva $M_{Extraida}$ 58
4.26 Ataque: compactação JPEG com máxima qualidade. Mostramos a $I_{Litigio}$ com 1,1 MB e a respectiva $M_{Extraida}$ 58
4.27 Ataque: compactação JPEG com média qualidade. Mostramos a $I_{Litigio}$ com 0,4 MB e a respectiva $M_{Extraida}$ 59
4.28 Ataque: rotação 90º. Mostramos a $I_{Litigio}$ rotacionada em 90º e a respectiva $M_{Extraida}$ 60
4.29 Ataque: conversão para P&B. Mostramos a $I_{Litigio}$ rconvertida para P&B e a respectiva $M_{Extraida}$ 60
4.30 Ataque: deformação horizontal. Mostramos a $I_{Litigio}$ com 1.000×500 pixels deformada na horizontal e a respectiva $M_{Extraida}$ 61
4.31 Ataque: deformação vertical. Mostramos a $I_{Litigio}$ com 200×1.000 pixels deformada na vertical e a respectiva $M_{Extraida}$ 62
4.32 Ataque: filtro cor automática do Photoshop. Mostramos a $I_{Litigio}$ com 1.000×1.000 pixels transformada pelo filtro cor automática do Photoshop e a respectiva $M_{Extraida}$ 62
4.33 Ataque: corte. Mostramos a $I_{Litigio}$ com 1.000×500 pixels obtida pelo corte de 50% da área de $I_{Marcada}$ e a respectiva $M_{Extraida}$ 63

Lista de Tabelas

2.1 Para cada pixel em X, todas as combinações em A, B e C e os resultados das sobreposições $A + B$, $A + C$ e $B + C$ 18
2.2 Construção de A' e B' a partir de um pixel branco 19
2.3 Construção de A' e B' a partir de um pixel preto 19

4.1 Resumo dos ataques realizados a $I_{Marcada}$ 64

Capítulo 1

Introdução

1.1 Considerações Preliminares

Um dos grandes desafios de segurança da nossa era digital é a proteção dos direitos autorais de imagem estática. No Brasil, conforme a Lei 9.610 de 19 de fevereiro de 1998, o desrespeito ao direito autoral é considerado crime. Porém a lei não é suficiente para garantir que o autor tenha seus direitos preservados em ambiente digital.

Em Criptografia, técnicas são geralmente utilizadas para *esconder* uma informação, *proteger* sua integridade e *identificar* o respectivo autor. No entanto, na Internet, estas mesmas técnicas são inócuas para evitar que a imagem seja *copiada*, *adulterada* ou *plagiada* sem autorização do verdadeiro autor. Métodos específicos são necessários para combater estes problemas, não tratados pela Criptografia tradicional.

Entre estes métodos encontramos a tecnologia de Marca d'Água. Deve ficar claro que o uso de Marca d'Água não impede que o conteúdo multimídia seja *copiado*, *adulterado* ou *plagiado* sem autorização do verdadeiro autor. O real benefício da Marca d'Água é que, em caso de *litígio*, ela pode ser utilizada como prova da verdadeira autoria.

Inúmeras técnicas de Marca d'Água são encontradas na literatura. Muitos destes métodos não apresentam determinadas propriedades desejáveis em Marca d'Água, como a de ser *Imperceptível por segurança perfeita* ou de apresentar *Alta entropia*.

Neste cenário, uma nova solução de Marca d'Água para proteção dos direitos autorais de imagem estática precisa ser projetada e ter sua segurança avaliada. Esta nova técnica deve prover que a Marca d'Água seja *Imperceptível por segurança perfeita* e que tenha *Alta entropia*.

1.2 Objetivos

O objetivo do nosso trabalho é descrever um novo algoritmo, denominado MACV, que procura apresentar as seguintes propriedades desejáveis de Marca d'Água: armazenada na própria imagem; visualmente imperceptível quando inserida; visualmente significativa quando

extraída; irreproduzível por terceiros não autorizados; publicamente verificável; indelével por manipulação não autorizada; resistente a certas operações de processamento de imagem; aplicável a formatos com e sem perdas, e a imagens binárias, em níveis de cinza e coloridas; eficiente em tempo de processamento e espaço de armazenamento; sem ambiguidade; imperceptível por segurança perfeita; alta entropia.

O MACV apresenta um conjunto de propriedades únicas em relação às técnicas de Marca d'Água encontradas na literatura. Em particular, ressaltamos como diferenciais do MACV as propriedades *Imperceptível por segurança perfeita*, *Alta entropia*, *Armazenamento na própria imagem* e *Sem ambiguidade*. Detalharemos estas propriedades adiante.

A propriedade *Imperceptível por segurança perfeita* é obtida pelo uso da Criptografia Visual, método criado por NAOR e SHAMIR [NS95]. Ao contrário da grande maioria dos processos criptográficos conhecidos, onde computação é exigida na etapa de decodificação, na Criptografia Visual a visão humana é suficiente para a revelação do segredo.

Além da conceituação teórica do MACV apresentaremos também os resultados experimentais obtidos e a análise de sua segurança. Um levantamento bibliográfico atualizado de técnicas de Marca d'Água baseadas em Criptografia Visual será mostrado e cada uma das soluções descritas será comparada ao MACV.

1.3 Contribuições originais

As contribuições originais deste trabalho são as seguintes:

1. Proposta de método original (MACV) para proteção dos direitos autorais de imagem estática (Capítulo 4)

2. Segmentação da solução de Marca d'Água em 2 camadas independentes: *transporte* e *conteúdo* (Capítulo 3.4)

3. Acréscimo de novas propriedades ideais de Marca d'Água: *Sem ambiguidade*; *Imperceptível por segurança perfeita*; *Alta entropia*; *Capacidade de absorção da Marca d'Água* (Seção 3.3)

4. Conceituação de *valor comercial* de uma imagem: embora ataques possam eliminar qualquer traço da Marca d'Água contida em uma imagem, muitos destes ataques degradam de tal forma a imagem resultante que ela não possui qualquer *valor comercial*, frustrando o intento do atacante (Seção 3.2.3)

5. Práticas complementares para minimizar os riscos do autor verdadeiro da imagem (Seção 3.2.3)

1.4 Organização do Trabalho

No Capítulo 2, apresentamos a tecnologia Criptografia Visual em seu Modelo Básico $(2,2) - VCS$[1] e em seu Modelo Geral $(k,n) - VCS$; provamos a *segurança perfeita* do $(2,2) - VCS$ e mostramos uma vulnerabilidade do $(2,n) - VCS$ para $n > 2$.

No Capítulo 3 introduzimos o conceito de Marca d'Água e exploramos suas classificações e propriedades; estudamos a técnica proposta por Cox e a segmentamos em camadas de *transporte* e *conteúdo*; analisamos os principais tipos de ataque à Marca d'Água.

No Capítulo 4 descrevemos o algoritmo original proposto MACV; segmentamos o MACV em seus módulos principais: Gerador, Encapsulador e Verificador de Marca d'Água; apresentamos os resultados experimentais obtidos; discutimos os aspectos de segurança do MACV; comparamos o MACV com as técnicas similares encontradas na literatura disponível.

O Capítulo final 5 apresenta nossas conclusões sobre o tema estudado e sugere temas para futuras pesquisas.

[1] Visual Cryptography Scheme

Capítulo 2

Criptografia Visual

2.1 Objetivo do capítulo

O objetivo deste capítulo é apresentar o sistema criptográfico denominado Criptografia Visual criado por NAOR e SHAMIR [NS95]. Ao contrário da grande maioria dos processos criptográficos conhecidos, onde computação é exigida na etapa de decodificação, na Criptografia Visual a visão humana é suficiente para a revelação do segredo.

Este capítulo está estruturado da seguinte forma: a Seção 2.2 apresenta uma breve introdução à Criptografia Visual incluindo uma amostra do experimento realizado. A Seção 2.3 apresenta o denominado Modelo Básico: $(2,2)-VCS$ e demonstra sua condição de segurança perfeita. A Seção 2.4 apresenta o Shamir's Secret Sharing, do qual o $(k,n)-VCS$ pode ser considerado uma variação visual. A Seção 2.5 apresenta o Modelo Geral: $(k,n)-VCS$ e estuda alguns casos particulares. A Seção 2.6 mostra uma vulnerabilidade do $(k,n)-VCS$, especialmente quando $k=2$ e $n>2$. A Seção 2.7 apresenta um resumo dos benefícios da Criptografia Visual e mostra o cenário de uso prático da tecnologia.

2.2 Introdução à Criptografia Visual

Criptografia Visual é um processo criptográfico originalmente apresentado por NAOR e SHAMIR [NS95]. Neste processo, extremamente simples em sua concepção, uma imagem cifrada é revelada sem a necessidade de computação.

Para facilitar o entendimento deste processo vamos utilizar o modelo básico da Criptografia Visual. Considere X uma imagem binária, isto é, formada por pixels brancos e pretos. Esta imagem X é nosso segredo. A partir de X serão geradas duas transparências (ou *shares*, como usualmente encontramos na literatura) A e B. Cada uma destas transparências, individualmente, não revela qualquer informação sobre X. No entanto, a simples sobreposição da transparência A por B (ou B por A) revelará a imagem secreta X. Denominaremos o modelo básico da Criptografia Visual como $(2,2)-VCS$. A Figura 2.1 apresenta um exemplo do processo do $(2,2)-VCS$.

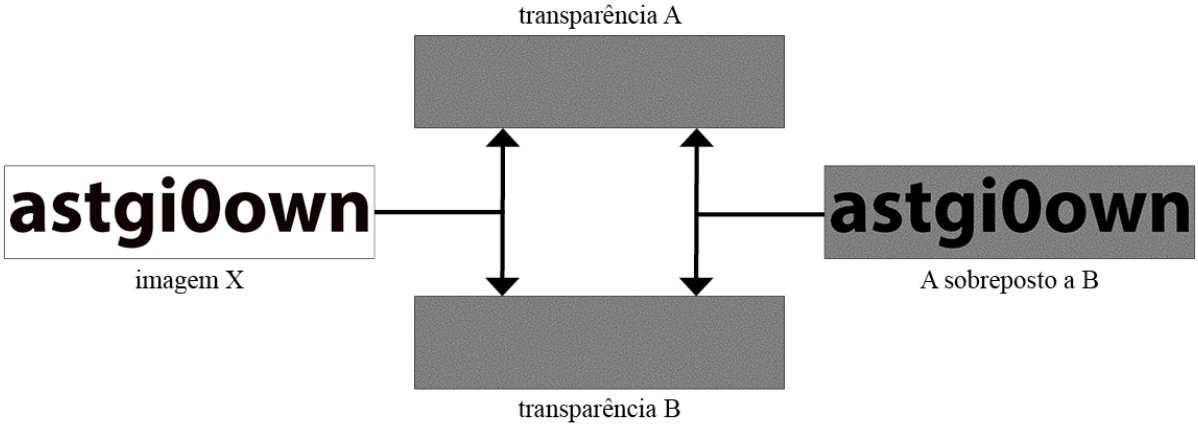

Figura 2.1: *Exemplo do* $(2,2) - VCS$. *As transparências A e B são geradas a partir de X. A sobreposição de A por B (ou B por A) revela X.*

Em seu modelo geral a Criptografia Visual pode ser entendida como uma derivação visual do processo de distribuição segura de chaves de SHAMIR [Sha79] denominado SSS (*Shamir's Secret Sharing*). No SSS, um segredo D é segmentado em n pedaços. O conhecimento de k ou mais pedaços torna D facilmente calculado. O conhecimento de quaisquer $k-1$ ou menos pedaços tornaria computacionalmente difícil calcular D.

Na Criptografia Visual, a partir de uma imagem X são geradas n transparências. A sobreposição de k ou mais destas transparências revela o segredo X. A sobreposição de quaisquer $k-1$ destas transparências não tem qualquer significado visual. Denominaremos o modelo geral da Criptografia Visual como $(k,n) - VCS$.

Além do modelo original criado por NAOR e SHAMIR [NS95], baseado em imagens binárias, encontramos na literatura (WEIR e YAN [WY12]) variações que tratam: imagens em tons de cinza, imagens coloridas, transparências que revelam múltiplos segredos, etc. O estudo destas variações não faz parte do escopo desta dissertação.

Embora seja notável o interesse do mundo acadêmico por esta recente e promissora tecnologia, percebemos que as aplicações práticas e comerciais ainda estão em estágio embrionário.

2.3 Modelo Básico: $(2,2) - VCS$

Como visto na Seção 2.2, o $(2,2) - VCS$ codifica uma imagem secreta X em 2 transparências A e B. As principais características do $(2,2) - VCS$ são:

- A imagem X é binária, isto é, formada por pixels brancos (codificados como 0) e pretos (codificados como 1). O pixel branco também pode ser interpretado como um pixel transparente. Há variações do $(2,2) - VCS$ específicas para imagens em tons de cinza ou coloridas

- O processo de geração das transparências A e B é realizado pixel a pixel a partir de

X. Cada pixel será codificado de forma independente dos demais. Desta forma pode-se entender a mensagem a ser protegida (a imagem X de n pixels) como um conjunto de n mensagens (de 1 pixel cada)

- Observadas de forma isolada, as transparências A e B se assemelham a ruído aleatório e não revelam qualquer informação sobre a imagem secreta X. É possível perceber essa característica das transparências A e B no exemplo mostrado na Figura 2.1. Veremos neste capítulo que, mesmo que um espião obtenha A ou B, ele não consegue qualquer pista do segredo contido na imagem X

- Conforme sugerido por NAOR e SHAMIR [NS95], a transparência A (ou a B) pode ser entendida como a chave do segredo

- A visão humana é suficiente para revelar a imagem secreta X a partir da sobreposição da transparência A pela B (ou B pela A). As transparências precisam estar perfeitamente alinhadas para que a sobreposição revele o segredo. Não é necessária qualquer computação nesta etapa

O algoritmo de geração das transparências A e B a partir da imagem secreta X pelo $(2,2) - VCS$ é simples e eficiente. Já citamos que cada pixel de X é tratado de forma independente. Desta forma, todo e qualquer pixel P de X é codificado por uma matriz M contendo 2×2 sub-pixels, sendo que 2 sub-pixels recebem valor 0 e 2 sub-pixels recebem valor 1. Nesta Seção vamos utilizar como exemplo de imagem secreta X a Figura 2.2 com 1181×295 pixels.

Figura 2.2: *Imagem secreta X com* 1181×295 *pixels*

A Figura 2.3 apresenta todas as 6 possíveis combinações de codificação da matriz M. A escolha entre estas 6 combinações de M é aleatória. É possível perceber as transparências A e B terão o dobro de linhas e o dobro de colunas em relação à imagem original X (tamanho das transparências A e B $= 4 \times n$).

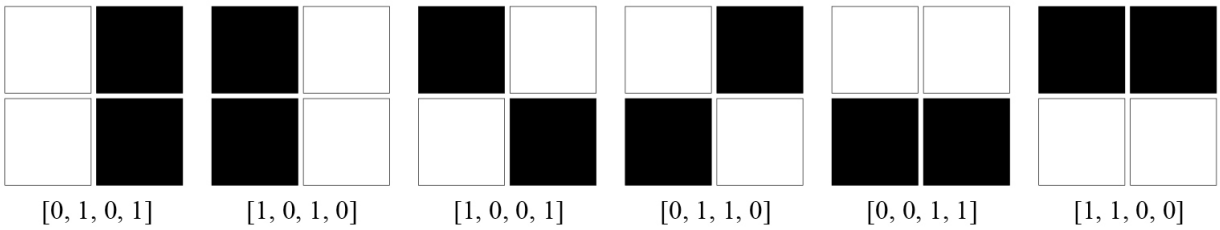

Figura 2.3: *Matriz M contendo* 2×2 *sub-pixels: há 6 combinações possíveis de dois 0 e dois 1*

A codificação percorre X de cima para baixo, da esquerda para a direita, construindo as transparências A e B simultaneamente e na mesma orientação:

- Se o pixel P é branco então as transparências A e B recebem a mesma matriz M
- Se o pixel P é preto então A recebe a matriz M e B recebe a matriz M' (com 2×2 sub-pixels), onde cada sub-pixel de M' é obtido pela operação XOR[1] de 1 com o respectivo sub-pixel de M

A Figura 2.4 ilustra todas as possíveis codificações de um pixel P (branco ou preto) nas transparências A e B. É possível observar que se um espião intercepta uma das duas transparências ele não consegue deduzir se o pixel P original era branco ou preto.

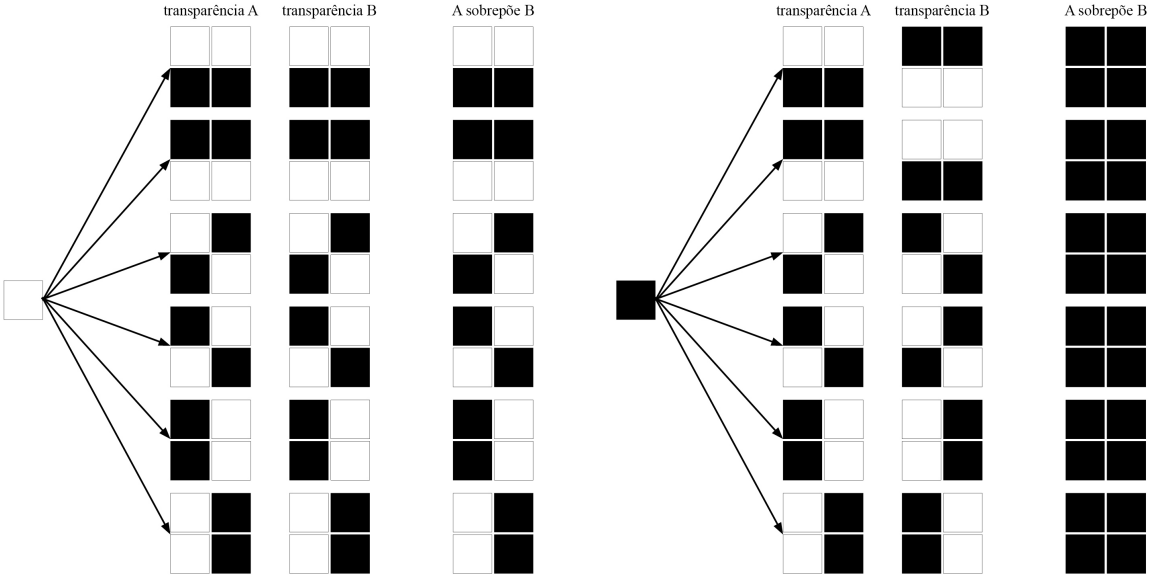

Figura 2.4: *Codificação de um pixel P (branco ou preto) nas transparências A e B*

A Figura 2.5 ilustra a ampliação e a codificação de alguns pixels da imagem secreta X. Observa-se que a sobreposição das transparências A e B codificadas a partir de um pixel preto é uma matriz com 4 sub-pixels pretos que, para a visão humana, será percebida como preto. Por outro lado, a mesma sobreposição a partir de um pixel branco será percebida pela visão como um cinza médio.

As 2 transparências geradas pela aplicação do $(2,2) - VCS$ sobre a imagem secreta X são exibidas na Figura 2.6. Cada uma destas 2 transparências tem 2362×590 pixels.

Um exemplo da sobreposição das transparências A e B pode ser observado na Figura 2.7. O cérebro humano facilmente identifica o segredo como uma sequência de letras (maiúsculas e minúsculas) e números. Isso é possível pois a imagem sobreposta das transparências A e B apresenta contraste suficiente entre pretos e cinzas. O papel do cérebro humano no processo de decodificação (revelar o segredo pela sobreposição das transparências A e B) elimina a

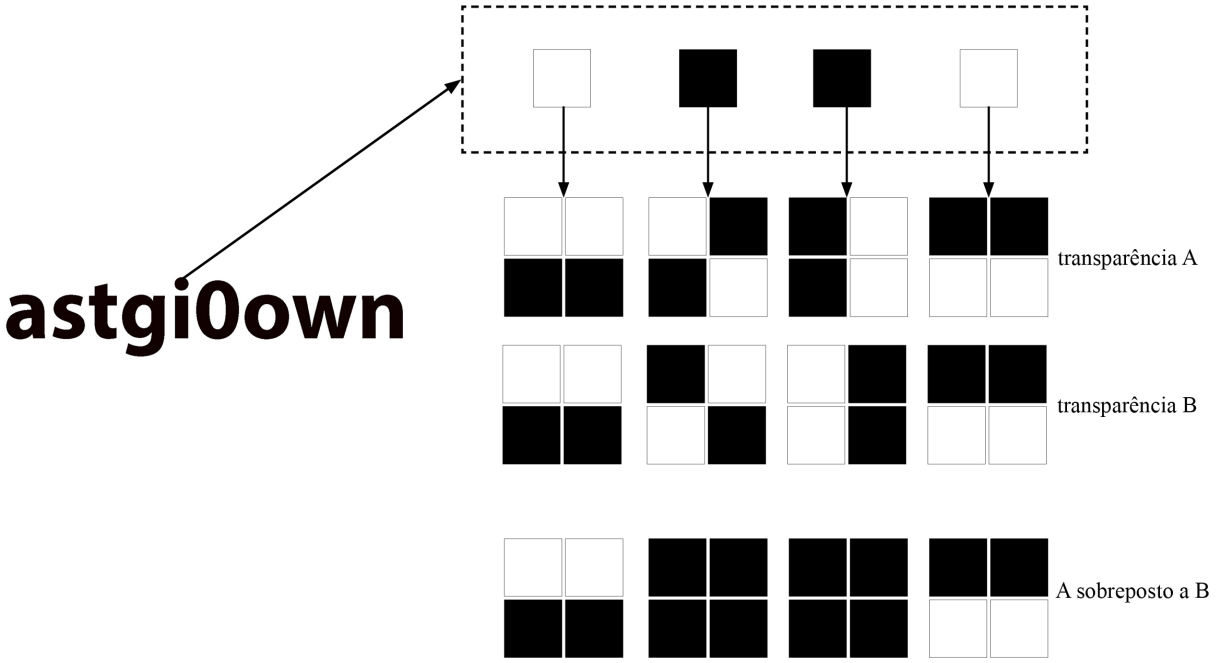

Figura 2.5: *Ampliação e a codificação de alguns pixels da imagem secreta X*

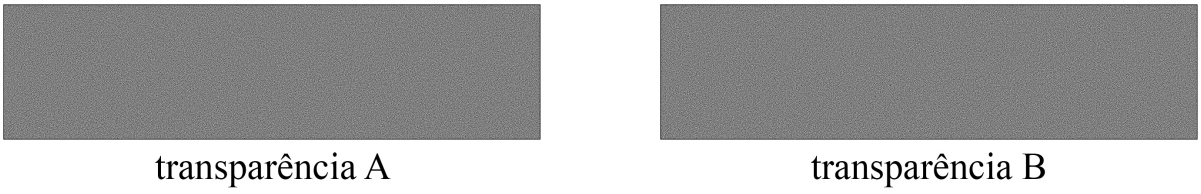

Figura 2.6: *Transparências A e B têm* 2362×590 *pixels cada*

Figura 2.7: *Imagem obtida pela sobreposição das transparências A e B*

necessidade de qualquer cálculo computacional.

Embora o contraste entre pretos e cinzas da sobreposição das transparências A e B seja suficiente para a interpretação do cérebro, é possível que uma aplicação prática necessite que a imagem sobreposta tenha o aspecto da imagem original X, onde vemos brancos no lugar de cinzas. O uso de um filtro computacional transforma o cinza em branco, permitindo que o aspecto visual da imagem original X seja recuperado, conforme exibido na Figura 2.8.

Neste momento já é possível elencar alguns pontos positivos do $(2,2) - VCS$:

- Simplicidade da implementação
- Eficiência da implementação, com tempo de execução linear em relação ao tamanho

[1]XOR (ou exclusivo) é uma operação lógica entre duas variáveis em que o resultado é 1 se e somente se uma das entradas é 1 e a outra é 0

Figura 2.8: *Sobreposição das transparências A e B após filtro computacional*

da imagem original X

- Independência de uso de computador para decodificação (caso as transparências A e B estejam em mídia física)
- Segurança. Esta característica será detalhada na Seção 2.3.1.

Embora classificadas como pontos negativos, as características abaixo são facilmente contornáveis:

- Perda de contraste da imagem sobreposta: em ambiente computacional o cinza médio pode ser convertido para branco conforme o exemplo visto em Figura 2.7 e Figura 2.8. No $(k, n) - VCS$, para k e n grandes, esse problema é mais significativo
- Espaço de armazenamento: tanto as transparências A e B como a imagem sobreposta tem 4 vezes o tamanho da imagem original X

Um dos primeiros artigos que aborda especificamente a expansão de pixels foi apresentado por KUWAKADO e TANAKA [KT99]. Eles mostram uma variação do $(2, 2) - VCS$ na qual não há expansão de pixels: um pixel preto é codificado como um pixel branco na transparência A e um pixel preto na transparência B (ou vice versa). Um pixel branco é codificado como 2 transparências com um pixel branco ou como 2 transparências com um pixel preto. A segurança é a mesma do $(2, 2) - VCS$ original, porém há um efeito colateral na imagem sobreposta: o cinza médio identifica um pixel branco enquanto o branco ou o preto identificam um pixel preto.

2.3.1 Segurança perfeita

Na seção anterior afirmamos que o $(2, 2) - VCS$ é um processo seguro mesmo que um espião conheça uma das duas transparências (A ou B). Essa possibilidade ocorre, por exemplo, se o espião instala uma escuta no canal em que a transparência é transmitida. Nesta seção é provado formalmente que o $(2, 2) - VCS$ apresenta segurança perfeita.

Em Criptografia entendemos como segurança perfeita que, mesmo que o espião intercepte uma mensagem criptografada, ele não consegue nenhuma informação sobre a mensagem legível original. Segundo SHANNON [Sha49], a afirmação acima vale mesmo que o espião possua recursos computacionais infinitos.

Para provar formalmente que o $(2, 2) - VCS$ satisfaz as condições de segurança perfeita são definidos:

- M = $\{x_1, \ldots, x_n\}$ o conjunto finito de n textos legíveis
- C = $\{y_1, \ldots, y_u\}$ o conjunto finito de u textos ilegíveis
- K = $\{k_1, \ldots, k_r\}$ o conjunto finito de r chaves. Cada chave k_s define uma função criptográfica: $f_{ks}(x_i) = y_j$
- $p(x_i)$ = probabilidade calculada a priori de x_i ser a mensagem a ser criptografada
- $p(x_i|y_j)$ = probabilidade calculada a posteriori de x_i ter sido criptografado dado que y_j foi interceptado

Conforme TERADA [Ter00], há segurança perfeita se e somente se:

$$\forall x_i \forall y_j : p(x_i|y_j) = p(x_i)$$

Teorema 2.3.1. $(2, 2) - VCS$ *é um algoritmo de segurança perfeita.*

Demonstração. Conforme WEIR e YAN [WY12], observamos que o conjunto finito de textos legíveis é:

$$M = \{0, 1\}$$

Observamos também que o conjunto finito de textos ilegíveis é:

$$C = \{[1, 1, 0, 0], [0, 0, 1, 1], [1, 0, 0, 1], [0, 1, 1, 0], [1, 0, 1, 0], [0, 1, 0, 1]\}$$

Considerando que o conjunto de todas as imagens binárias preto e branco tem distribuição uniforme, as probabilidades a priori de que o pixel a ser criptografado seja branco (0) ou preto (1):

$$p(0) = p(1) = 1/2$$

Consideremos que um $c_j \in C$ seja interceptado pelo espião, $0 \leq j \leq 5$. Para reconstruir um pixel original branco este c_j deveria ser combinado com a transparência de mesmo padrão. Para reconstruir um pixel original preto deveria ser combinado com a transparência complementar. Porém, como não temos conhecimento da outra transparência, o c_j interceptado pode ter sido originado por um pixel branco ou por um pixel preto, ou seja, as probabilidades a posteriori:

$$p(0|c_j) = p(1|c_j) = 1/2$$

□

2.4 Shamir's Secret Sharing: $(k, n) - TS$

Antes de apresentar o $(k, n) - VCS$ vamos abordar o $(k, n) - TS$ (*threshold scheme*), denominação do Shamir's Secret Sharing (ou SSS), processo de distribuição segura de chaves de SHAMIR [Sha79]. O $(k, n) - VCS$ pode ser considerado uma variação visual do $(k, n) - TS$.

Um exemplo de problema para o qual o $(k,n)-TS$ foi desenhado: imagine um laboratório com 11 cientistas que trabalham em um projeto secreto. Os segredos são guardados em um cofre que é aberto somente se seis ou mais cientistas estão presentes. Se resolvêssemos este problema por meio de chaves e fechaduras teríamos uma solução inviável na prática: seriam necessárias 462 fechaduras e cada cientista precisaria carregar 252 chaves!

A solução encontrada por SHAMIR [Sha79] estabelece que uma informação secreta D substituirá chaves e fechaduras. A informação D é dividida em n pedaços:

$$D = \{D_1, D_2, \ldots, D_n\}$$

Sendo que:

- O conhecimento de k ou mais pedaços D_i torna D facilmente calculado
- O conhecimento de quaisquer $k-1$ ou menos pedaços D_i torna computacionalmente difícil calcular D

A solução do $(k,n)-TS$ é baseada em interpolação polinomial: dados k pontos no plano bidimensional $(x_1, y_1), \ldots, (x_k, y_k)$ com distintos x_i's, há um e somente um polinômio $f(x)$ de grau $k-1$ tal que $f(x_i) = y_i$ para todo i. Para construir tal polinômio definimos o segredo $s \in$ Grupo de Galois F de tamanho p onde:

- p é primo
- $0 < k \leq n < p$
- $s < p$

Em seguida escolhemos aleatoriamente inteiros positivos $a_1, \ldots, a_{k-1}$ com $a_i < p$ e adotamos $a_0 = s$ (segredo). Construímos $f(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{(k-1)} \pmod p$ e distribuímos as chaves: o participante i recebe o par $(i, f(i))$ sendo $1 \leq i \leq n$.

A apresentação de k pares permite a recuperação computacionalmente simples do segredo s. Embora o $(k,n)-TS$ ofereça um esquema robusto para distribuir chaves, ele não impede, por exemplo, a união de $k-1$ traidores.

2.5 Modelo Geral: $(k,n) - VCS$

O $(k,n)-VCS$ pode ser entendido como uma variação visual do $(k,n)-TS$. No lugar da informação secreta D do $(k,n)-TS$ passaremos a usar a imagem secreta X no $(k,n)-VCS$. A imagem X será dividida em n pedaços, ou melhor, transformada em n transparências. A presença de k transparências revela o segredo.

De forma análoga ao $(2,2) - VCS$, o $(k,n) - VCS$ foi definido por NAOR e SHAMIR [NS95] como:

- A imagem X é binária, isto é, formada por pixels brancos (codificados como 0) e pretos (codificados como 1). O pixel branco também pode ser interpretado como um pixel transparente. Há variações do $(k, n) - VCS$ específicas para imagens em tons de cinza ou coloridas

- O processo de geração das n transparências é realizado pixel a pixel a partir de X. Cada pixel será codificado de forma independente dos demais. Desta forma pode-se entender a mensagem a ser protegida (a imagem X de p pixels) como um conjunto de p mensagens (de 1 pixel cada)

- Observadas de forma isolada, todas as n transparências geradas pelo $(k, n) - VCS$ se assemelham a ruído aleatório e não revelam qualquer informação sobre a imagem secreta X

- A visão humana é suficiente para revelar a imagem secreta X a partir da sobreposição de quaisquer k transparências. As transparências precisam estar perfeitamente alinhadas para que a sobreposição revele o segredo. Não é necessária qualquer computação nesta etapa

- Quaisquer $k - 1$ transparências sobrepostas não revelam qualquer pista visual sobre a imagem secreta X

No $(k, n) - VCS$ são definidos também:

- No processo de codificação $(k, n) - VCS$:
  - Cada pixel pode ser entendido como uma coleção de n transparências com m sub-pixels cada. Esta coleção será representada como uma matriz booleana $S = [s_{i,j}]$ (a Figura 2.9 apresenta um exemplo da matriz booleana $S = [s_{i,j}]$):
    * $s_{i,j} = 1$ se o j-ésimo sub-pixel na i-ésima transparência é preto
    * $s_{i,j} = 0$ se o j-ésimo sub-pixel na i-ésima transparência é branco

- No processo de decodificação $(k, n) - VCS$:
  - Consideremos o vetor $V = [v_j]$ obtido pela operação OR[2] entre todas as linhas de $S = [s_{i,j}]$ da seguinte forma: $\forall j, 1 \leq j \leq m, v_j = s_{1,j} \vee s_{2,j} \vee \ldots \vee s_{n,j}$
  - Consideremos $H(V)$ a distância de Hamming (OLIVEIRA e NUSSENZEVEIG [ON99]) do vetor V. O resultado será interpretado como:
    * Pixel preto se $H(V) \geq d$ para um patamar $1 \leq d \leq m$
    * Pixel branco se $H(V) \leq d - \alpha m$ para uma diferença relativa $\alpha > 0$

[2]OR (ou) é uma operação lógica entre duas variáveis em que o resultado é 1 se e somente se pelo menos uma das entradas é 1

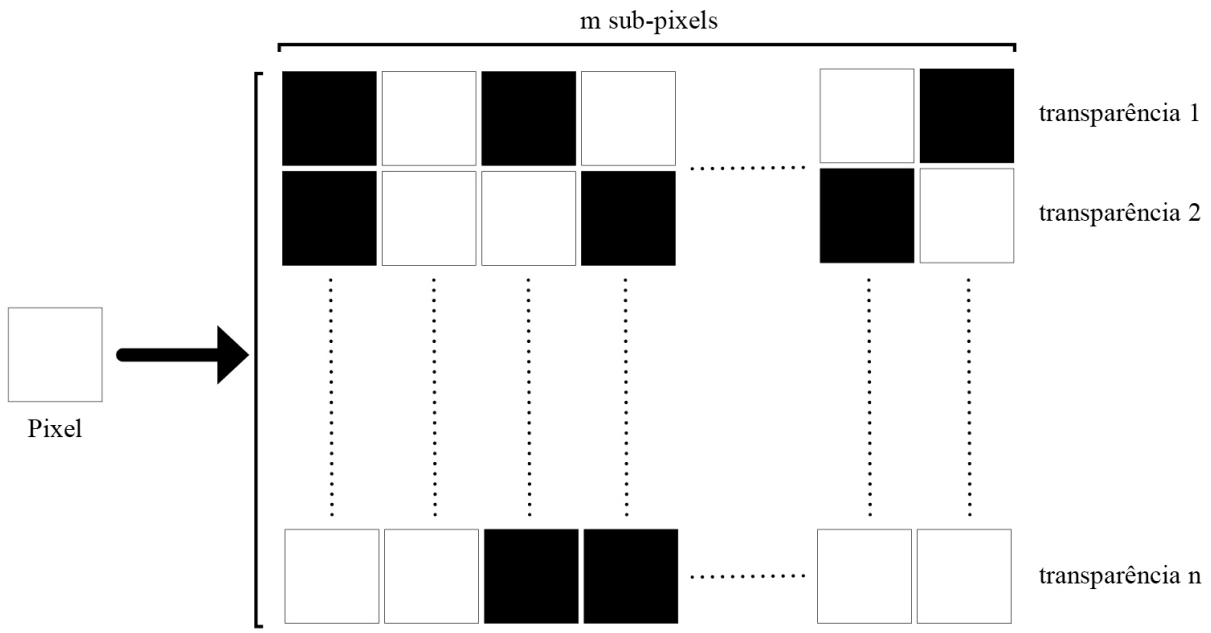

Figura 2.9: *Exemplo da matriz booleana* $S = [s_{i,j}]$*, onde* $1 \leq i \leq n$ *e* $1 \leq j \leq m$

A Figura 2.10 apresenta um exemplo do vetor V, obtido a partir da Figura 2.9, supondo que os sub-pixels não exibidos (entra a transparência 3 e a transparência $n-1$) são brancos. Neste exemplo $H(V) = m - 5$.

Figura 2.10: *Exemplo de vetor* V *com* $H(V) = m - 5$

A diferenciação entre pixel branco e preto é realizada naturalmente pelo cérebro. Um tom cinza mais claro será interpretado como branco e um cinza mais escuro como preto. Em ambiente computacional um filtro de contraste realiza a transformação para branco ou preto. Porém, para n ou k grandes, a perda de contraste pode dificultar a interpretação do cérebro humano.

Para completar a definição do $(k, n) - VCS$ precisamos mostrar o algoritmo de geração das n transparências. No caso do $(2, 2) - VCS$ escolhíamos aleatoriamente uma entre as seis combinações possíveis de matrizes 2×2 de 0 e 1.

No $(k, n) - VCS$ definimos 2 matrizes booleanas $n \times m$ C_0 e C_1:

- Para criptografar um pixel branco: escolha aleatoriamente uma matriz em C_0
- Para criptografar um pixel preto: escolha aleatoriamente uma matriz em C_1

As matrizes C_0 e C_1 são construídas segundo critérios pré-estabelecidos. A solução acima é válida se e somente se 3 condições são satisfeitas:

- Para qualquer $S \in C_0$, a operação $V = OR$ entre quaisquer k linhas de S satisfazem $H(V) \leq d - \alpha m$ (branco)
- Para qualquer $S \in C_1$, a operação $V = OR$ entre quaisquer k linhas de S satisfazem $H(V) \geq d$ (preto)
- Para qualquer subconjunto $\{i_1, i_2, \ldots, i_q\}$ de $\{1, 2, \ldots, n\}$ com $q < k$, as duas coleções de $q \times n$ matrizes D_t para $t \in \{0, 1\}$ obtidas pela limitação de cada matriz $n \times m$ em C_t (onde $t = 0, 1$) às linhas $i_1, i_2, \ldots, i_q$ são indistinguíveis no sentido de que elas contêm as mesmas matrizes com as mesmas frequências.

Segundo NAOR e SHAMIR [NS95], a terceira condição impõe que, ao se inspecionar $q < k$ transparências, o atacante não teria qualquer pista se o pixel original era branco ou preto. As duas primeiras condições definem o contraste e a terceira estabelece a segurança. Veremos na Seção 2.6 uma vulnerabilidade do $(k, n) - VCS$.

A sobreposição de transparências do $(k, n) - VCS$ nos permite analisar 2 parâmetros importantes:

- Número de pixels m em cada transparência:
  - Representa perda em resolução
  - É desejável m menor possível
- Distância relativa α:
  - Representa perda em contraste
  - É desejável α maior possível. A busca do contraste ótimo em criptografia visual é o tema abordado em BLUNDO, D'ARCO, DE SANTIS e STINSON [BDDSS03]

No artigo original de NAOR e SHAMIR [NS95] é demonstrado que:

- Há uma solução ótima para $(k, k) - VCS$:
  - $m = 2^{k-1}$
  - $\alpha = 1/2^{k-1}$
  - Como exemplo: $k = 4$, $m = 8$, $\alpha = 1/8$, $d = 7$, preto ≥ 7 e branco ≤ 6
- Há uma solução (não necessariamente ótima) para $(k, n) - VCS$:
  - $m = \log n \cdot 2^{O(k \log k)}$
  - $\alpha = 1/2^{O(k)}$

No exemplo de decodificação $(2,2)-VCS$ observado na Figura 2.7, onde $k=2$ e $n=2$, usamos $d=4$ e $\alpha=1/2$ para separar pretos e brancos e recuperar a imagem original. Há uma solução alternativa com $d=2$ que também apresenta segurança perfeita. Porém, embora o uso de matriz 2×1 $(d=2)$ utilize menos espaço de memória em relação à 2×2, o efeito colateral é a duplicação do número de colunas (sem afetar o número de linhas) gerando uma imagem distorcida.

Para melhor entendimento do $(k,n)-VCS$ vamos estudar 2 casos particulares nas próximas seções: $(2,n)-VCS$ e $(3,3)-VCS$.

2.5.1 Caso particular $(2,n)-VCS$

A solução do $(2,n)-VCS$ pode ser obtida utilizando-se os seguintes conjuntos de matrizes $n\times n$:

- C_0 = conjunto de todas as matrizes $n\times n$ obtidas pela permutação das colunas de

$$\begin{pmatrix} 100\cdots 0 \\ 100\cdots 0 \\ \cdots \\ 100\cdots 0 \end{pmatrix}$$

- C_1 = conjunto de todas as matrizes $n\times n$ obtidas pela permutação das colunas de

$$\begin{pmatrix} 100\cdots 0 \\ 010\cdots 0 \\ \cdots \\ 000\cdots 1 \end{pmatrix}$$

A análise individual de qualquer transparência pertencente a C_0 ou C_1 não revela se o pixel original era branco ou preto, pois todas as transparências de C_0 e C_1 possuem um 1 e $n-1$ 0.

A sobreposição de quaisquer 2 transparências de uma matriz escolhida de C_0 tem distância de Hamming igual a 1 (cinza claro) que indica que o pixel original era branco.

A sobreposição de 2 transparências de qualquer matriz escolhida de C_1 tem distância de Hamming igual a 2 (cinza mais escuro) que indica que o pixel original era preto.

2.5.2 Caso particular $(3,3)-VCS$

A solução do $(3,3)-VCS$ pode ser obtida utilizando-se os seguintes conjuntos de matrizes 4×3:

- C_0 = conjunto de todas as matrizes 4×3 obtidas pela permutação das colunas de

$$\begin{pmatrix} 0011 \\ 0101 \\ 0110 \end{pmatrix}$$

- C_1 = conjunto de todas as matrizes 4×3 obtidas pela permutação das colunas de

$$\begin{pmatrix} 1100 \\ 1010 \\ 1001 \end{pmatrix}$$

A análise individual de qualquer transparência pertencente a C_0 ou C_1 não revela se o pixel original era branco ou preto, pois todas as transparências de C_0 e C_1 possuem dois 1 e dois 0.

A sobreposição de quaisquer 2 transparências de uma matriz escolhida de C_0 (ou de C_1) tem distância de Hamming igual a 3. Logo, com uma ou duas transparências é impossível distinguir o pixel original.

A sobreposição de 3 transparências de C_0 tem 3/4 pretos e de C_1 tem 4/4 pretos. Assim o segredo é revelado.

2.6 Vulnerabilidade em Criptografia Visual

Vimos na Seção 2.3.1 que o $(2,2) - VCS$ apresenta segurança perfeita, ou seja, mesmo que o espião tenha conhecimento de uma das duas transparências ele não consegue qualquer pista sobre o segredo. No entanto, como veremos nesta Seção, o $(2,n) - VCS$ para $n > 2$ pode ser fraudado.

No caso do $(2,n) - VCS$ para $n > 2$, conforme demonstrado por HORNG, CHEN e TSAI [HCT06], os participantes $P_1, \cdots, P_{n-1}$ podem conspirar e produzir transparências falsas $T'_1, \cdots, T'_{n-1}$. O participante P_n será enganado ao sobrepor sua transparência verdadeira T_n a qualquer uma das transparências falsas $T'_1, \cdots, T'_{n-1}$.

A Figura 2.11 apresenta um fluxo para fraudar o $(2,3) - VCS$. A partir da imagem secreta X são geradas 3 transparências: A, B e C. As sobreposições A com B, A com C e B com C recuperam o segredo X. No entanto, em caso de conluio entre A e B, é possível produzir transparências falsas A' e B' de tal forma que C verá a imagem falsa Y ao sobrepor sua transparência à A' ou B'.

Conforme descrito por DOMINGUES e XEXÉO [DX13], para mostrar que esta fraude é possível, vamos considerar a solução do $(2,3) - VCS$ obtida pelos conjuntos de matrizes 3×3:

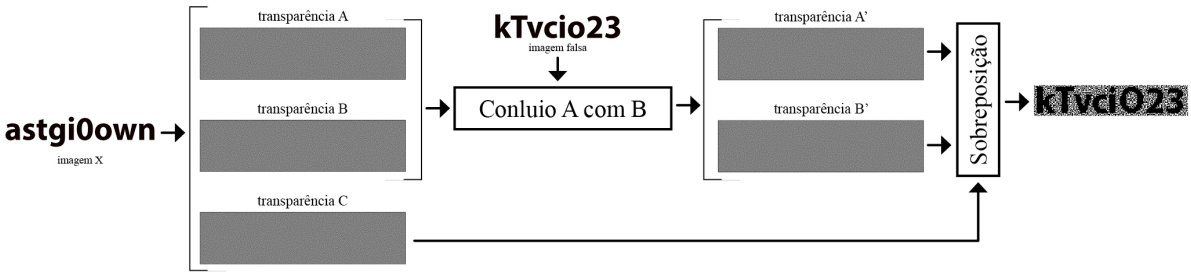

Figura 2.11: *Fraude contra o* $(2,3) - VCS$

- C_0 = conjunto de todas as matrizes 3×3 obtidas pela permutação das colunas de

$$\begin{pmatrix} 100 \\ 100 \\ 100 \end{pmatrix}$$

- C_1 = conjunto de todas as matrizes 3×3 obtidas pela permutação das colunas de

$$\begin{pmatrix} 100 \\ 010 \\ 001 \end{pmatrix}$$

A sobreposição de quaisquer 2 transparências de uma matriz escolhida de C_0 tem distância de Hamming igual a 1 (cinza claro) que indica que o pixel original era branco. A sobreposição de 2 transparências de qualquer matriz escolhida de C_1 tem distância de Hamming igual a 2 (cinza mais escuro) que indica que o pixel original era preto. A Tabela 2.1 ilustra, para cada pixel em X, todas as combinações em A, B e C e os resultados das sobreposições $A+B$, $A+C$ e $B+C$.

Caso	Pixel em X	A	B	C	$A+B$	$A+C$	$B+C$
1	branco	100	100	100	100	100	100
	branco	010	010	010	010	010	010
	branco	001	001	001	001	001	001
2	preto	100	010	001	110	101	011
	preto	100	001	010	101	110	011
	preto	010	100	001	110	011	101
	preto	001	100	010	101	011	110
	preto	010	001	100	011	110	101
	preto	001	010	100	011	101	110

Tabela 2.1: *Para cada pixel em* X*, todas as combinações em* A*,* B *e* C *e os resultados das sobreposições* $A+B$*,* $A+C$ *e* $B+C$

Conhecendo as informações de suas respectivas transparências, a ideia da fraude a ser realizada por A e B é criar novas transparências A' e B' que enganem C ao sobrepor $A'+C$

ou $B' + C$. Tais A' e B' são construídas conforme as Tabelas 2.2 (para pixel branco) e 2.3 (para pixel preto). Em ambas as tabelas, como os pixels de A e B são conhecidos pelos respectivos participantes, é fácil descobrir o pixel em C e criar o pixel falso (branco ou preto, conforme a imagem falsa Y desejada por A e B) nas novas transparências A' e B'. Com esta metodologia C seria induzido a pensar que a imagem falsa Y é a imagem secreta autêntica.

Caso	Pixel em X	A	B	C	Pixel Falso	A'	B'
1	branco	100	100	100	branco	100	100
	branco	010	010	010	branco	010	010
	branco	001	001	001	branco	001	001
2	branco	100	100	100	preto	010	001
	branco	100	100	100	preto	001	010
	branco	010	010	010	preto	100	001
	branco	010	010	010	preto	001	100
	branco	010	001	100	preto	100	010
	branco	001	001	001	preto	010	100

Tabela 2.2: *Construção de A' e B' a partir de um pixel branco*

Caso	Pixel em X	A	B	C	Pixel Falso	A'	B'
1	preto	100	010	001	branco	001	001
	preto	010	100	001	branco	001	001
	preto	100	001	010	branco	010	010
	preto	001	100	010	branco	010	010
	preto	010	001	100	branco	100	100
	preto	001	010	100	branco	100	100
2	preto	100	010	001	preto	100	010
	preto	010	100	001	preto	010	100
	preto	100	001	010	preto	100	001
	preto	001	100	010	preto	001	100
	preto	010	001	100	preto	010	001
	preto	001	010	100	preto	001	010

Tabela 2.3: *Construção de A' e B' a partir de um pixel preto*

Demonstramos que o $(2,3) - VCS$ é vulnerável à fraude. Logo o $(2,n) - VCS$ para $n > 3$ também é vulnerável pois A e B poderiam enganar qualquer um dos demais $n-2$ participantes utilizando a mesma metodologia.

É possível evitar esta fraude se C desconfiar que as transparências A' ou B' não são autênticas. Em Criptografia Visual o conceito de autenticação foi criado por NAOR e PINKAS [NP97]. Foi apresentada por HORNG, CHEN e TSAI [HCT06] uma solução à vulnerabilidade do $(2,3) - VCS$ baseada em autenticação. Na sequencia desta dissertação utilizaremos exclusivamente o $(2,2) - VCS$, que apresenta segurança perfeita. Portanto não há necessidade de detalhar a autenticação em Criptografia Visual nesta dissertação.

2.7 Conclusões do capítulo

A Criptografia Visual é um sistema criptográfico de simples implementação e que não exige computadores na fase de decriptografia. O modelo básico $(2,2)-VCS$ apresenta segurança perfeita. Seus benefícios são robustos se comparados a diversos sistemas criptográficos existentes. Porém o modelo geral $(k,n)-VCS$, derivação visual do Shamir's Secret Sharing (ou SSS), apresenta vulnerabilidade quando $k = 2$ e $n > 2$.

Se por um lado a criação de NAOR e SHAMIR [NS95] é relativamente recente, ela tem despertado interesse do mundo acadêmico e é possível encontrar uma boa quantidade de estudos com melhorias e variações em Criptografia Visual.

Curiosamente, embora a Criptografia Visual apresente benefícios tangíveis e atraia interesse científico, é difícil encontrar aplicações práticas e úteis baseadas na tecnologia. Consequentemente as aplicações comerciais são ainda mais raras, provavelmente inexistentes.

Uma das motivações deste estudo é encontrar uma utilidade prática da Criptografia Visual.

Capítulo 3

Marca d'Água

3.1 Objetivo do capítulo

O objetivo deste capítulo é apresentar o sistema criptográfico denominado Marca d'Água, onde um conteúdo multimídia é alterado, preferencialmente de modo imperceptível, para carregar informações relevantes sobre sua autoria. Estas informações, em caso de disputa, devem revelar o verdadeiro autor.

Este capítulo está estruturado da seguinte forma: a Seção 3.2 apresenta uma breve introdução à Marca d'Água e sua contextualização em três momentos distintos: em ambiente protegido, em ambiente desprotegido (na Internet, por exemplo) e em caso de disputa (em um tribunal). Em 3.2.1 encontramos as classificações de um sistema de Marca d'Água conforme o objetivo desejado, tipo de conteúdo multimídia a ser protegido, visibilidade, resistência, domínio de desenvolvimento e dependência da imagem original. Em 3.2.2 veremos a classificação da proposta de Marca d'Água que é objeto desta dissertação. Em 3.2.3 introduzimos uma breve reflexão sobre o *valor comercial* de uma imagem estática, conceito que será útil na discussão sobre ataques à Marca d'Água. A Seção 3.3 apresenta uma lista ampliada com as propriedades desejadas de uma solução de Marca d'Água aplicada à proteção dos direitos autorais de imagem. A Seção 3.4 apresenta a solução de Marca d'Água proposta por COX, KILIAN, LEIGHTON e SHAMOON [CKLS97], segmentada em camadas *conteúdo* (em 3.4.1) e *transporte* (em 3.4.2). A Seção 3.5 apresenta os ataques que a Marca d'Água pode sofrer e analisa o impacto destes ataques sobre o *valor comercial* da imagem protegida. A Seção 3.6 apresenta nosso ponto de vista sobre Marca d'Água e estabelece a base da nova solução que será proposta no próximo capítulo.

3.2 Introdução à Marca d'Água

A proteção dos direitos autorais de conteúdos multimídia (imagens estáticas, vídeos, áudios, etc) é um dos desafios atuais da Criptografia. No contexto em que vivemos, de rápida propagação das redes sociais, ambientes compartilhados e aplicativos celulares, temos o uso

massivo da Internet por mais de 3 bilhões de pessoas no mundo [Exa15]. Muitas das quais estão, neste exato momento, mesmo que inconscientemente, copiando e alterando músicas, fotografias e vídeos sem o devido respeito aos direitos autorais. Para agravar o problema, mais de 4 bilhões de pessoas residentes nos países mais pobres do planeta seguem sem estar conectadas.

Em Criptografia, técnicas são geralmente utilizadas para *esconder* uma informação, *proteger* sua integridade e *identificar* o respectivo autor. No entanto, na Internet, estas mesmas técnicas são inócuas para evitar que a informação seja *copiada*, *adulterada* ou *plagiada* sem autorização do verdadeiro autor. Métodos específicos são necessários para combater estes problemas, não tratados pela Criptografia tradicional. Entre estes métodos encontramos a tecnologia de Marca d'Água.

O uso de Marca d'Água como mecanismo de proteção dos direitos autorais de conteúdos multimídia pode ser sintetizado da seguinte forma:

- Um dado relevante M_{Autor} (Marca d'Água), conhecido apenas pelo autor A, é inserido na obra multimídia original $I_{Original}$ de forma imperceptível. A informação resultante é denominada $I_{Marcada}$

- Este dado relevante M_{Autor} deve ser capaz de preservar sua integridade, sendo resistente às manipulações que possam ser realizadas na informação $I_{Marcada}$. A informação resultante da manipulação de $I_{Marcada}$ é denominada $I_{Litigio}$

- Ao ser extraído de $I_{Marcada}$ ou de $I_{Litigio}$, M_{Autor} apresenta informações relevantes sobre o autor A e sua obra original $I_{Original}$

Deve ficar claro que o uso de Marca d'Água não impede que o conteúdo multimídia seja *copiado*, *adulterado* ou *plagiado* sem autorização do verdadeiro autor. O real benefício da Marca d'Água é que, em caso de *litígio*, ela pode ser utilizada como prova da verdadeira autoria.

A Figura 3.1 apresenta o fluxo de uso de Marca d'Água em 3 momentos:

- Em ambiente protegido, o autor A incorpora a Marca d'Água M_{Autor} na obra multimídia original $I_{Original}$. A obra $I_{Marcada}$ obtida neste processo será divulgada pelo autor

- Em ambiente desprotegido (na Internet, por exemplo), qualquer Usuário pode copiar, adulterar ou plagiar a obra divulgada $I_{Marcada}$

- Em caso de litígio (no tribunal, por exemplo), o verdadeiro autor pode defender seus direitos a partir da extração da Marca d'Água de $I_{Litigio}$ e provar sua autoria

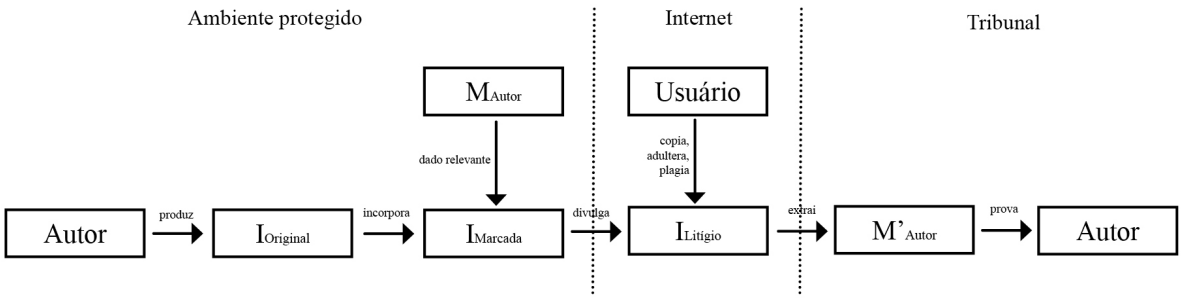

Figura 3.1: *Fluxo de uso de Marca d'Água em 3 momentos: em ambiente protegido, em ambiente desprotegido (neste fluxo, na Internet) e em caso de litígio*

3.2.1 Classificação de Marcas d'Água

Os sistemas baseados em Marca d'Água podem ser classificados de diversas formas conforme observado em LOPES [Lop06] e revisado sob nosso ponto de vista:

1. Em relação ao objetivo da Marca d'Água:
   - Proteção dos direitos autorais da imagem
   - Integridade da imagem. É observado em MACQ e QUISQUATER [BQ95] que os objetivos de proteção dos direitos autorais e de garantia de integridade não podem ser simultaneamente atingidos. Não encontramos na literatura uma solução que aborde os dois objetivos simultaneamente
2. Em relação ao conteúdo multimídia protegido:
   - Imagem estática (fotografia, gravura, etc)
   - Vídeo (imagens em movimento)
   - Áudio
3. Em relação à visibilidade da Marca d'Água:
   - Visível: quando a Marca d'Água é visualmente percebida no conteúdo protegido, como, por exemplo, o logotipo de uma emissora de TV em sua programação. Embora de simples implementação, a principal desvantagem desta técnica é que o atacante sabe exatamente onde deve agir para tentar eliminar a Marca d'Água. Um exemplo de Marca d'Água visível é mostrado na Figura 3.2
   - Invisível: quando a Marca d'Água não é visualmente percebida pelo olhar humano. Isso é possível pois a visão humana não consegue perceber a diferença entre o conteúdo original e o conteúdo marcado se bits do conteúdo multimídia a ser protegido são ligeiramente alterados para incorporar a Marca d'Água. O exemplo de Marca d'Água invisível mostrado na Figura 3.3 apresenta lado a lado a imagem original e uma imagem marcada com 33,5% dos bits sutilmente alterados sem aparentes diferenças

Figura 3.2: *Exemplo de Marca d'Água visível*

Figura 3.3: *Exemplo de Marca d'Água invisível: a imagem original e sua respectiva imagem marcada com 33,5% dos bits sutilmente alterados sem aparentes diferenças*

4. Em relação à resistência da Marca d'Água encapsulada. BARRETO [Bar03] define esta classificação como a dificuldade em remover a Marca d'Água de um conteúdo marcado. Para efeito desta dissertação adotaremos a definição proposta por LOPES [Lop06], que usa a capacidade da Marca d'Água encapsulada em resistir às operações de processamento de imagem. Desta forma uma Marca d'Água pode ser:

    - Robusta: para proteção dos direitos autorais é desejável que a Marca d'Água encapsulada no conteúdo protegido permaneça relevante mesmo que este conteúdo sofra alterações como compactação, corte, rotação, filtros, etc. Uma Marca d'Água com esta característica é dita robusta
    - Frágil: para garantia de integridade, é desejável que qualquer alteração do conteúdo protegido corrompa a Marca d'Água encapsulada. Uma Marca d'Água deste tipo é dita frágil

5. Em relação ao domínio utilizado em seu desenvolvimento (conforme descrito em WEIR e YAN [WY12]):

- Espaço: neste domínio o encapsulamento da Marca d'Água no conteúdo a ser protegido é realizado pela alteração de alguma(s) de sua(s)característica(s): brilho, saturação, contraste, etc. Para que esta alteração não seja percebida pela visão humana são utilizados os bits menos significativos. Porém, os bits menos significativos são os mais afetados em operações simples e corriqueiras como a compactação de conteúdo. Esta é a principal desvantagem do domínio espacial quando o objetivo é obter uma Marca d'Água robusta
- Frequência: neste domínio o encapsulamento da Marca d'Água no conteúdo a ser protegido é realizado pela alteração dos coeficientes gerados nas transformações em frequência. As principais transformadas utilizadas em Marca d'Água são:
  - DCT (Discrete Cosine Transform)
  - DWT (Discrete Wavelet Transform)

6. Em relação ao processo de extração da Marca d'Água (segundo KUTTER e PETITCOLAS [KP99]):
   - Extração cega: quando o conteúdo original *não é* necessário para a extração da Marca d'Água do conteúdo protegido
   - Extração não cega: quando o conteúdo original *é* necessário para a extração da Marca d'Água do conteúdo protegido

3.2.2 Definições

Nesta dissertação temos como foco principal a utilização de Marca d'Água com as seguintes características: proteção dos direitos autorais, aplicada a imagens estáticas coloridas ou monocromáticas (fotografias, gravuras, etc), invisível ao olhar humano, robusta em relação às operações de processamento de imagem, desenvolvida no domínio da frequência pela transformada DCT e com extração não cega.

A partir deste momentos definimos:

- $I_{Original}$ = imagem original, a ser protegida
- A = autor verdadeiro da imagem $I_{Original}$
- M_{Autor} = imagem Marca d'Água contendo informação relevante sobre o autor verdadeiro A e sua obra $I_{Original}$
- $I_{Marcada}$ = imagem $I_{Original}$ marcada com M_{Autor} (ou, de forma mais completa, imagem resultante do encapsulamento da Marca d'Água M_{Autor} em $I_{Original}$)
- $I_{Litigio}$ = imagem com autoria incerta
- A_{falso} = autor falso, que pleiteia a autoria da imagem $I_{Litigio}$

3.2.3 Valor comercial de uma imagem

Na literatura encontramos uma grande quantidade de trabalhos com foco nas questões tecnológicas da Marca d'Água. Por outro lado, uma quantidade muito menor de trabalhos dedica-se ao estudo de sua aplicabilidade prática.

Neste ponto é importante fazer uma breve reflexão sobre o *valor comercial* de uma imagem. De forma genérica, podemos afirmar que uma imagem tem *valor comercial* se há intenção de uma pessoa P de remunerar o autor A por uma quantidade de dinheiro maior do que zero pelo uso/ cópia/ alteração da imagem original $I_{Original}$. Vamos citar situações em que uma imagem estática tem *valor comercial*:

- Leilão de imagem impressa (em [For12] encontramos uma fotografia leiloada em 2011 por US$ 4,3 milhões pela conceituada casa de leilões americana Christie's)
- Serviço de download de ilustração digital (planta de arquitetura, por exemplo)
- Publicação de desenho em jornal ou revista
- Exibição de fotografia numa peça publicitária
- Criação de uma logomarca

Neste contexto, encontramos em KATZENBEISSER e PETITCOLAS [KP00] e IGUCHI [Igu07] a definição de que uma técnica de Marca d'Água é considerada robusta se apresenta duas características:

- A eliminação da Marca d'Água M_{Autor} da imagem marcada $I_{Marcada}$ somente é possível se a imagem resultante degrada a ponto de não possuir *valor comercial*
- A eliminação da Marca d'Água M_{Autor} da imagem marcada $I_{Marcada}$ sem degradação da imagem resultante é *computacionalmente cara*[1]

Em ambas as situações acima temos um valor financeiro envolvido, característica relevante na aplicação prática de qualquer sistema. A primeira característica indica que não compensa para o verdadeiro autor disputar judicialmente por algo que não tem *valor comercial*; a segunda característica inibe a ação de falsos autores. Logo, pela definição de KATZENBEISSER e PETITCOLAS [KP00], basta ao autor verdadeiro aplicar um método robusto de Marca d'Água em suas imagens para que não haja nenhuma necessidade de disputa de autoria.

Figura 3.4 apresenta lado a lado uma imagem marcada $I_{Marcada}$ com resolução 1.000×1.000 pixels e uma imagem manipulada $I_{Litigio}$ com resolução 100×100 pixels. A Marca d'Água presente em $I_{Marcada}$ foi eliminada em $I_{Litigio}$ porém resultando em uma imagem sem qualquer *valor comercial*.

[1] computacionalmente cara significa: os algoritmos publicados mais rápidos para encontrar uma solução são todos de complexidade maior que qualquer polinônio.

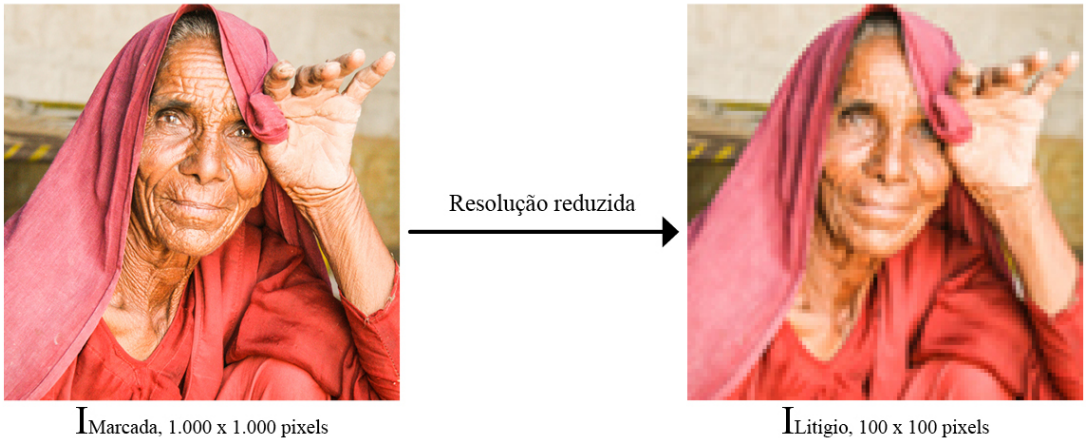

Figura 3.4: *Exemplo de eliminação de Marca d'Água com evidente perda de* valor comercial

Porém, é fácil notar que um método robusto de Marca d'Água não resolve por si só todos os problemas relacionados à proteção dos direitos autorais de imagens estáticas. Para minimizar riscos o autor verdadeiro A da imagem $I_{Original}$ pode adotar algumas práticas complementares:

1. Armazenar em ambiente protegido o acervo de imagens originais em alta resolução. Caso contrário, um falso autor poderia copiar os arquivos e aplicar sua própria Marca d'Água às imagens originais. Há inúmeras aplicações comerciais para armazenamento seguro de imagens de forma local e remota, logo este ponto não será detalhado neste estudo

2. Aplicar uma metodologia robusta de encapsulamento de Marca d'Água às suas imagens originais em alta resolução. Esta metodologia deve prover ferramentas para que o verdadeiro autor possa defender seus interesses em um eventual processo judicial de disputa de autoria. Uma metodologia com este objetivo será detalhada no próximo capítulo desta dissertação

3. Gerar cópias para divulgação em baixa resolução de suas imagens. A partir de suas imagens originais encapsuladas com a Marca d'Água o autor deve gerar as respectivas cópias em baixa resolução (por exemplo, máximo de 200 pixels na horizontal e na vertical, em 72 dpi[2], dimensões que possuem baixo ou nenhum *valor comercial*)

4. Acrescentar uma Marca d'Água visível em suas imagens para divulgação. Uma identificação visível do autor na imagem (nome, site, logotipo, etc) irá facilitar sua busca por usuários legítimos realmente interessados na aquisição da imagem. É preciso ter algum critério de qualidade para evitar que esta marca visível não comprometa o *valor comercial* da imagem

[2] *Dots per inch*, ou pontos por polegada, indica a densidade de agrupamento de pontos em uma imagem. Para exibição web bastam 72 dpi; para impressão recomenda-se 300 dpi

5. No caso de comercialização de uma imagem, assinar contrato de uso com a pessoa adquirente. Tal contrato estabelecerá as condições nas quais a imagem pode ser utilizada (quantidade de cópias, abrangência geográfica, período de divulgação, mídias, etc) e aquelas nas quais ela não pode ser usada (cópias não autorizadas, etc). Este contrato trará segurança jurídica ao autor em caso de má utilização de sua imagem pelo adquirente

6. A imagem fornecida ao adquirente será a $I_{Marcada}$ na resolução contratada. A $I_{Original}$ nunca é revelada pelo autor verdadeiro, exceto na condição de disputa em um tribunal

3.3 Propriedades ideais da Marca d'Água

Em BARRETO [Bar03] encontramos uma relação das propriedades ideais de Marca d'Água no contexto de localização de alterações maliciosas em uma imagem marcada. Estas mesmas propriedades podem ser avaliadas e adaptadas para o nosso foco em proteção dos direitos autorais de uma imagem. Estas propriedades são:

- Armazenada na própria imagem. Como cada imagem marcada carrega sua informação de segurança de forma intrínseca, não existe a necessidade de criação de uma entidade independente de certificação (EIC) para registro e verificação das imagens originais. Provavelmente o custo e a complexidade de tal entidade inviabilizariam sua existência

- Visualmente imperceptível quando inserida. Esta característica é importante para garantir que a imagem marcada, mesmo em baixa resolução, não perca seu *valor comercial* por algum ruído indesejável causado pela inserção da Marca d'Água

- Visualmente significativa quando extraída. A informação de segurança extraída da imagem marcada deve apresentar significado relevante para ser útil numa disputa de autoria

- Irreproduzível por terceiros não autorizados. Com esta característica pretende-se evitar que uma imagem maliciosa seja atribuída a um autor com o objetivo de prejudicá-lo. A metodologia deve garantir que somente o autor consegue encapsular sua própria Marca d'Água

- Capaz de localizar alterações maliciosas na imagem hospedeira com resolução suficiente. Conforme observamos na seção 3.2.1, quando falamos sobre as classificações de Marca d'Água, os objetivos de proteção de direitos autorais e de garantia de integridade não podem ser simultaneamente atingidos. Adicionalmente, não há relação entre a localização das alterações em uma imagem e a prova de sua autoria. Logo, esta propriedade não se aplica ao nosso contexto de proteção dos direitos autorais e será desconsiderada

- Publicamente verificável. Para que esta propriedade seja considerada válida em uma disputa de direito autoral é fundamental que a metodologia utilize ferramentas de conhecimento público e que possam ser reproduzidas por ambas as partes litigantes

- Indelével por manipulação não autorizada. Com esta propriedade pretende-se garantir que um falsário não consegue modificar/ excluir (com sucesso) uma Marca d'Água encapsulada em uma imagem

- Resistente a certas operações de processamento de imagem (como mudar o nível de compressão). É corriqueiro o uso de técnicas de processamento de imagem como compactação, rotação, corte e transformação em preto e branco aplicadas às imagens estáticas. Esta propriedade visa garantir que a Marca d'Água não sofrerá perdas significativas pela aplicação destas técnicas de forma razoável. É possível intuir que a manipulação radical de uma imagem marcada, mantendo somente resquícios da imagem original e criando uma imagem completamente nova, provavelmente degradará a Marca d'Água a ponto de torná-la inútil. Uma discussão que não faz parte deste estudo é a divisão precisa entre manipulação razoável e radical. Pode-se adotar como regra a jurisprudência existente para a resolução de casos de *plágio*. A consequência de um caso recente de *plágio* de repercussão internacional foi o cancelamento da logomarca das Olimpíadas de 2020 em Tóquio [G115]

- Aplicável a formatos com e sem perdas, e a imagens binárias, em níveis de cinza e coloridas. Com esta propriedade pretende-se deixar a metodologia independente da codificação de imagens estáticas em ambiente computacional

- Eficiente em tempo de processamento e espaço de armazenamento

Adicionalmente, sugerimos novas propriedades com o objetivo de enriquecer o modelo de Marca d'Água para proteção dos direitos autorais de uma imagem (conforme será detalhado no capítulo seguinte):

- Sem ambiguidade. Com esta nova propriedade pretende-se que a extração da Marca d'Água da imagem em litígio irá minimizar a existência de dúvida em um processo de disputa de autoria

- Imperceptível por segurança perfeita. Esta propriedade deve garantir que a busca da Marca d'Água inserida em uma imagem marcada será *computacionalmente cara*, tornando-a inviável para o atacante

- Alta entropia. A cada imagem original deve corresponder uma única Marca d'Água

- Capacidade de absorção da Marca d'Água. Adaptamos esta propriedade a partir da descrição encontrada em IGUCHI [Igu07] sobre a capacidade de inserção das técnicas de marca d'água (*watermarking payload*). A capacidade de absorção de uma Marca

d'Água é o menor tamanho necessário da imagem original (em pixels) para que a inserção da Marca d'Água minimize a degradação visual da imagem resultante

Conforme BARRETO [Bar03], obter todas estas características em uma única solução é muito difícil na prática. A maioria dos algoritmos de Marca D'água concentra-se apenas em alguns poucos destes requisitos, e portanto são aplicáveis numa gama restrita de circunstâncias. Outra dificuldade reside no fato que os requisitos acima podem ser contraditórios, no sentido de que algumas operações em imagens são ora exigidas, ora proibidas.

3.4 Marca d'Água de Cox

Em WEIR e YAN [WY12] encontramos que os sistemas de Marca d'Água podem ser agrupados conforme seus domínios: espacial ou frequência.

Genericamente, um método de Marca d'Água que utiliza o domínio espacial realiza alterações nos bits que codificam as cores, modificando os aspectos de brilho, saturação e contraste. Neste domínio o pressuposto utilizado é de que o encapsulamento da Marca d'Água não deve afetar a qualidade visual da imagem resultante; para isso utilizam-se métodos LSB (*Least Significant Bit*) como o encontrado em SCHYNDEL, TIRKEL e OSBORNE [VSTO94], considerado por alguns como o primeiro trabalho sobre Marca d'Água na literatura. Porém, a utilização do LSB para *esconder* a Marca d'Água não tem resistência significativa às operações de processamento de imagem. Portanto, métodos do domínio espacial não são adequados ao nosso objetivo de Marca d'Água robusta, necessário para a proteção de direitos autorais.

No caso do domínio de frequência, o princípio utilizado é o encapsulamento da Marca d'Água nas partes visualmente mais significativas da imagem original. Pelo uso de transformadas como a DWT (*Discrete Wavelet Transform*) e a DCT (*Discrete Cosine Transform*) identificamos os coeficientes mais significativos da imagem; estes coeficientes são alterados para encapsular a Marca d'Água e a imagem resultante é obtida pela transformada inversa. Usualmente no domínio de frequência a Marca d'Água é verificada utilizando-se métodos estatísticos, porém nesta dissertação adotaremos uma solução visual conforme veremos no próximo capítulo. Os métodos de Marca d'Água que utilizam o domínio de frequência provaram-se mais resistentes às operações de processamento de imagem do que os do domínio espacial. Encontramos em COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] o uso pioneiro do domínio de frequência no contexto de Marca d'Água.

Apresentado em 1997, o método de COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] continua sendo referência nos estudos sobre Marca d'Água. Os componentes da marca são espalhados por toda a imagem a ser protegida usando técnica similar à da comunicação *spread spectrum*. A imagem original é segmentada em pequenos blocos e a cada um destes blocos é aplicada a transformada DCT. O uso desse recurso minimiza a degradação da qualidade visual da imagem marcada desde que os parâmetros sejam devidamente ajustados.

Para uso desta dissertação vamos segmentar o método de Cox em duas camadas: *conteúdo* e *transporte*. Como *conteúdo* entende-se a Marca d'Água propriamente dita, ou seja, a informação relevante que será usada como ferramenta de proteção dos direitos autorais. Como *transporte* entende-se o método de encapsular a Marca d'Água na imagem a ser protegida tornado-a robusta.

Neste estudo poderíamos ter adotado outros métodos de Marca d'Água robusta. No entanto a solução de Cox nos é conveniente por diversos motivos:

- É um método bastante conhecido
- Foi testado de forma exaustiva
- Permite simples e completa segmentação entre as camadas *conteúdo* e *transporte*
- Possibilita que a camada *conteúdo*, que Cox define como um dado pseudo-aleatório, seja integralmente substituída por uma informação com significado e sem prejuízo dos resultados
- Fornece uma camada de *transporte* bastante resistente às operações de processamento de imagem e outras transformações

3.4.1 Camada *conteúdo* de Cox

Sobre a camada de *conteúdo*, COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] defendem que a Marca d'Água deve ser construída como um vetor randômico gaussiano independente e identicamente distribuído. Muitas variações são possíveis, embora a mais básica consiste em uma sequência de números reais $X = x_1, \cdots, x_n$. Na prática cada valor x_i da Marca d'Água é escolhido de forma independente de acordo com $N(0,1)$, onde $N(\mu, \sigma^2)$ indica uma distribuição com valor esperado μ e variância σ^2. Distribuições alternativas para construção da Marca d'Água incluem a escolha uniforme de x_i em $\{1, -1\}$, $\{0, 1\}$ ou [0, 1].

3.4.2 Camada *transporte* de Cox

A estratégia de COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] para transportar a Marca d'Água e resistir aos possíveis ataques (como veremos na Seção 3.5) tem 2 pilares:

1. Encapsular a Marca d'Água nas partes visualmente mais significativas da imagem. Nesta abordagem, utilizando a transformada DCT, identificamos os coeficientes mais significativos da imagem. Quanto menor for a alteração destes coeficientes pela inserção da Marca d'Água, menor será o impacto visual na imagem resultante
2. Espalhar a Marca d'Água por toda a superfície da imagem. Nesta abordagem a imagem é segmentada em pequenos blocos de 8×8 pixels, cada um destes blocos sendo candidato a receber parte da Marca d'Água. Deste modo evita-se que a Marca d'Água

fique concentrada em uma região pré-determinada da imagem (como no *header*, por exemplo), dificultando a ação de um possível atacante

A imagem $I_{Original}$ a ser marcada é segmentada em blocos de $N \times N$ pixels. Para realizar a transformada DCT em cada um destes blocos, conforme encontramos em CABEEN e GENT [CG98], adotamos $N = 8$ e calculamos a matriz T de transformação onde a i, j-ésima posição é dada pela equação:

$$T_{i,j} = \left\{ \begin{array}{ll} \frac{1}{\sqrt{N}} & \text{se } i = 0 \\ \\ \sqrt{\frac{2}{N}} \cos[\frac{(2j+1)i\pi}{2N}] & \text{se } i > 0 \end{array} \right\} \text{ onde } 0 \leq i \leq 7,\ 0 \leq j \leq 7$$

A matriz 8×8 resultante é

$$T = \begin{bmatrix} 0.3535 & 0.3535 & 0.3535 & 0.3535 & 0.3535 & 0.3535 & 0.3535 & 0.3535 \\ 0.4903 & 0.4157 & 0.2777 & 0.0975 & -0.0975 & -0.2777 & -0.4157 & -0.4903 \\ 0.4619 & 0.1913 & -0.1913 & -0.4619 & -0.4619 & -0.1913 & 0.1913 & 0.4619 \\ 0.4157 & -0.0975 & -0.4903 & -0.2777 & 0.2777 & 0.4903 & 0.0975 & -0.4157 \\ 0.3535 & -0.3535 & -0.3535 & 0.3535 & 0.3535 & -0.3535 & -0.3535 & 0.3535 \\ 0.2777 & -0.4903 & 0.0975 & 0.4157 & -0.4157 & -0.0975 & 0.4903 & -0.2777 \\ 0.1913 & -0.4619 & 0.4619 & -0.1913 & -0.1913 & 0.4619 & -0.4619 & 0.1913 \\ 0.0975 & -0.2777 & 0.4157 & -0.4903 & 0.4903 & -0.4157 & 0.2777 & -0.0975 \end{bmatrix}$$

Vamos tomar como exemplo um bloco M com 8×8 pixels de uma imagem $I_{Original}$ monocromática:

$$M = \begin{bmatrix} 83 & 83 & 88 & 98 & 94 & 113 & 139 & 172 \\ 86 & 88 & 93 & 99 & 113 & 154 & 170 & 181 \\ 84 & 87 & 92 & 100 & 135 & 175 & 184 & 180 \\ 91 & 79 & 98 & 128 & 151 & 169 & 164 & 169 \\ 92 & 81 & 117 & 149 & 149 & 148 & 143 & 165 \\ 104 & 107 & 128 & 157 & 152 & 148 & 156 & 182 \\ 120 & 126 & 139 & 159 & 153 & 156 & 150 & 188 \\ 134 & 146 & 143 & 147 & 142 & 147 & 143 & 177 \end{bmatrix}$$

A transformada DCT do bloco M é obtida pela fórmula $D = TMT'$, onde T' é a matriz

transposta de T. O resultado da DCT no exemplo acima é[3]:

$$D = \begin{bmatrix} 37.0000 & -206.3680 & 9.1980 & -16.3405 & 22.5000 & -0.9269 & 15.6731 & -13.8338 \\ -93.3543 & -63.8359 & 33.9540 & 21.0395 & -21.1099 & 8.6541 & -1.9775 & 5.3151 \\ -6.9155 & 39.1190 & 42.4684 & -24.6166 & -7.7601 & -10.3731 & -3.7626 & -6.4321 \\ -21.9525 & 14.6422 & -10.6960 & -29.7600 & 23.1325 & 7.9486 & -0.7291 & -1.5707 \\ -25.5000 & 22.7440 & -7.1204 & -4.8823 & 12.5000 & 1.5397 & 2.4082 & -0.0566 \\ 2.9318 & -4.9011 & -1.5122 & 3.3525 & 4.5834 & -2.1836 & -0.6942 & -0.7144 \\ -3.6299 & -0.6353 & 6.2374 & -0.6174 & -4.3624 & -1.6041 & -5.9684 & 2.9824 \\ -5.4511 & 0.7189 & -2.1476 & -2.7975 & -0.0791 & -0.9887 & 3.1460 & 0.7796 \end{bmatrix}$$

Uma das propriedades da transformada DCT é que os coeficientes mais significativos (em módulo) tendem a se agrupar na parte superior esquerda da matriz D e representam as áreas visualmente mais significativas deste bloco. Analogamente, os menos significativos (em módulo) tendem a se agrupar na parte inferior direita de D. Algoritmos de compactação como o JPEG utilizam esta propriedade como estratégia para economizar espaço e desprezam os coeficientes menos significativos.

Da transformação DCT aplicada a todos os blocos com 8×8 pixels de $I_{Original}$ extraímos a sequência com os n maiores coeficientes (em módulo) $V = v_1, \cdots, v_n$ na qual será inserida a Marca D'água $X = x_1, \cdots, x_n$ para obter a sequência ajustada $V' = v'_1, \cdots, v'_n$. Em seguida inserimos V' no lugar de V e calculamos a IDCT (DCT inversa[4]) de cada bloco de 8×8 pixels para obter a imagem marcada $I_{Marcada}$. O resultado destas operações é o espalhamento da Marca D'água pela superfície de $I_{Marcada}$.

A Figura 3.5 ilustra o método de inserção de Cox, considerando uma Marca d'Água na distribuição $\{0, 1\}$.

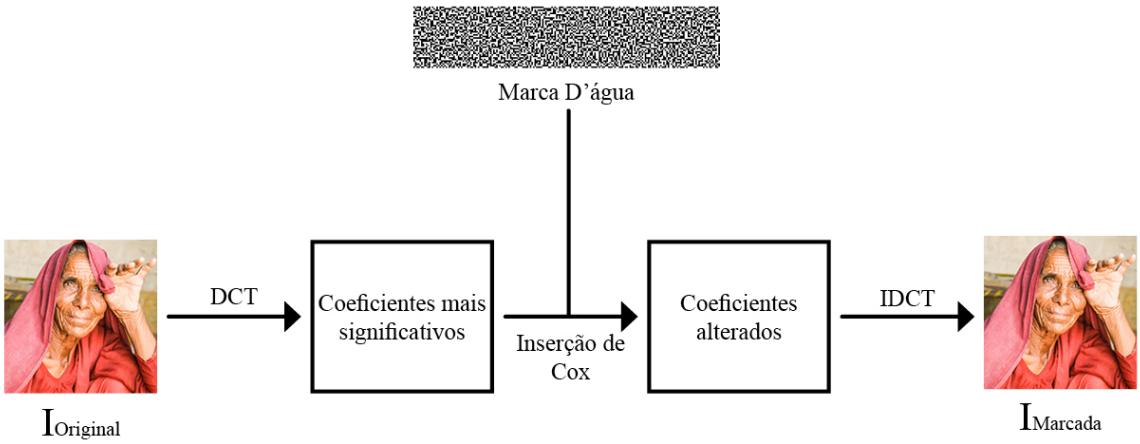

Figura 3.5: *Fluxo de inserção de Cox, considerando uma Marca d'Água na distribuição* $\{0, 1\}$

[3]Um detalhe de implementação: cada elemento da matriz M deve ser subtraído de 128 antes da multiplicação de matrizes $D = TMT'$

[4]A DCT inversa é calculada por $M' = T'DT$. Um detalhe de implementação: cada elemento da matriz M deve ser acrescido de 128 após a multiplicação de matrizes

Um atacante pode alterar $I_{Marcada}$ e produzir uma nova imagem $I_{Litigio}$ onde uma Marca D'água $X^* = x_1^*, \cdots, x_n^*$ (possivelmente $X \neq X^*$) pode ser extraída (conforme veremos em seguida).

Segundo COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] há três fórmulas naturais para calcular V':

$$v_i' = v_i + \alpha x_i \tag{3.1}$$

$$v_i' = v_i(1 + \alpha x_i) \tag{3.2}$$

$$v_i' = v_i(e^{\alpha x_i}). \tag{3.3}$$

O fator de escala α indica o quanto a Marca D'água X afetará os coeficientes V. Um valor pequeno de α torna a Marca D'água mais suscetível a ruídos; um valor grande de α pode afetar visualmente a imagem resultante. Nos experimentos de Cox foram utilizados: $\alpha = 0.1$ e a fórmula $v_i' = v_i(1 + \alpha x_i)$.

Se $v_i > 0$ então $v_i' = v_i(1 + \alpha x_i)$ é inversível, permitindo que a Marca D'água $X = x_1, \cdots, x_n$ seja extraída de $I_{Marcada}$. De forma análoga a Marca D'água $X^* = x_1^*, \cdots, x_n^*$ pode ser extraída de $I_{Litigio}$.

A Figura 3.6 ilustra o método de extração de Cox, obtendo a Marca d'Água X^* na distribuição $\{0, 1\}$.

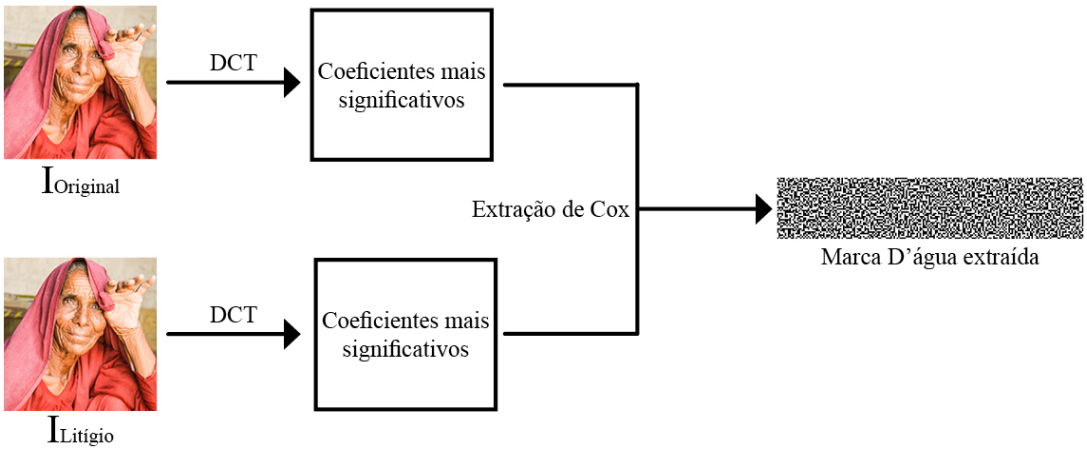

Figura 3.6: *Fluxo de extração de Cox, obtendo uma Marca d'Água na distribuição* $\{0, 1\}$

Cox também define uma medida para avaliar a similaridade entre a Marca d'Água original X e a Marca d'Água extraída X^*:

$$sim(X, X^*) = \frac{X^*.X}{\sqrt{X^*.X^*}}$$

Aceita-se a correspondência entre X e X^* se $sim(X, X^*) > T$, sendo T um limiar pré-definido. Nos testes de Cox utilizou-se $T = 6$. Em nosso método, conforme será apresentado no próximo capítulo, a comparação entre X e X^* será visual e não estatística.

3.5 Ataques à Marca d'Água

Para reforçar as vantagens do uso do domínio de frequência, COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] examinam os diferentes *ataques* que uma imagem pode sofrer em ambiente desprotegido. Usaremos o termo *ataques* para designar as operações que afetam a estrutura da imagem, motivadas pela ação maliciosa de um atacante para remover (ou danificar) a Marca d'Água ou simplesmente uma ação inconsciente (sem intenção maliciosa) do usuário. Para enriquecer as conclusões de Cox, incluiremos também nossa análise sobre o impacto causado pelos *ataques* no *valor comercial* da imagem.

A Figura 3.7 categoriza tais possibilidades de ataque, que detalhamos a seguir:

Figura 3.7: *Fluxo com categorias de ataque (intencional ou não)*

- Compressão com perdas: o objetivo deste tipo de alteração é reduzir o tamanho de armazenamento com possível prejuízo da qualidade da imagem resultante. Quanto menor o tamanho desejado, maior será o prejuízo visual. Nesta categoria encontramos algoritmos como o JPEG[5] que atuam no domínio de frequência e priorizam as perdas nas áreas menos significativas da imagem. Logo, os coeficientes mais significativos da imagem e que contêm a Marca d'Água serão os menos afetados pela compressão. Se a compressão for utilizada maliciosamente para prejudicar a Marca d'Água, compactando a imagem a ponto de afetar os coeficientes mais significativos, a imagem resultante também estará degradada e perderá seu *valor comercial*

- Distorção geométrica: nesta categoria incluímos transformações como rotação, translação, escala e corte. Em geometria, conforme FAUGERAS [Fau93], a rotação e a translação podem ser desfeitas pela escolha manual de 4 pontos correspondentes na imagem original e na imagem transformada. No entanto, no corte e na redução de escala temos perda de informação, inclusive nos coeficientes mais significativos da imagem. Ainda assim, o espalhamento da Marca d'Água de Cox por toda a extensão da imagem

[5] *Joint Photographics Experts Group*

permite que, pelo menos parcialmente, ela seja recuperável. Embora o corte e a redução de escala possam prejudicar seriamente a Marca d'Água, provavelmente a imagem resultante de nada adiantará ao atacante pois estará degradada e perderá seu *valor comercial*

- Processamento de sinal: neste categoria encontramos distorções de imagem como *resampling*, *dithering* e *recompression* assim como ajustes de contraste e cor. Segundo COX, KILIAN, LEIGHTON e SHAMOON [CKLS97], dado que temos a imagem original no processo de extração da Marca d'Água, é possível desfazer algumas operações de processamento de sinal, pelo menos de forma aproximada. Neste tipo de ataque, se a Marca d'Água é seriamente prejudicada, não há como garantir que a imagem transformada perderá *valor comercial* pois o atacante poderá argumentar que é uma criação original sua. Este caso tem similaridade com o exemplo de plágio do logotipo das Olimpíadas que vimos anteriormente

- Conversão D/A ou A/D: é esperada alguma degradação da imagem quando esta é impressa ou escaneada. Degradação é sinônimo de perda de *valor comercial*

A Marca d'Água deve ser resistente aos *ataques* acima mencionados, tanto de forma isolada quanto combinada. Deve também proteger a Marca d'Água contra falsificações e conluio (COX, KILIAN, LEIGHTON e SHAMOON [CKLS97]).

3.6 Conclusões do capítulo

Vimos que soluções baseadas em Marca d'Água podem ser relevantes na proteção dos direitos autorais de imagem em caso de disputa. Vimos também que não há uma solução absolutamente segura, como em outras áreas da Criptografia, embora métodos complementares possam ser utilizados na proteção dos direitos autorais de imagem para minimizar riscos. Introduzimos o conceito de *valor comercial* de uma imagem, e vimos que ataques realizados à Marca d'Água podem resultar em imagens sem *valor comercial* para o atacante.

Apresentamos o bastante conhecido método de Marca d'Água de COX, KILIAN, LEIGHTON e SHAMOON [CKLS97] e o segmentamos em camadas de *conteúdo* e *transporte*. No próximo capítulo apresentaremos um novo método derivado de Cox no qual a camada de *conteúdo* será substituída por uma nova informação relevante.

Capítulo 4

Método MACV

4.1 Objetivo do capítulo

O objetivo deste capítulo é apresentar o MACV, método original que utiliza as tecnologias combinadas de Marca d'Água e Criptografia Visual para a proteção dos direitos autorais de imagens estáticas.

Este capítulo está estruturado da seguinte forma: a Seção 4.2 apresenta as ideias principais na construção do MACV. A Seção 4.3 descreve de forma detalhada os três algoritmos que compõem o MACV: o MACV.g - Gerador de Marca d'Água, o MACV.e - Encapsulador de Marca d'Água e o MACV.v - Verificador de Marca d'Água. A Seção 4.4 mostra os resultados práticos do MACV obtidos por sua aplicação a imagens em tons de cinza e a imagens coloridas. A Seção 4.5 traz a análise de segurança do MACV segundo suas camadas de *transporte* e *conteúdo*. A Seção 4.6 analisa as propriedades ideais de Marca d'Água atendidas pelo MACV. A Seção 4.7 traz um resumo bibliográfico contendo outras soluções para proteção de direitos autorais de imagens estáticas baseadas em Marca d'Água e Criptografia Visual. Finalmente a Seção 4.8 apresenta as conclusões deste capítulo.

4.2 ideias principais do MACV

Vimos no capítulo 3 que a proteção dos direitos autorais de imagens estáticas é um problema atual e relevante em Criptografia. Neste exato momento pessoas estão copiando e alterando imagens estáticas sem o devido respeito aos direitos autorais. A tecnologia de Marca d'Água é um dos métodos utilizados para aumentar a segurança dos verdadeiros autores no sentido de provar a autoria de determinada imagem. Observamos também que, até onde nossa pesquisa alcançou, não há uma solução de Marca d'Água (ou de qualquer outra metodologia similar) absolutamente segura para a proteção dos direitos autorais de imagens estáticas.

O princípio básico das técnicas de Marca d'Água é o encapsulamento de informações relevantes de autoria na própria imagem a ser protegida, resultando no que denominamos

imagem marcada. É desejável que as diferenças entre a imagem original e sua correspondente imagem marcada não sejam visualmente perceptíveis. A qualquer momento a Marca d'Água pode ser extraída da imagem marcada, fornecendo informações sobre seu verdadeiro autor em um eventual processo de disputa. Porém, em ambiente desprotegido (na Internet, por exemplo), esta imagem marcada pode sofrer manipulações que prejudiquem (ou eliminem totalmente) a Marca d'Água encapsulada. Desta forma, é desejável que o método de Marca d'Água utilizado seja robusto, ou seja, que as informações encapsuladas na imagem marcada permaneçam relevantes mesmo que esta imagem sofra manipulações como compactação, corte, rotação, filtros, etc. Em um método robusto, a eliminação da Marca d'Água somente é possível se a imagem resultante degrada a tal ponto de não possuir *valor comercial*.

A Figura 4.1 ilustra a expectativa de perda de *valor comercial* de uma imagem marcada ao ser atacada (por manipulação da imagem, involuntária ou não) para remoção da Marca d'Água robusta.

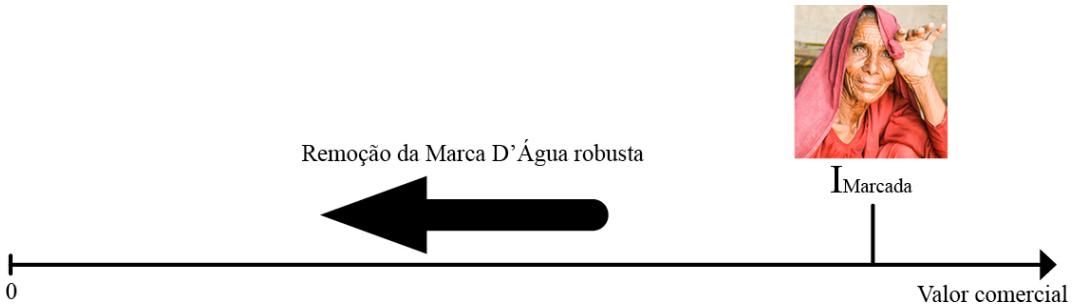

Figura 4.1: *Perda de* valor comercial *de uma imagem marcada ao ser atacada para remoção da Marca d'Água robusta*

Entre os métodos existentes de Marca d'Água robusta adotamos como referência a solução de COX, KILIAN, LEIGHTON e SHAMOON [CKLS97]. Neste método, por meio da transformada DCT, identificamos as partes visualmente mais significativas da imagem original e as alteramos para encapsular a Marca d'Água. Um das características positivas desta técnica é que a imagem marcada não apresenta diferenças visuais relevantes em relação à sua original; outra propriedade interessante é o espalhamento da Marca d'Água em toda a superfície da imagem marcada. Os resultados apontam que o método de Cox é bastante resistente às operações de processamento de imagem, embora apresente vulnerabilidades como vimos na Seção 3.5.

Para efeito desta dissertação segmentamos o método de Cox em camadas de *conteúdo* e *transporte*. Sem perda de generalidade, na proposição do nosso novo método utilizaremos a camada *transporte* de Cox em sua forma original (ver Seção 3.4). Embora não proteja a Marca d'Água contra todos os possíveis ataques, ela é suficientemente robusta para mostrar os nossos resultados dado que nenhum outro método conhecido de Marca d'Água é completamente seguro. Se tal método completamente seguro existir, é esperado que seja possível segmentá-lo em camadas de *conteúdo* e *transporte*; neste caso, é possível que a camada *con-*

teúdo possa ser substituída da mesma forma que estamos realizando com o método de Cox. Assim sendo concentraremos nossos esforços em melhorias da camada de *conteúdo* de Cox conforme veremos em seguida. Com as premissas acima construiremos o nosso novo método em 5 etapas.

A Figura 4.2 ilustra a Etapa 1/ 5 de construção do novo método de Marca d'Água. Adotaremos a camada *transporte* de Cox em nosso novo método e substituiremos a camada de *conteúdo* por outra a ser definida.

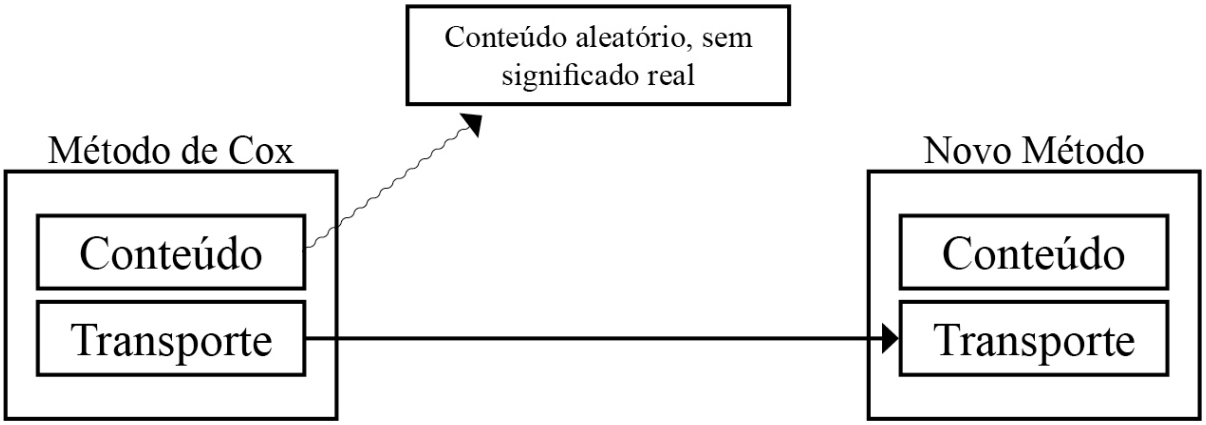

Figura 4.2: *Construção do novo método - Etapa 1/ 5: adotar a camada* transporte *de Cox e excluir a camada* conteúdo

Como camada de *conteúdo* entendemos a Marca d'Água propriamente dita, ou seja, o conjunto de informações relevantes e suficientes para provar a autoria da imagem original. Na camada *conteúdo* de Cox este conjunto de informações é uma sequência pseudo-aleatória sem qualquer significado . Variações foram propostas à camada de *conteúdo* de Cox utilizando-se, por exemplo, o logotipo do autor da imagem. Porém, neste caso, o atacante pode ter uma vantagem adicional se utilizar o conhecimento do logotipo do autor para encontrar, alterar ou excluir a Marca d'Água da imagem marcada.

A Figura 4.3 ilustra a Etapa 2/ 5 de construção do novo método de Marca d'Água. Não usaremos na camada de *conteúdo* informações com significado porém de baixa entropia (como o logotipo do autor, por exemplo).

Entendemos como deficiências da camada *conteúdo* tanto o uso de informação sem significado como o uso de informação de baixa entropia. Estas deficiências nos inspiraram a construir um novo método de Marca d'Água no qual a informação relevante encapsulada na imagem marcada tem significado e apresenta *segurança perfeita*. Em Criptografia temos *segurança perfeita* quando, mesmo que o atacante intercepte uma mensagem criptografada, ele não consegue nenhuma informação sobre a mensagem legível original (TERADA [Ter00]).

A Figura 4.4 ilustra a Etapa 3/ 5 de construção do novo método de Marca d'Água. Usaremos uma informação, em forma de imagem, com significado e que apresente segurança perfeita na camada de *conteúdo*.

No Capítulo 2 introduzimos o conceito de Criptografia Visual, método no qual um se-

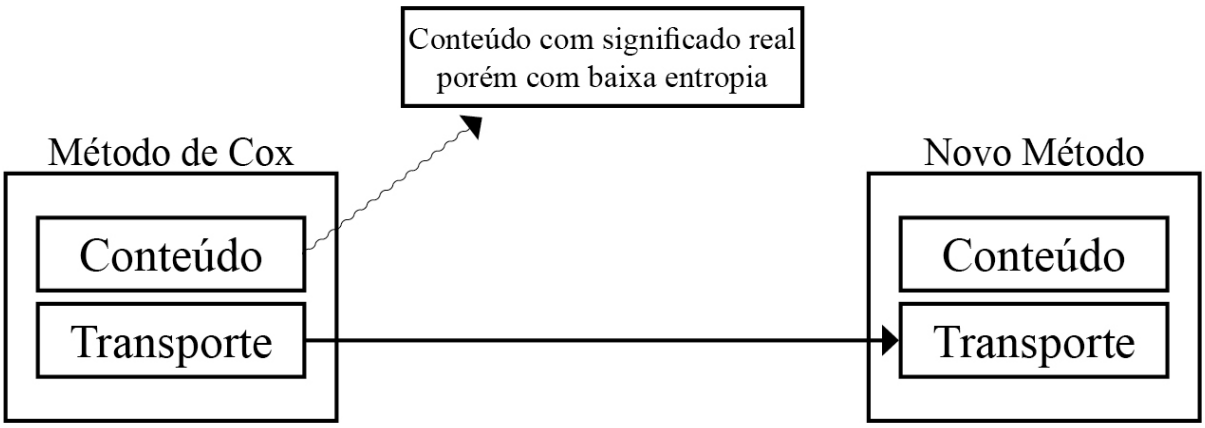

Figura 4.3: *Construção do novo método - Etapa 2/ 5: não usar informações de baixa entropia na camada* conteúdo

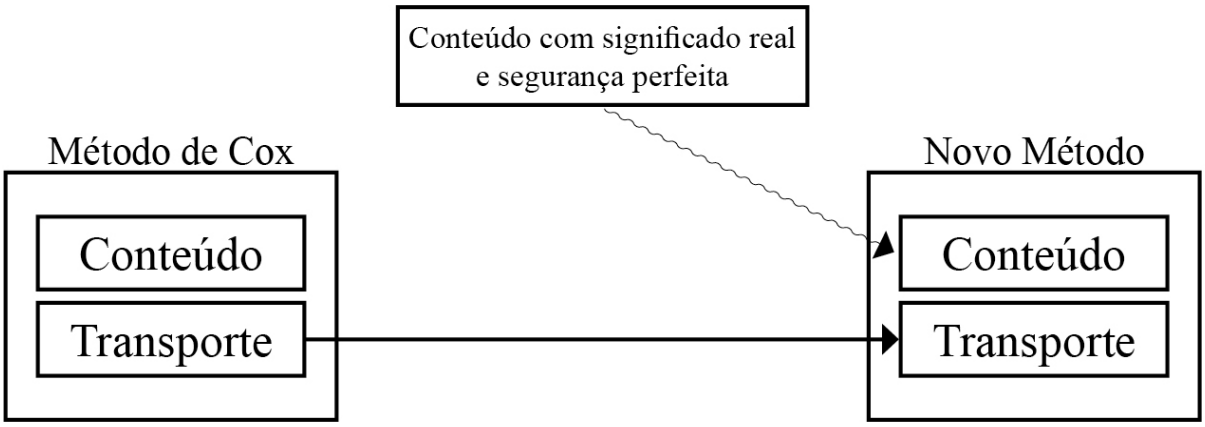

Figura 4.4: *Construção do novo método - Etapa 3/ 5: usar informação com significado e segurança perfeita na camada* conteúdo

gredo é criptografado em n transparências. A sobreposição de k ou mais destas transparências revela o segredo; este modelo é denominado $(k, n) - VCS$. Vimos na Seção 2.3 que o Modelo Básico $(2, 2) - VCS$ apresenta a propriedade de *segurança perfeita*: mesmo que o atacante obtenha uma das duas transparências criptografadas ele não consegue nenhuma pista adicional sobre o segredo. Desta forma, em nosso novo método, adotaremos umas das duas transparências geradas $(2, 2) - VCS$ em substituição à camada *conteúdo* de Cox. Esta transparência estará encapsulada na imagem marcada no papel de Marca d'Água; ao ser extraída e sobreposta à outra transparência, o segredo é revelado.

A Figura 4.5 ilustra a Etapa 4/ 5 de construção do novo método de Marca d'Água. Usaremos uma informação com significado criptografada pelo $(2, 2) - VCS$ na camada de *conteúdo*.

Alcançado o requisito de *segurança perfeita*, o próximo passo para completar nosso novo método é definir uma informação que tenha significado e que seja de alta entropia. Esta informação com significado , em forma de imagem, será submetida ao $(2, 2) - VCS$ e posteriormente usada como Marca d'Água. O uso do logotipo do autor é descartado por sua baixa

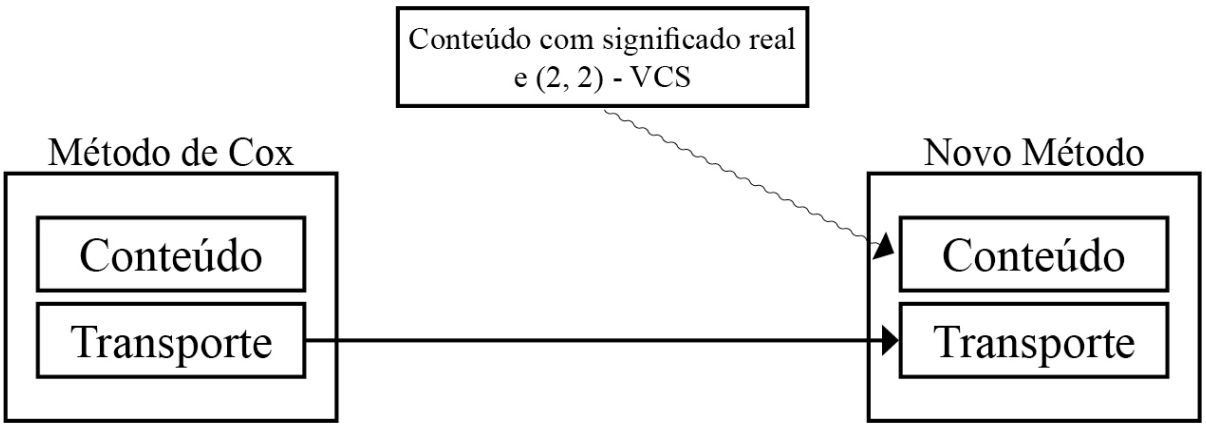

Figura 4.5: *Construção do novo método - Etapa 4/ 5: usar informação com significado criptografada pelo* $(2,2) - VCS$ *na camada* conteúdo

entropia. Uma escolha natural é o uso de uma função de hashing aplicada à imagem que será protegida. Adotamos como função de hashing nesta dissertação o SHA (ver TERADA [Ter00]) que, sendo um algoritmo público, pode ser verificável por ambas as partes em uma eventual disputa na justiça pelos direitos autorais desta imagem. Pela característica de modularidade da solução MACV que apresentaremos a seguir, o SHA poderia ser substituído por outra função de hashing considerada mais segura.

A Figura 4.6 ilustra a Etapa 5/ 5 de construção do novo método de Marca d'Água. Aplicaremos um algoritmo de hash na imagem a ser protegida; o resultado será criptografado pelo $(2,2) - VCS$ e inserido na camada de *conteúdo*.

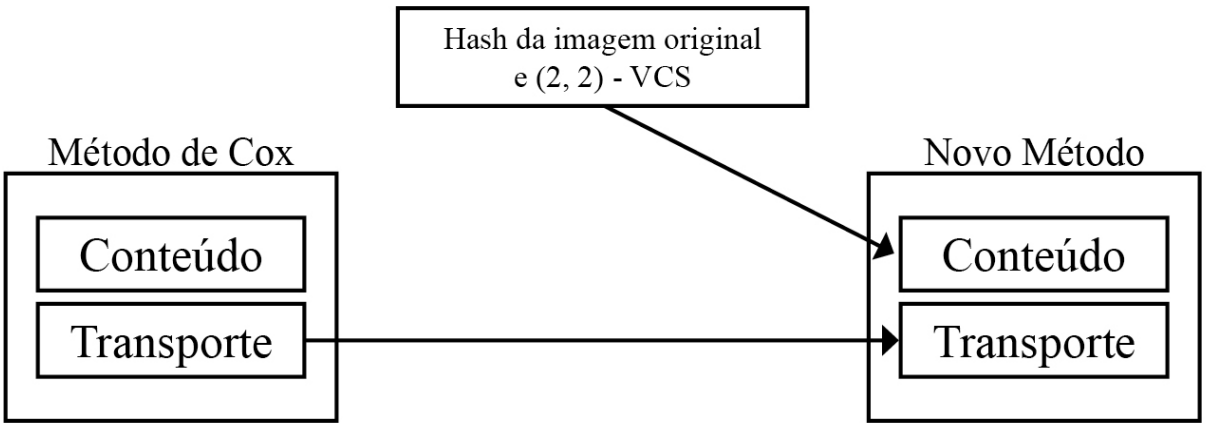

Figura 4.6: *Construção do novo método - Etapa 5/ 5: aplicar hash e* $(2,2) - VCS$ *para produzir a camada* conteúdo

Expostas as ideias principais do nosso novo método, a partir deste momento denominado MACV (iniciais de Marca d'Água e Criptografia Visual), detalharemos nas próximas seções seu funcionamento.

4.3 Algoritmos do MACV

O MACV pode ser segmentado pelos seus algoritmos:

- MACV.g - Gerador de Marca d'Água do MACV: gera as transparências $M_{Publica}$ e $M_{Secreta}$ pela aplicação do $(2,2) - VCS$ ao resultado do hash sobre a imagem a ser protegida. É o gerador da camada *conteúdo*
- MACV.e - Encapsulador de Marca d'Água do MACV: calcula a transformada DCT da imagem a ser protegida e ajusta os maiores coeficientes utilizando $M_{Publica}$ como parâmetro de entrada. É o gerador da camada *transporte* e segue o método de Cox
- MACV.v - Verificador de Marca d'Água do MACV: extrai $M_{Publica}$ e sobrepõe a $M_{Secreta}$ para verificação visual do resultado do hash da imagem original. Não tem qualquer paralelo com o método de Cox

A Figura 4.7 ilustra a execução do MACV. O MACV.g e o MACV.e são executados pelo autor em ambiente protegido (conforme práticas de segurança descritas na Seção 3.2.3). Em caso de disputa (no tribunal, por exemplo), utiliza-se o MACV.v para defender a autoria da imagem original $I_{Original}$.

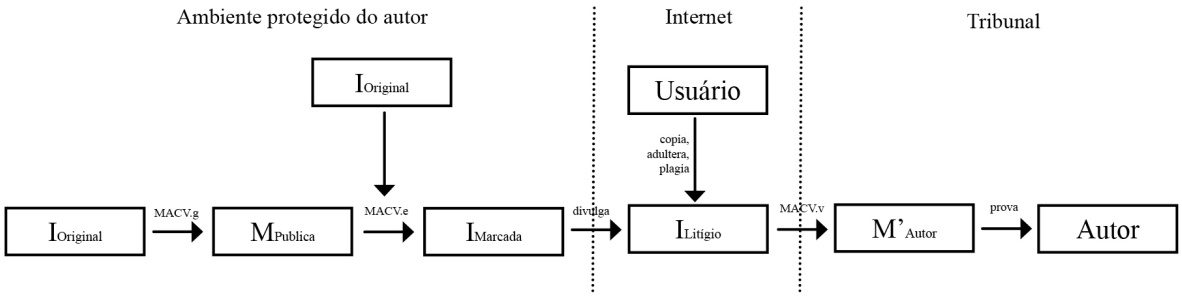

Figura 4.7: *Fluxo de execução do MACV em diferentes ambientes: protegido, Internet e no tribunal*

Veremos a seguir a descrição de cada um destes algoritmos.

4.3.1 MACV.g - Gerador de Marca d'Água do MACV

O objetivo do MACV.g, ou Gerador de Marca d'Água do MACV, é produzir a Marca d'Água que será encapsulada na imagem a ser protegida. Primeiro o MACV.g calcula a marca M_{Autor} em forma de imagem e posteriormente calcula duas transparências pelo $(2,2) - VCS$: $M_{Publica}$ e $M_{Secreta}$. Quando observadas isoladamente, $M_{Publica}$ e $M_{Secreta}$ se assemelham a ruído aleatório e não apresentam qualquer significado visual. Porém, quando sobrepostas, revelam M_{Autor}.

A Figura 4.8 ilustra o fluxo de execução do MACV.g.

- Parâmetros de entrada: a imagem original $I_{Original}$

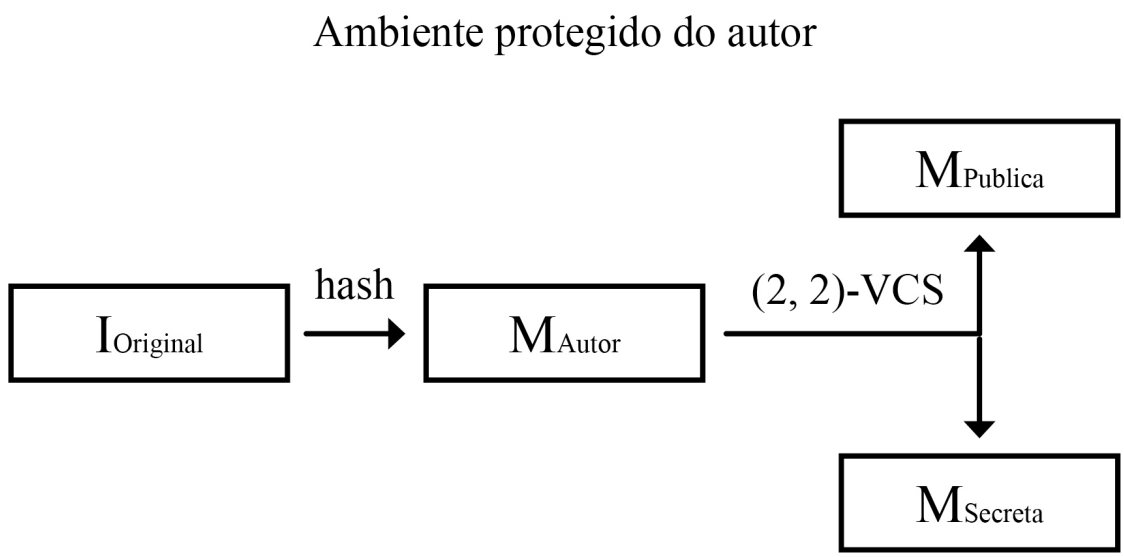

Figura 4.8: *Fluxo de execução do MACV.g*

- Parâmetros de saída: as transparências $M_{Publica}$ e $M_{Secreta}$. A marca M_{Autor} é um resultado temporário do algoritmo e poderá ser descartado ao final da execução de MACV.g
  - $M_{Publica}$: faz o papel de Marca d'Água e será encapsulada em $I_{Original}$ pelo MACV.e como veremos em seguida
  - $M_{Secreta}$: deve ser mantida secreta pelo autor, supondo como premissa a execução das práticas de segurança descritas na Seção 3.2.3. Somente será revelada em caso de disputa por $I_{Original}$
- Passos de execução do MACV.g
  1. Calcule M_{Autor} pela aplicação da função de hashing SHA-256 sobre a imagem $I_{Original}$
  2. Converta M_{Autor} para o formato de imagem binária
  3. Execute o $(2,2) - VCS$ sobre M_{Autor} obtendo-se $M_{Publica}$ e $M_{Secreta}$

4.3.2 MACV.e - Encapsulador de Marca d'Água do MACV

O objetivo do MACV.e, ou Encapsulador de Marca d'Água do MACV, é encapsular a marca $M_{Publica}$ na imagem original $I_{Original}$. É uma implementação da camada de *transporte* de Cox para inserção de Marca d'Água.

A Figura 4.9 ilustra o fluxo de execução do MACV.e.

- Parâmetros de entrada: a imagem original $I_{Original}$ e a respectiva Marca d'Água $M_{Publica}$
- Parâmetros de saída: a imagem marcada $I_{Marcada}$

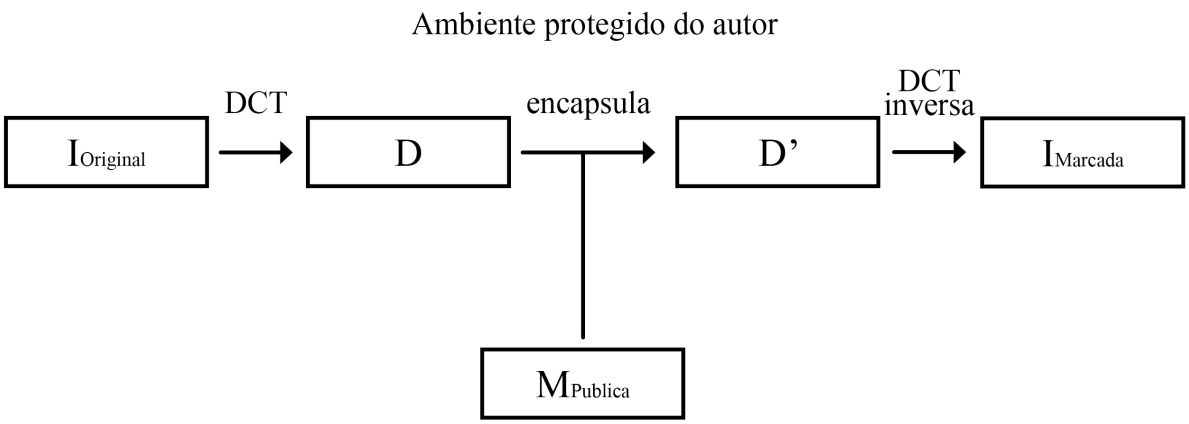

Figura 4.9: *Fluxo de execução do MACV.e*

- Passos de execução do MACV.e

  1. Considere os n bits de $M_{Publica}$ como $x_1, \cdots, x_n$
  2. Calcule a matriz D pela transformação DCT de $I_{Original}$
  3. Calcule a sequência ordenada $v_1, \cdots, v_n$ com os n maiores coeficientes DCT de D
  4. Calcule $v'_i = v_i(1 + \alpha x_i)$ para $1 \leq i \leq n$
  5. Substitua $v_1, \cdots, v_n$ por $v'_1, \cdots, v'_n$ em D obtendo a matriz D'
  6. Calcule $I_{Marcada}$ pela transformação DCT inversa de D'

4.3.3 MACV.v - Verificador de Marca d'Água do MACV

Os objetivos do MACV.v, ou Verificador de Marca d'Água do MACV, são:

1. Extrair a marca $M_{Extraida}$ da imagem marcada $I_{Marcada}$. É a implementação da camada de *transporte* de Cox para extração da Marca d'Água. É um algoritmo de domínio público e que pode ser usado por ambas as partes litigantes

2. Calcular o hash SHA-256 de $I_{Original}$ obtendo-se M_{Autor}. O SHA-256 é um algoritmo de domínio público

3. Sobrepor $M_{Extraida}$ a $M_{Secreta}$ obtendo-se M'_{Autor} (segredo revelado conforme o $(2,2) - VCS$, algoritmo de domínio público)

O MACV.v foi construído para ser utilizado em situações de disputa por direitos autorais de uma imagem estática. Por exemplo, em um tribunal. Utiliza algoritmos de domínio público (camada de *transporte* de Cox, SHA-256 e $(2,2) - VCS$) e, desta forma, pode ser executado por ambas as partes litigantes. A comparação entre M_{Autor} e M'_{Autor} é visual.

A Figura 4.10 ilustra o fluxo de execução do MACV.v.

- Parâmetros de entrada:

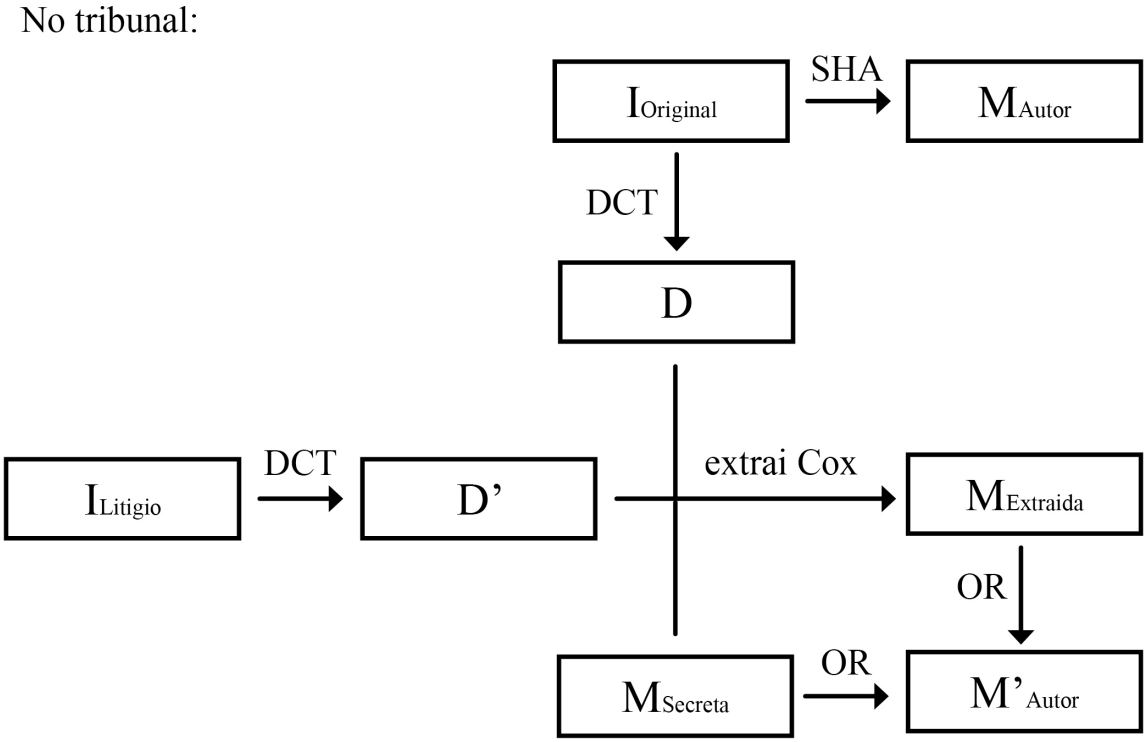

Figura 4.10: *Fluxo de execução do MACV.v. A imagem original $I_{Original}$ e a respectiva transparência $M_{Secreta}$ são apresentadas pelo autor verdadeiro. A imagem em disputa $I_{Litigio}$ é apresentada pela outra parte litigante*

  - A imagem original $I_{Original}$ e a respectiva transparência $M_{Secreta}$, apresentadas pelo autor verdadeiro
  - A imagem em disputa $I_{Litigio}$, apresentada pela outra parte litigante

- Parâmetros de saída: as Marcas d'Água M_{Autor} (obtida de $I_{Original}$) e M'_{Autor} (obtida de $I_{Litigio}$)

- Passos de execução do MACV.v

  1. Calcule M_{Autor} como a aplicação da função de hashing SHA-256 sobre a imagem $I_{Original}$
  2. Considere n a quantidade de bits de $M_{Secreta}$
  3. Calcule a matriz D pela transformação DCT de $I_{Original}$
  4. Calcule a matriz D' pela transformação DCT de $I_{Litigio}$
  5. Calcule a sequência ordenada $v_1, \cdots, v_n$ com os n maiores coeficientes DCT de D, onde $v_i = D[p_i, q_i]$
  6. Considere a sequência $v'_1, \cdots, v'_n$ onde $v'_i = D'[p_i, q_i]$
  7. Calcule a sequência $x'_1, \cdots, x'_n$ onde $x'_i = (v'_i/v_i - 1)/\alpha$
  8. Considere os n bits de $M_{Extraida}$ como $x'_1, \cdots, x'_n$
  9. Calcule bit a bit $M'_{Autor} = M_{Extraida}$ OR $M_{Secreta}$

4.4 Resultados do MACV

A inserção de uma Marca d'Água $M_{Publica}$ em uma imagem $I_{Original}$ pode provocar degradação visual da imagem resultante $I_{Marcada}$. Esta degradação pode acarretar em perda do *valor comercial* da imagem, situação que deve ser evitada. Por outro lado, se nossa única preocupação for evitar a degradação visual, é possível que a Marca d'Água encapsulada seja muito frágil e facilmente removível. Há que se encontrar um equilíbrio entre a possível degradação visual da imagem e a robustez da Marca d'Água encapsulada.

Analisando o MACV.e, descrito em 4.3.2, percebe-se que este equilíbrio é controlado por dois parâmetros:

- Tamanho da Marca d'Água n: indica a quantidade de coeficientes DCT elegíveis da $I_{Original}$ que serão afetados pela inserção de $M_{Publica}$. É fácil notar que, quanto maior o n, maior será o impacto visual em $I_{Marcada}$. Por outro lado, a utilização de n pequeno reduz a resistência da Marca d'Água aos ataques por processamento de imagem pois uma das principais características de segurança do MACV é o espalhamento da Marca d'Água pela superfície da $I_{Marcada}$

- Fator de escala α: indica a escala utilizada para alterar cada um dos coeficientes DCT elegíveis da $I_{Original}$. Quanto maior o α, maior o impacto visual em $I_{Marcada}$. Quanto menor o α, maior o risco de ruído na Marca d'Água resultante de processamento de imagem sobre $I_{Marcada}$

Vimos que uma das propriedades da Marca d'Água é sua capacidade de absorção, ou seja, o menor tamanho necessário da imagem $I_{Original}$ para que a inserção da Marca d'Água não degrade visualmente a imagem resultante. Chamaremos a capacidade de absorção da Marca d'Água de λ. A determinação de λ está diretamente relacionada com o tamanho n da Marca d'Água $M_{Publica}$.

No MACV adotamos a função de hashing SHA-256 para calcular M_{Autor}. A saída do SHA-256 é composta por 32 dígitos hexadecimais. No MACV.g estes 32 dígitos são convertidos para o formato de imagem binária, cada um deles assumindo, em média, 20 pixels de largura por 15 pixels de altura. Consequentemente, M_{Autor} teria 20×480 pixels em sua forma de imagem.

Em nossa implementação adotamos o $(2,2) - VCS$ em sua forma clássica, ou seja, cada pixel da imagem a ser criptografada se transforma em uma matriz de 2×2 sub-pixels em suas transparências resultantes. Logo, ao se aplicar o $(2,2) - VCS$ sobre M_{Autor} obtemos a Marca d'Água $M_{Publica}$ com 40×960 pixels. O impacto de se usar $M_{Publica}$ com tal magnitude é a degradação visual da imagem marcada; ou, alternativamente, a necessidade de se incrementar a capacidade de absorção λ de $M_{Publica}$, reduzindo-se assim sua aplicabilidade prática. Poderíamos reduzir tais impactos implementando-se o $(2,2) - VCS$ sem expansão de pixels como o de KUWAKADO e TANAKA [KT99] ou, de forma complementar, poderíamos agrupar o resultado do SHA-256 em 16 pares de 2 dígitos, cada par sendo representado

por um caractere ASCII[1]. A solução imediata que adotamos, sem perda de generalidade, foi a de usar M_{Autor} com 20×100 pixels (adotamos a representação os sete primeiros dígitos do resultado do SHA-256).

Em termos de espaço de armazenamento, considerando o ambiente do autor, o MACV precisa manter 2 arquivos auxiliares, $M_{Publica}$ e $M_{Secreta}$, a cada $I_{Original}$. Isso representa um acréscimo de 4% ao espaço ocupado por $I_{Original}$. A $I_{Marcada}$ é gerada conforme a necessidade e não precisa ser armazenada no ambiente do autor. Não estamos considerando o espaço ocupado por $I_{Marcada}$ em ambientes públicos como a Internet.

Sobre tempo de execução o ponto crítico do MACV é a ordenação dos coeficientes DCT. Isso ocorre no MACV.g e no MACV.v, uma vez por imagem. Quanto maior a imagem tratada, maior será o tempo de execução, que cresce de forma não linear em relação ao tamanho da imagem. Não foi objeto desta dissertação a pesquisa de algoritmos eficientes para esta ordenação.

Com as considerações acima podemos afirmar que nossos experimentos com o MACV apresentaram bons resultados, como veremos em seguida, utilizando-se os seguintes parâmetros:

- Tamanho da Marca d'Água $n = 8.000$ pixels ($M_{Publica}$ com 200 pixels de largura por 40 pixels de altura)
- Fator de escala $\alpha = -0,01$ (comparativamente, Cox utilizou $\alpha = +0,1$ em seus experimentos
- Capacidade de absorção da Marca d'Água $\lambda = 0,67$ milhões de pixels

4.4.1 Aplicação dos algoritmos MACV a uma imagem em tons de cinza

Nos experimentos desta Seção será usada a imagem original em tons de cinza *Pescador Tailandês*, mostrada na Figura 4.11, com 1.000×666 pixels, cada pixel representando 256 tons de cinza.

4.4.1.1 Aplicação do MACV.g a uma imagem em tons de cinza

O resultado da função de hashing SHA-256 aplicada à imagem *Pescador Tailandês* é:

$$6DE9058C70EFC45CB9345B50512A2A66B37EFC04$$

Na Figura 4.12 temos o fluxo de aplicação do MACV.g à $I_{Original}$. Os resultados da aplicação de MACV.g à imagem *Pescador Tailandês* são:

- As transparências $M_{Publica}$ e $M_{Secreta}$, ambas com 200×40 pixels

[1] *American Standard Code for Information Interchange*

Figura 4.11: $I_{Original}$*: imagem em tons de cinza* Pescador Tailandês *com* 1.000×666 *pixels*

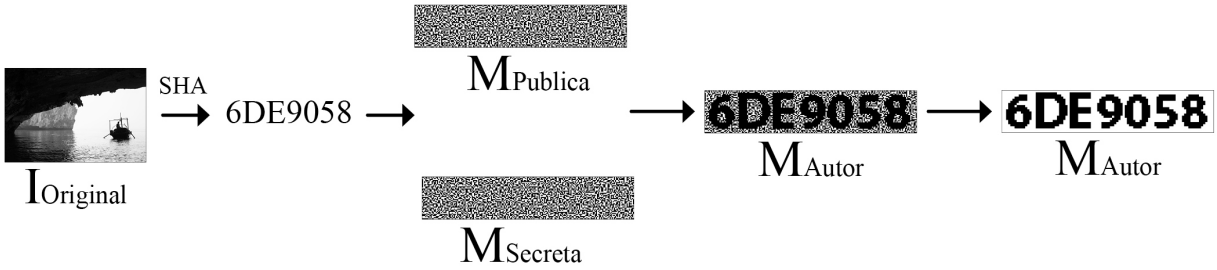

Figura 4.12: *Aplicação de MACV.g à imagem original* Pescador Tailandês*. As resultantes* $M_{Publica}$ *e* $M_{Secreta}$ *têm* 200×40 *pixels*

- A imagem M_{Autor} de 200×40 pixels, obtida pela sobreposição (OR) de $M_{Publica}$ e $M_{Secreta}$
- A imagem M_{Autor} após aplicação de filtro para maior visibilidade

4.4.1.2 Aplicação do MACV.e a uma imagem em tons de cinza

O fluxo de encapsulamento da Marca d'Água $M_{Publica}$ pelo MACV.e na imagem original *Pescador Tailandês* pode ser observado na Figura 4.13. A imagem resultante $I_{Marcada}$ pode ser vista na Figura 4.14

Como não há diferença visual significativa entre $I_{Original}$ e $I_{Marcada}$ pode-se supor que, se $I_{Original}$ tem *valor comercial*, então a respectiva $I_{Marcada}$ tem o mesmo valor.

Uma dos parâmetros utilizados nesta implementação é o fator de escala α definido por

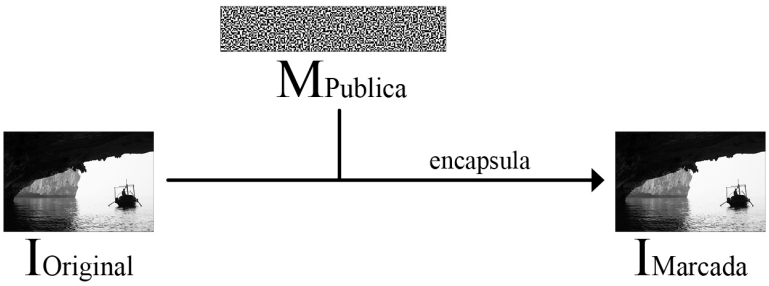

Figura 4.13: *Fluxo de encapsulamento da Marca d'Água $M_{Publica}$ pelo MACV.e na imagem original* Pescador Tailandês

Figura 4.14: *A imagem $I_{Marcada}$ tem* 1.000×666 *pixels e é o resultado da aplicação do MACV.e à imagem original* Pescador Tailandês. *As imagens $I_{Original}$ e a respectiva $I_{Marcada}$ são visualmente muito parecidas, embora tenham 38,4% pixels diferentes entre si*

Cox. Em seus experimentos Cox utilizou $\alpha = 0,1$, em nossa implementação utilizamos $\alpha = -0,01$. Caso utilizássemos $\alpha = 0,1$ teríamos um impacto visual muito maior em $I_{Marcada}$ conforme observamos na Figura 4.15. Isso se explica por que no MACV alteramos oito vezes mais coeficientes DCT do que em Cox.

4.4.1.3 Aplicação do MACV.v a uma imagem em tons de cinza

No MACV.v as imagens $I_{Original}$ e $M_{Secreta}$ são fornecidas pelo autor verdadeiro. A imagem $I_{Litigio}$ é fornecida pela outra parte litigante. O uso do MACV.v somente faz sentido em um ambiente de resolução de disputas (no tribunal, por exemplo). Os resultados do

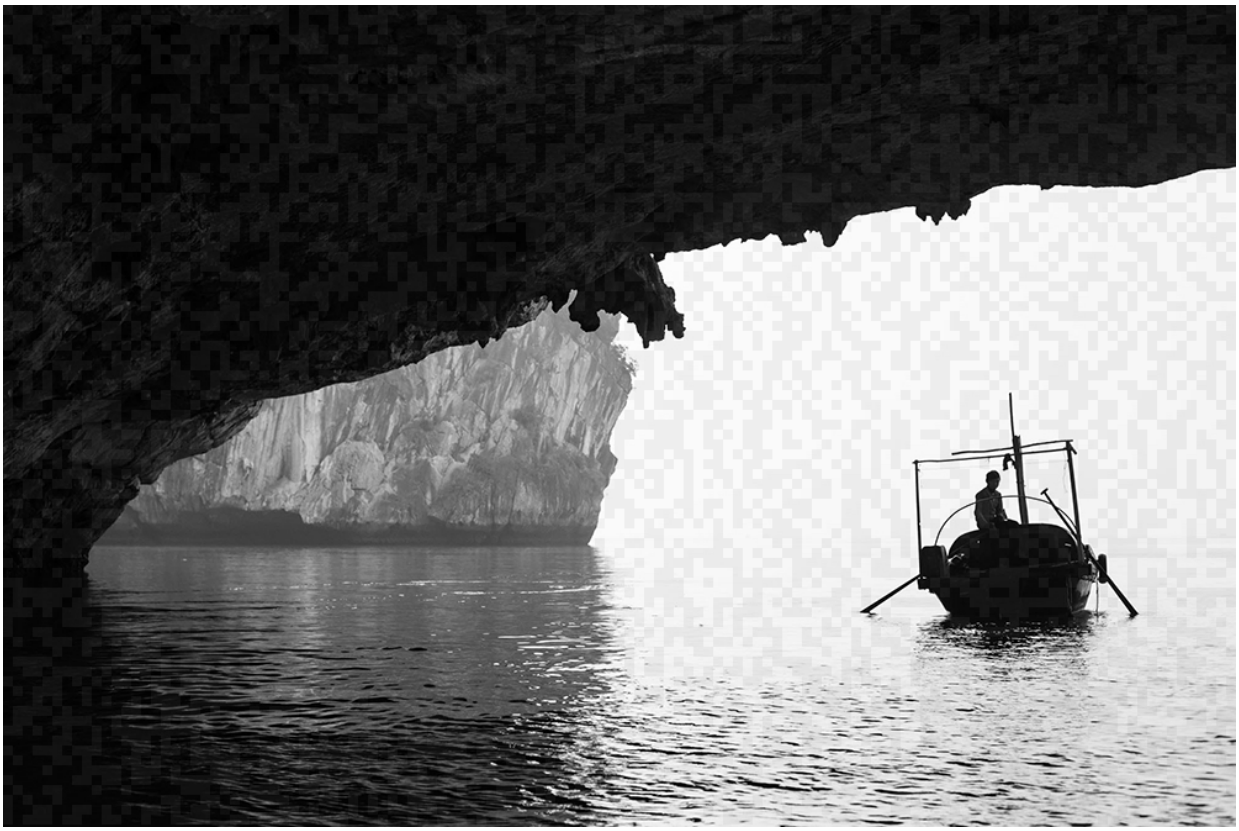

Figura 4.15: *A imagem* $I_{Marcada}$ *é gerada com* $\alpha = 0,1$ *ao inves de* $\alpha = 0,01$

MACV.v são as Marcas d'Água: M_{Autor} (obtida de $I_{Original}$) e M'_{Autor} (extraída de $I_{Litigio}$). A comparação entre M_{Autor} e M'_{Autor} é visual é não é atendida pelo MACV.v.

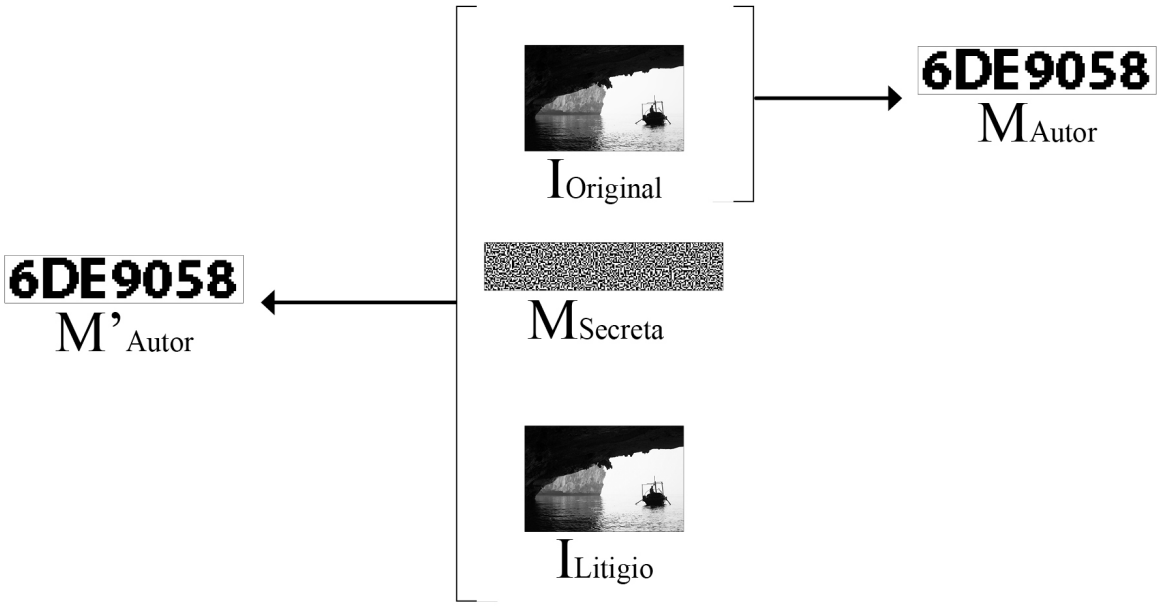

Figura 4.16: *Fluxo de verificação da Marca d'Água pelo MACV.v aplicado à imagem* Pescador Tailandês

O fluxo de verificação da Marca d'Água pelo MACV.v é mostrado pela Figura 4.16. Neste exemplo adotou-se $I_{Litigio} = I_{Marcada}$. Ou, dito de outra forma, $I_{Marcada}$ não sofreu qualquer

tipo de manipulação e foi levada ao tribunal para uma disputa de autoria. Percebe-se que, se o falso autor não alterar $I_{Marcada}$, teremos $M_{Autor} = M'_{Autor}$. Prova-se assim que o falso autor está afirmando equivocadamente que $I_{Marcada}$ é sua, pois, com alta probabilidade, somente o autor verdadeiro conseguiria calcular M_{Autor} e encapsulá-la em $I_{Marcada}$. Veremos mais sobre estas questões de disputa na Seção 4.5.

4.4.2 Aplicação dos algoritmos MACV a uma imagem colorida

Nos experimentos desta Seção será usada a imagem colorida *Senhora Indiana*, mostrada na Figura 4.17, com 1.000×1.000 pixels, cada pixel representando 256^3 cores.

Figura 4.17: *Imagem colorida* Senhora Indiana *com* 1.000×1.000 *pixels*

4.4.2.1 Aplicação do MACV.g a uma imagem colorida

O resultado da função de hashing SHA-256 aplicada à imagem *Senhora Indiana* é:

$$DC4D50B300B29AB65051EA814D7796063C7D3748$$

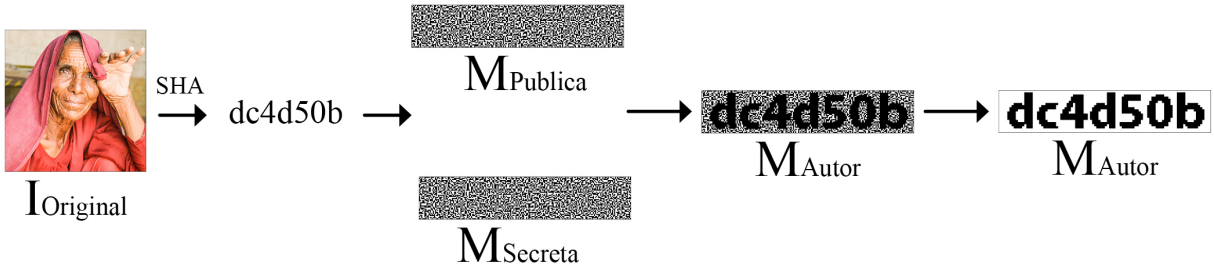

Figura 4.18: *Aplicação de MACV.g à imagem original* Senhora Indiana. *As resultantes* $M_{Publica}$ *e* $M_{Secreta}$ *têm* 200×40 *pixels*

Na Figura 4.18 temos o fluxo de aplicação do MACV.g à $I_{Original}$. Os resultados da aplicação de MACV.g à imagem *Senhora Indiana* são:

- As transparências $M_{Publica}$ e $M_{Secreta}$, ambas com 200×40 pixels
- A imagem M_{Autor} de 200×40 pixels, obtida pela sobreposição (OR) de $M_{Publica}$ e $M_{Secreta}$
- A imagem M_{Autor} após aplicação de filtro para maior visibilidade

4.4.2.2 Aplicação do MACV.e a uma imagem colorida

O fluxo de encapsulamento da Marca d'Água $M_{Publica}$ pelo MACV.e na imagem original *Senhora Indiana* pode ser observado na Figura 4.19. A imagem resultante $I_{Marcada}$ pode ser vista na Figura 4.20

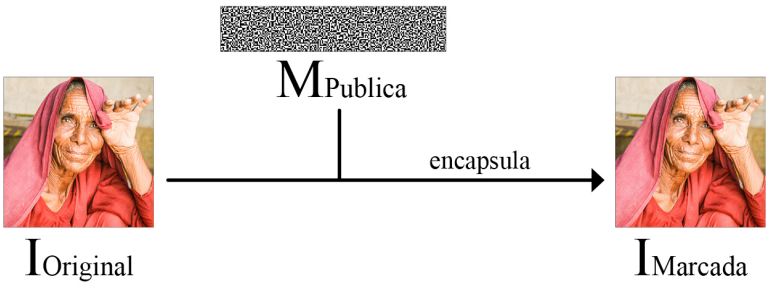

Figura 4.19: *Fluxo de encapsulamento da Marca d'Água* $M_{Publica}$ *pelo MACV.e na imagem original* Senhora Indiana

Como não há diferença visual significativa entre $I_{Original}$ e $I_{Marcada}$ pode-se supor que, se $I_{Original}$ tem *valor comercial*, então a respectiva $I_{Marcada}$ tem o mesmo valor.

Figura 4.20: *A imagem* $I_{Marcada}$ *tem* 1.000×1.000 *pixels e é o resultado da aplicação do MACV.e à imagem original* Senhora Indiana. *As imagens* $I_{Original}$ *e a respectiva* $I_{Marcada}$ *são visualmente muito parecidas, embora tenham 25,6% pixels diferentes entre si*

4.4.2.3 Aplicação do MACV.v a uma imagem colorida

No MACV.v as imagens $I_{Original}$ e $M_{Secreta}$ são fornecidas pelo autor verdadeiro. A imagem $I_{Litigio}$ é fornecida pela outra parte litigante. O uso do MACV.v somente faz sentido em um ambiente de resolução de disputas (no tribunal, por exemplo). Os resultados do MACV.v são as Marcas d'Água: M_{Autor} (obtida de $I_{Original}$) e M'_{Autor} (extraída de $I_{Litigio}$). A comparação entre M_{Autor} e M'_{Autor} é visual é não é realizada pelo MACV.v.

O fluxo de verificação da Marca d'Água pelo MACV.v é mostrado pela Figura 4.21. Neste exemplo adotou-se $I_{Litigio} = I_{Marcada}$, logo, M_{Autor} é idêntica a M'_{Autor}

Figura 4.21: *Fluxo de verificação da Marca d'Água pelo MACV.v aplicado à imagem* Senhora Indiana

4.5 Análise de segurança do MACV

A tecnologia atual não oferece recursos eficientes para impedir a cópia/ adulteração/ plágio de uma imagem sem autorização do verdadeiro autor. Desta forma, estas ações maliciosas ocorrem sem qualquer tipo de controle. O autor tem recursos escassos para provar sua autoria e defender seus direitos em uma eventual disputa judicial. Na época da fotografia analógica com filme, por exemplo, o negativo poderia ser usado como prova. Em nossa época digital essa possibilidade desapareceu.

Neste contexto, o uso do MACV permite que o autor *marque* suas imagens estáticas no momento de criação. Com o uso combinado do MACV.g e do MACV.e, uma imagem $I_{Original}$ é marcada e somente sua versão $I_{Marcada}$ é divulgada. Supomos como premissa que o autor mantém $I_{Original}$ e $M_{Secreta}$ armazenadas de forma segura. $I_{Original}$ e $M_{Secreta}$ terão papel relevante na prova de autoria, de forma análoga ao negativo de filme usado no passado.

Deve ser observado que $I_{Marcada}$ é passível de *ataque* assim que fica disponível em ambiente desprotegido. Entendemos como *ataque* qualquer operação que altere $I_{Marcada}$, seja de caráter malicioso (de um atacante para remover ou danificar a Marca d'Água) ou inconsciente (de um usuário sem intenção maliciosa).

Na definição do MACV adotamos duas camadas independentes: *transporte* e *conteúdo*. A camada *conteúdo* apresenta informações relevantes para a prova de autoria da imagem. A camada *transporte* deve ser suficientemente robusta para proteger a integridade da camada de *conteúdo* mesmo após os possíveis ataques à $I_{Marcada}$. Se a camada *transporte* falha em seu objetivo, a camada *conteúdo* possivelmente será corrompida e poderá se tornar inútil como prova de autoria.

Nossa pesquisa não encontrou qualquer camada de *transporte* que seja segura contra

todos os ataques possíveis. Simultaneamente nossos experimentos mostraram que a camada de *transporte* utilizada pelo MACV, que segue a solução de Cox, apresenta uma resistência razoável em diversas situações de ataque. A camada de *transporte* de Cox não é resistente ao ataque de marcações sucessivas. É possível que a evolução da tecnologia de Marca d'Água apresente uma nova camada de *transporte* que seja mais robusta do que as que temos no presente; neste caso, o MACV poderá ser beneficiado pela substituição da camada de *transporte* de Cox por esta nova mais robusta.

A seguir apresentaremos a análise de segurança do MACV com alguns dos possíveis ataques às camadas de *transporte* e *conteúdo*.

4.5.1 Segurança da camada de *transporte* do MACV

Todos os experimentos mostrados a seguir utilizam como $I_{Original}$ a *Senhora Indiana*, vista na Figura 4.17, com 1.000×1.000 pixels, cada pixel representando 256^3 cores. Aplicamos MACV.g e MACV.e a $I_{Original}$ e obtemos $M_{Publica}$ (vista na Figura 4.18) e $I_{Marcada}$ (vista na Figura 4.20). Em cada um dos ataques mostrados baixo a $I_{Marcada}$ é transformada em $I_{Litigio}$.

Quando utilizamos outras imagens coloridas, ou mesmo imagens em tons de cinza, as análises descritas abaixo são idênticas.

4.5.1.1 Redução de resolução a 36%

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ter sua resolução reduzida em 36% para 800×800 pixels. O resultado é visto na Figura 4.22, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

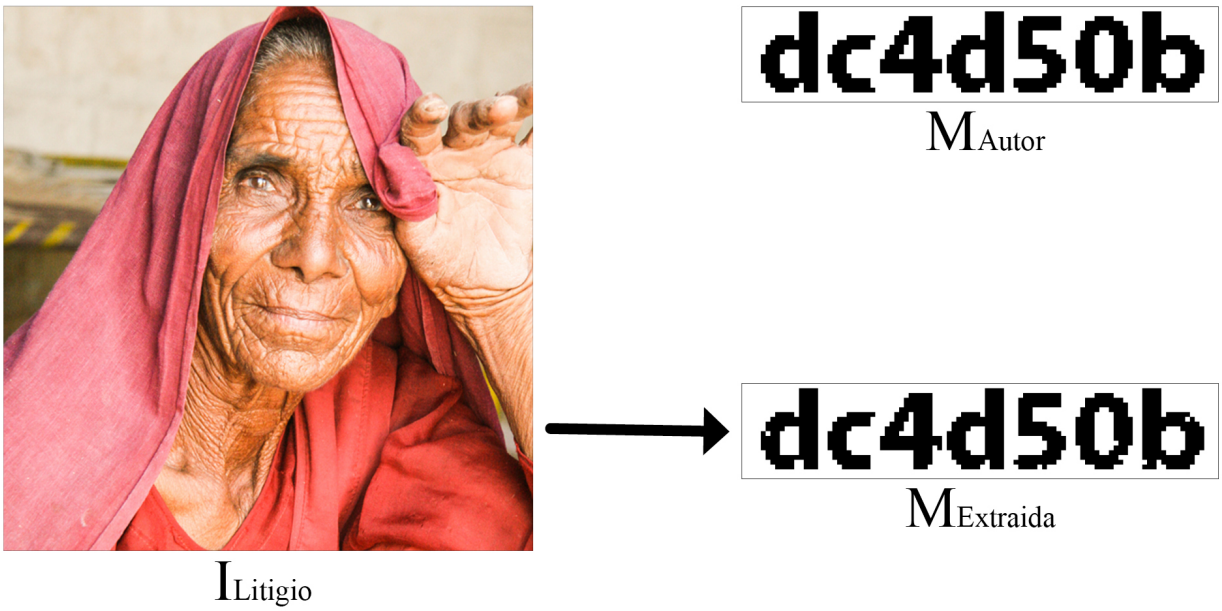

Figura 4.22: *Ataque: redução da resolução de $I_{Marcada}$ em 36%. Mostramos a $I_{Litigio}$ com 800×800 pixels e a respectiva $M_{Extraida}$*

Análise do resultado deste ataque:

- Não há degradação visual perceptível de $I_{Marcada}$ para $I_{Litigio}$
- Embora $I_{Litigio}$ tenha perdido 36% da informação contida em $I_{Marcada}$, a Marca d'Água obtida $M_{Extraida}$ sofreu pequena degradação em relação à M_{Autor}

4.5.1.2 Redução de resolução a 75%

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ter sua resolução reduzida em 75% para 500×500 pixels. O resultado é visto na Figura 4.23, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

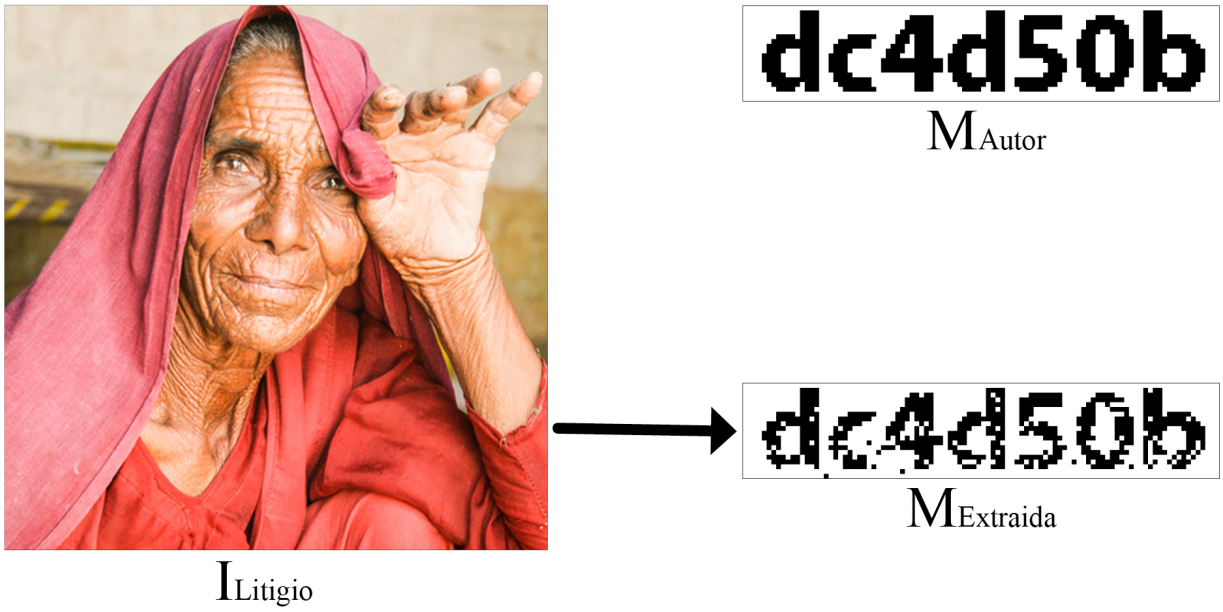

Figura 4.23: *Ataque: redução da resolução de $I_{Marcada}$ em 75%. Mostramos a $I_{Litigio}$ com* 500×500 *pixels e a respectiva $M_{Extraida}$*

Análise do resultado deste ataque:

- A degradação visual de $I_{Marcada}$ para $I_{Litigio}$ é levemente perceptível
- Embora $I_{Litigio}$ tenha perdido 75% da informação contida em $I_{Marcada}$, a Marca d'Água obtida $M_{Extraida}$ sofreu alguma degradação em relação à M_{Autor} porém permanece legível

4.5.1.3 Redução de resolução a 99%

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ter sua resolução reduzida em 99% para 100×100 pixels. O resultado é visto na Figura 4.24, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

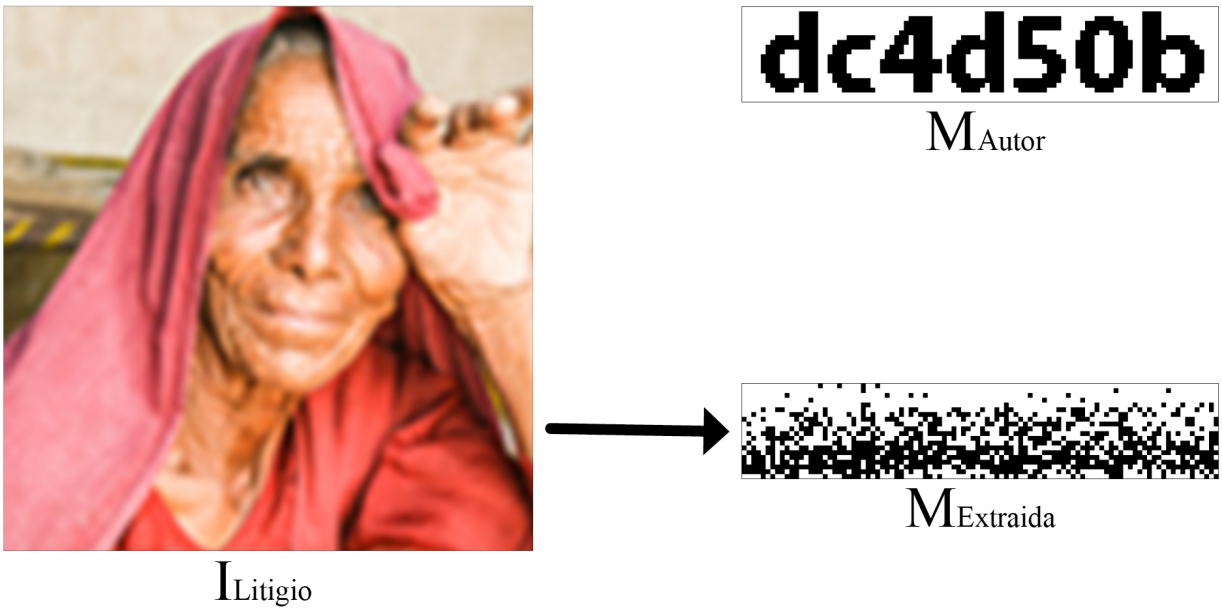

Figura 4.24: *Ataque: redução da resolução de $I_{Marcada}$ em 99%. Mostramos a $I_{Litigio}$ com* 100×100 *pixels e a respectiva $M_{Extraida}$*

- $I_{Litigio}$ perdeu 99% da informação contida em $I_{Marcada}$ e a Marca d'Água obtida $M_{Extraida}$ é completamente ilegível
- Embora o ataque tenha prejudicado seriamente a $M_{Extraida}$, a $I_{Litigio}$ está bastante degradada e não possui qualquer *valor comercial*

4.5.1.4 Interpolação a 400%

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ter sua resolução aumentada em 400% para 2.000×2.000 pixels. O resultado é visto na Figura 4.25, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- Não há degradação visual perceptível de $I_{Marcada}$ para $I_{Litigio}$
- Embora $I_{Litigio}$ tenha sido interpolada em 400% a partir de $I_{Marcada}$, a Marca d'Água obtida $M_{Extraida}$ sofreu pequena degradação em relação à M_{Autor}

4.5.1.5 Compactação JPEG com máxima qualidade

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ser compactada pelo JPEG com máxima qualidade. A compactação reduziu o tamanho do arquivo de 3 MB para 1,1 MB. O resultado é visto na Figura 4.26, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- Não há degradação visual perceptível de $I_{Marcada}$ para $I_{Litigio}$

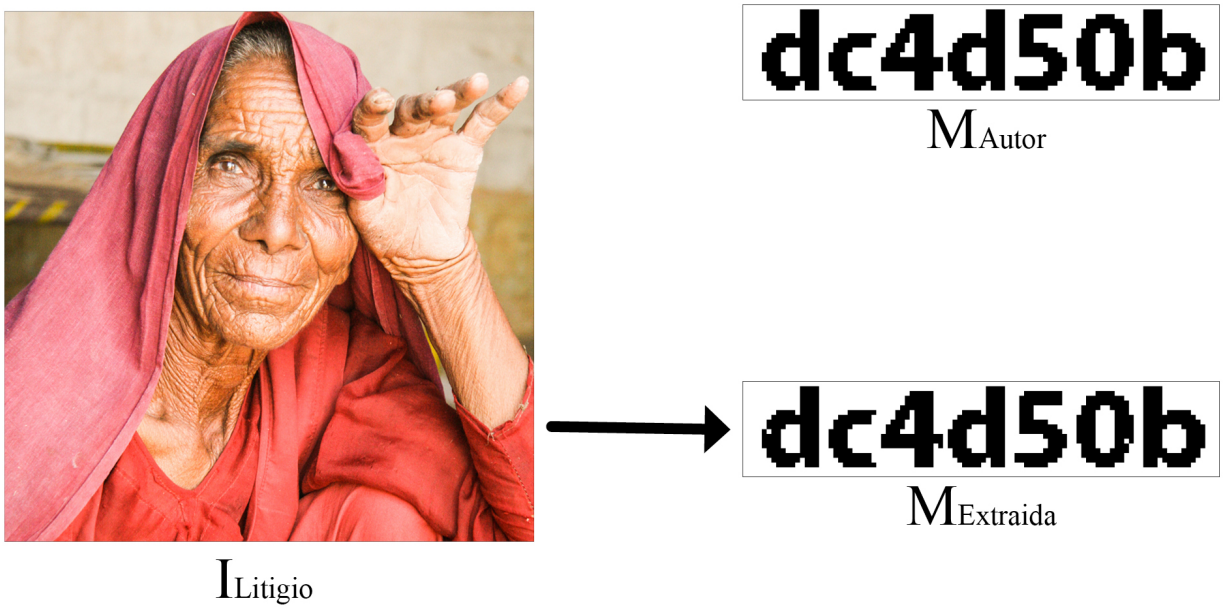

Figura 4.25: *Ataque: interpolação de $I_{Marcada}$ em 400%. Mostramos a $I_{Litigio}$ com* 2.000×2.000 *pixels e a respectiva $M_{Extraida}$*

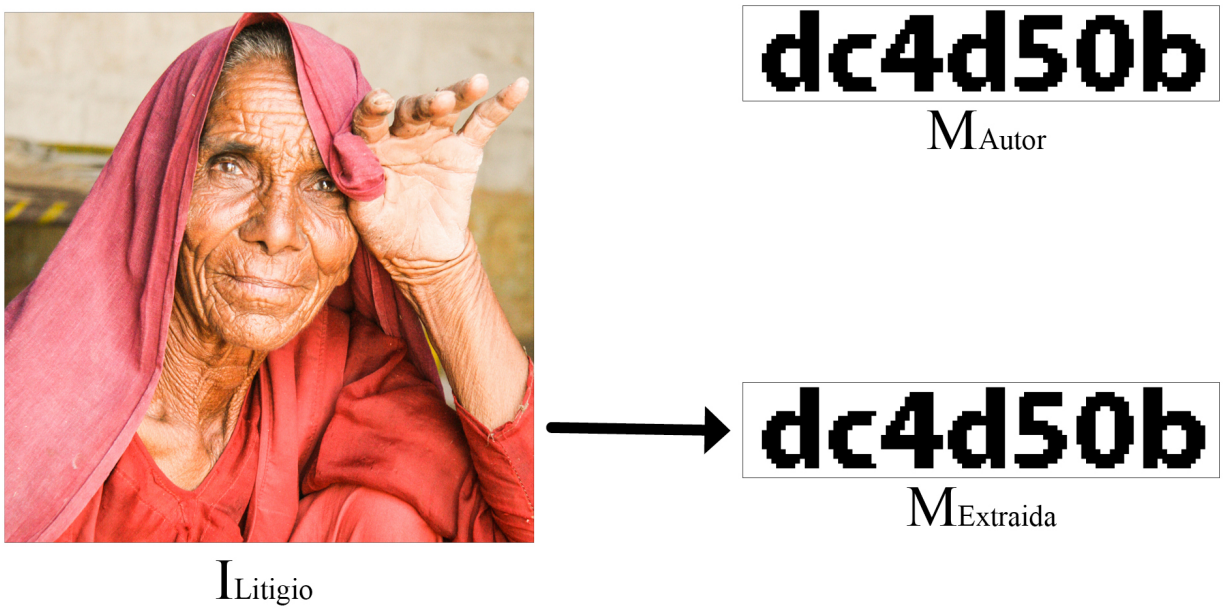

Figura 4.26: *Ataque: compactação JPEG com máxima qualidade. Mostramos a $I_{Litigio}$ com 1,1 MB e a respectiva $M_{Extraida}$*

- Embora o tamanho do arquivo $I_{Litigio}$ tenha sido reduzido em 63% em relação a $I_{Marcada}$, a Marca d'Água obtida $M_{Extraida}$ não mostra degradação em relação à M_{Autor}

4.5.1.6 Compactação JPEG com média qualidade

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ser compactada pelo JPEG com média qualidade. A compactação reduziu o tamanho do arquivo de 3 MB para 0,4 MB. O resultado é visto na Figura 4.27, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

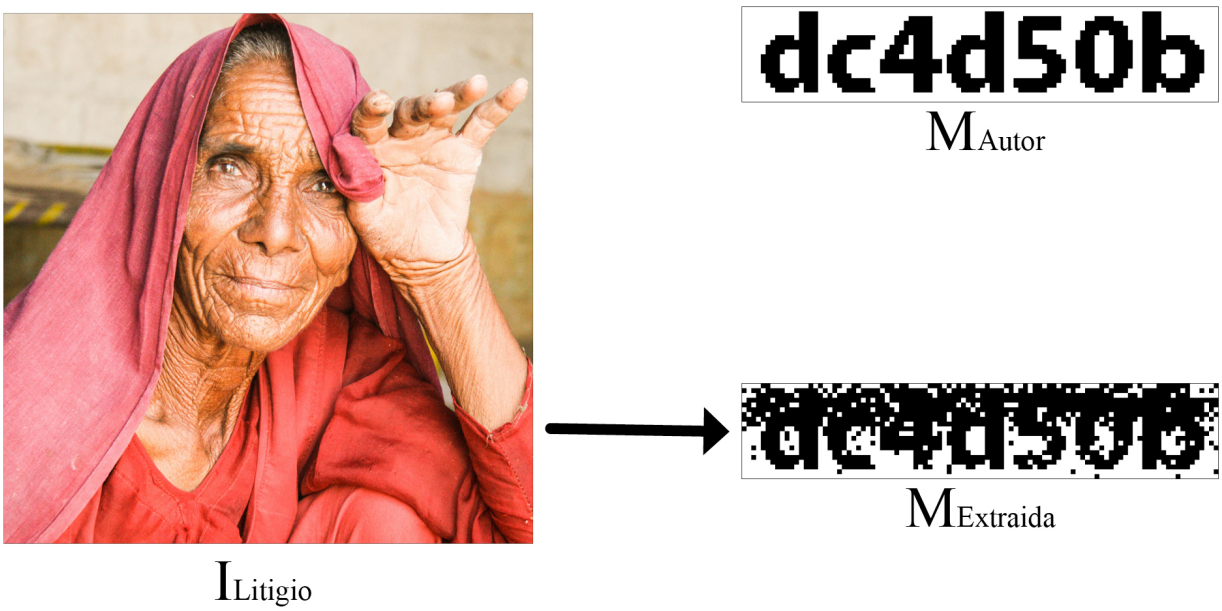

Figura 4.27: *Ataque: compactação JPEG com média qualidade. Mostramos a $I_{Litigio}$ com 0,4 MB e a respectiva $M_{Extraida}$*

Análise do resultado deste ataque:

- A degradação visual de $I_{Marcada}$ para $I_{Litigio}$ é levemente perceptível
- A redução do tamanho do arquivo $I_{Litigio}$ em 88% em relação a $I_{Marcada}$ degradou a Marca d'Água obtida $M_{Extraida}$ em relação à M_{Autor} embora traços de similaridade possam ser observados

4.5.1.7 Rotação 90º

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ser rotacionada em 90º. O resultado é visto na Figura 4.28, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- Não há degradação visual perceptível de $I_{Marcada}$ para $I_{Litigio}$
- Não há degradação visual perceptível de M_{Autor} para $M_{Extraida}$

4.5.1.8 Conversão para P&B

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ ao ser convertida para P&B. O resultado é visto na Figura 4.29, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- A alteração visual perceptível de $I_{Marcada}$ para $I_{Litigio}$ é somente a conversão para P&B

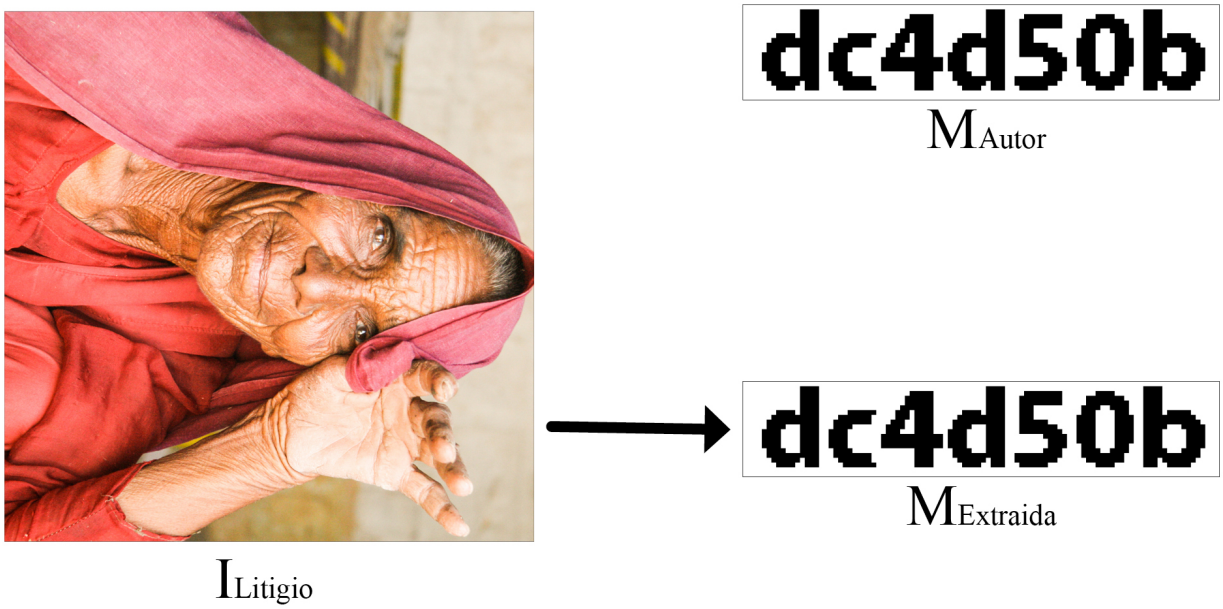

Figura 4.28: *Ataque: rotação 90º. Mostramos a $I_{Litigio}$ rotacionada em 90º e a respectiva $M_{Extraida}$*

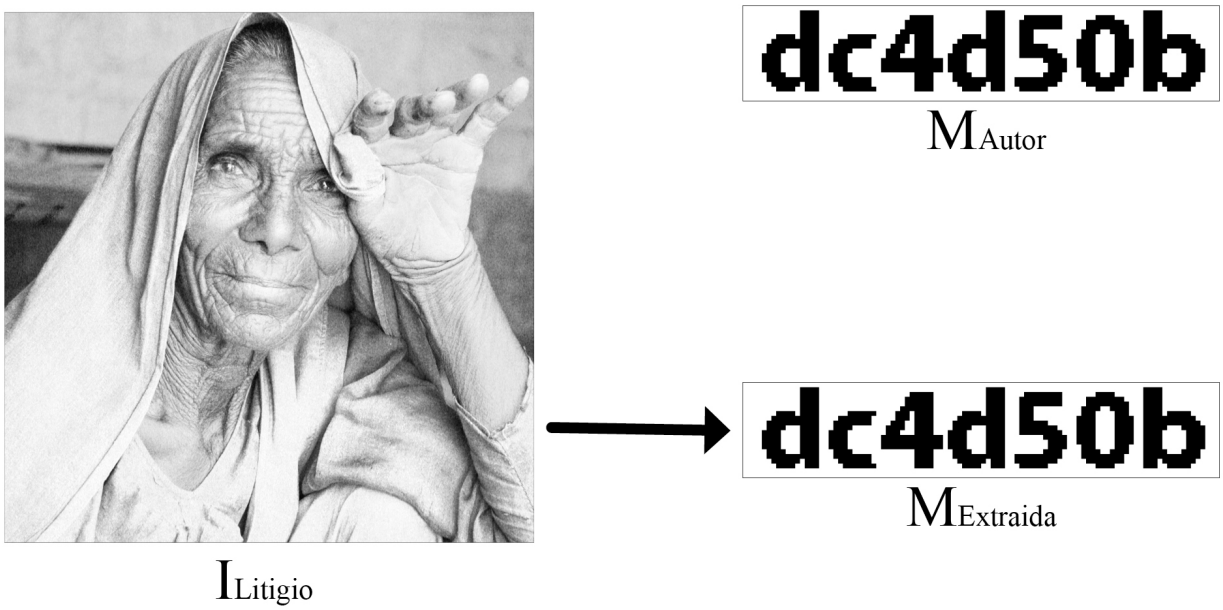

Figura 4.29: *Ataque: conversão para P&B. Mostramos a $I_{Litigio}$ rconvertida para P&B e a respectiva $M_{Extraida}$*

- Não há degradação visual perceptível de M_{Autor} para $M_{Extraida}$[2]

4.5.1.9 Deformação horizontal

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ de 1.000×500 pixels ao ser deformada na horizontal. O resultado é visto na Figura 4.30, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

[2]Nesta implementação consideramos somente os coeficientes DCT do canal Red em imagens RGB. No ataque de conversão para P&B mostrado na Figura 4.29 utilizou-se, favoravelmente, somente o canal Red; no caso de conversões para P&B nos canais Green ou Blue encontramos grande degradação de $M_{Extraida}$

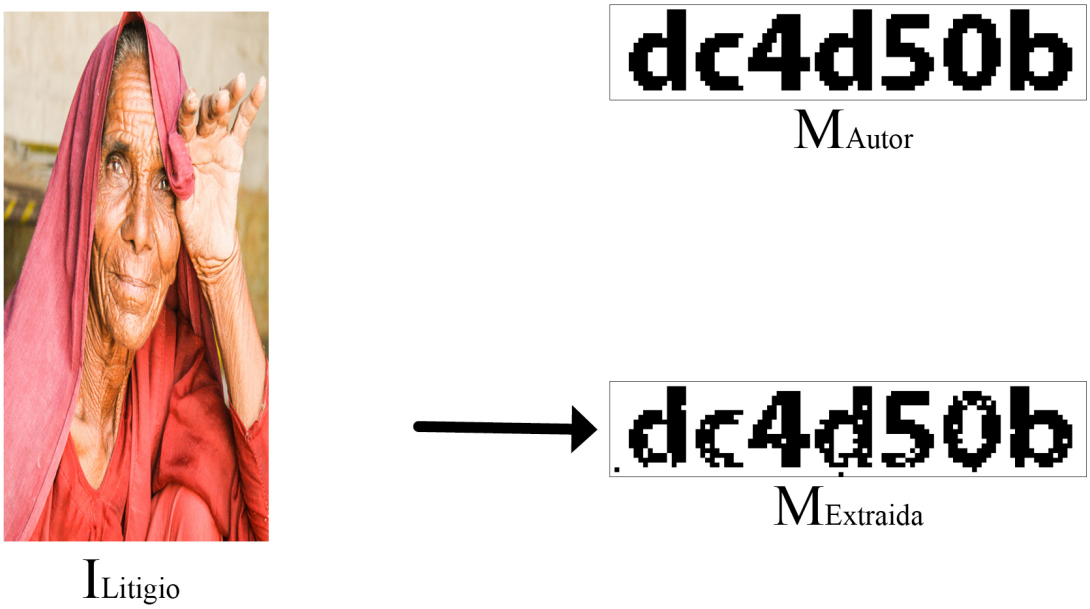

Figura 4.30: *Ataque: deformação horizontal. Mostramos a $I_{Litigio}$ com* 1.000×500 *pixels deformada na horizontal e a respectiva $M_{Extraida}$*

Análise do resultado deste ataque:

- Há degradação visual perceptível de $I_{Marcada}$ para $I_{Litigio}$
- Há degradação visual perceptível de M_{Autor} para $M_{Extraida}$ porém esta permanece legível

4.5.1.10 Deformação vertical

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ de 200×1.000 pixels ao ser deformada na vertical. O resultado é visto na Figura 4.31, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- Há grande degradação visual de $I_{Marcada}$ para $I_{Litigio}$
- Há degradação visual de M_{Autor} para $M_{Extraida}$ embora traços de similaridade possam ser observados

4.5.1.11 Filtro cor automática do Photoshop

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ de 1.000×1.000 pixels ao ser aplicado o filtro cor automática do Photoshop. O resultado é visto na Figura 4.32, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

- Há grande similaridade entre $I_{Marcada}$ e $I_{Litigio}$, embora as cores estejam perceptivelmente alteradas

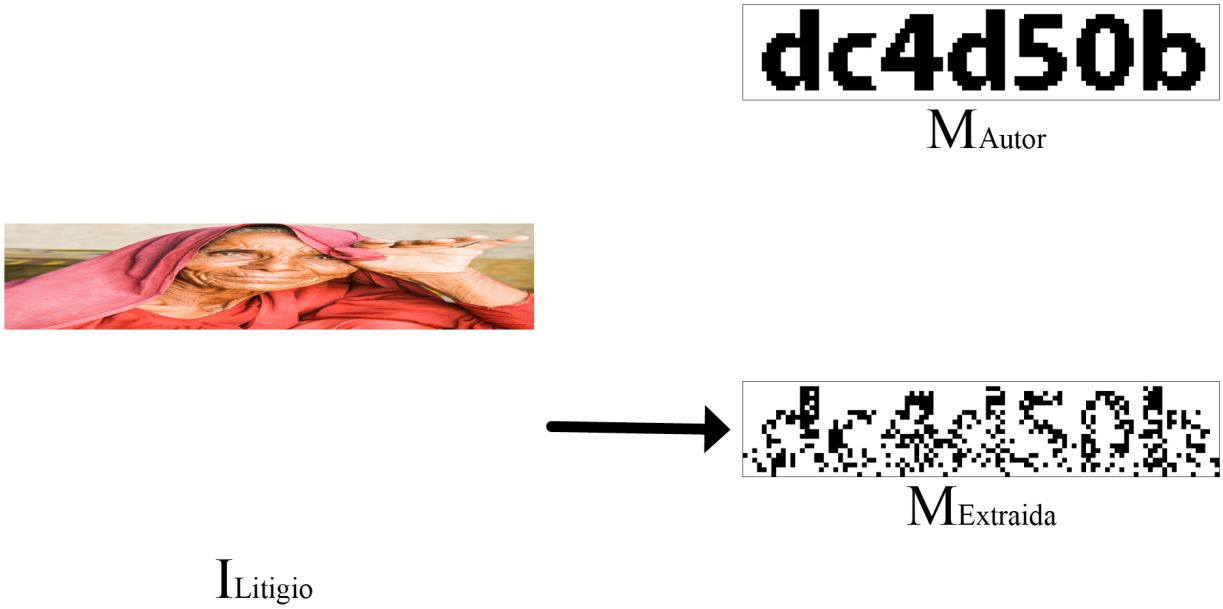

Figura 4.31: *Ataque: deformação vertical. Mostramos a $I_{Litigio}$ com 200×1.000 pixels deformada na vertical e a respectiva $M_{Extraida}$*

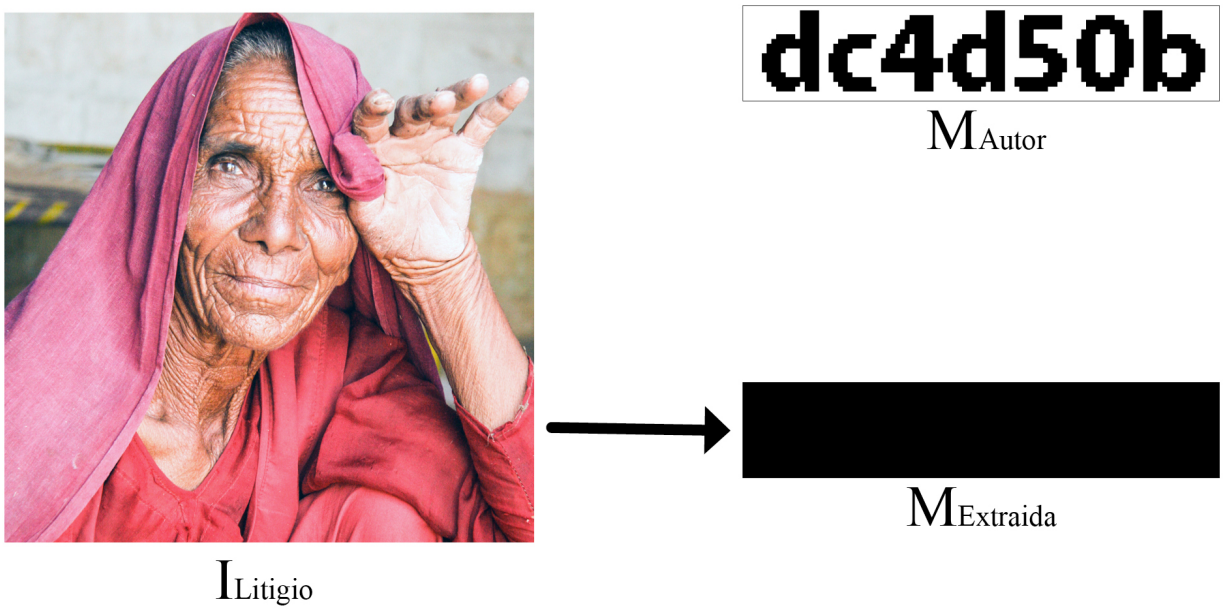

Figura 4.32: *Ataque: filtro cor automática do Photoshop. Mostramos a $I_{Litigio}$ com 1.000×1.000 pixels transformada pelo filtro cor automática do Photoshop e a respectiva $M_{Extraida}$*

- Há completa degradação de M_{Autor} para $M_{Extraida}$

4.5.1.12 Corte

Neste ataque a $I_{Marcada}$ de 1.000×1.000 pixels foi transformada em $I_{Litigio}$ de 1.000×500 pixels ao ser aplicado o corte em 50% da sua área. O resultado é visto na Figura 4.33, sendo $M_{Extraida}$ a Marca d'Água contida em $I_{Litigio}$.

Análise do resultado deste ataque:

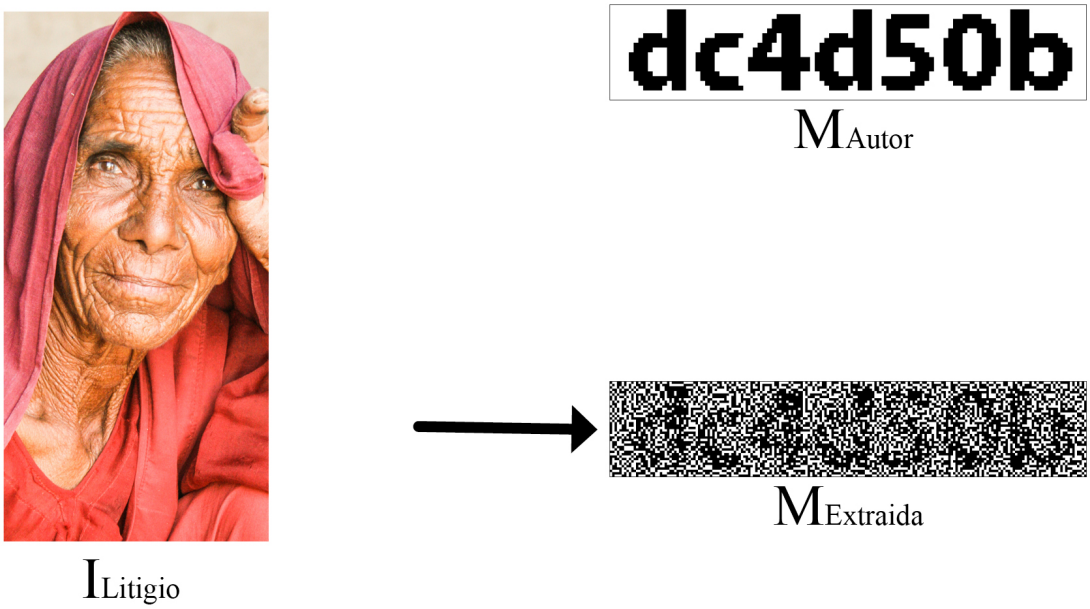

Figura 4.33: *Ataque: corte. Mostramos a* $I_{Litigio}$ *com* 1.000×500 *pixels obtida pelo corte de 50% da área de* $I_{Marcada}$ *e a respectiva* $M_{Extraida}$

- A área mais importante de $I_{Marcada}$, que é o rosto da *Senhora Indiana*, é claramente visível em $I_{Litigio}$
- Mesmo com o corte de 50% da área de $I_{Marcada}$ há degradação visual de M_{Autor} para $M_{Extraida}$, porém traços de similaridade possam ser observados

4.5.1.13 Resumo dos ataques à camada de transporte do MACV

A Tabela 4.1 apresenta o resumo dos resultados obtidos nos experimentos acima considerando os ataques realizados a $I_{Marcada}$. As colunas *Ataque* e *Descrição* identificam o experimento; a coluna *Resolução de* $I_{Litigio}$ mostra a resolução resultante do ataque; a coluna *Robusta* indica se a Marca D'Água foi resistente ao ataque; a coluna *Recomendação* mostra nossa avaliação de segurança, supondo que uma Marca D'Água parcialmente robusta tem elementos visuais suficientes para a prova de autoria.

Os experimentos foram realizados com ataques individuais, embora todos possam ser combinados entre si com esperada maior degradação tanto da imagem resultante como da Marca d'Água encapsulada.

Outras transformações poderiam ter sido utilizadas, mas o conjunto escolhido é abrangente o suficiente para verificar que a camada de *transporte* do MACV é robusta em determinadas situações porém não é imune a todos os ataques. Como vimos anteriormente, não temos notícia de algum método de Marca d'Água que seja seguro contra todos os ataques.

Ataque	Descrição	Resolução de $I_{Litigio}$	Robusta	Recomendação
1	Redução de resolução a 36%	800×800	Sim	Seguro
2	Redução de resolução a 75%	500×500	Parcial	Seguro
3	Redução de resolução a 99%	100×100	Não	Inseguro
4	Interpolação a 400%	2.000×2.000	Sim	Seguro
5	JPEG com máxima qualidade	1.000×1.000	Sim	Seguro
6	JPEG com média qualidade	1.000×1.000	Parcial	Seguro
7	Rotação 90^0	1.000×1.000	Sim	Seguro
8	Conversão para P&B	1.000×1.000	Sim	Seguro
9	Deformação horizontal	1.000×500	Sim	Seguro
10	Deformação vertical	200×1.000	Parcial	Seguro
11	Cor automática do Photoshop	1.000×1.000	Não	Inseguro
12	Corte	1.000×500	Parcial	Seguro

Tabela 4.1: *Resumo dos ataques realizados a* $I_{Marcada}$

4.5.2 Segurança da camada de *conteúdo* do MACV

Como vimos anteriormente, não há como impedir a cópia/ adulteração/ plágio de uma imagem estática sem autorização. A tecnologia atual de busca na internet não oferece recursos eficazes para encontrar as imagens resultantes de tais malfeitos. Desta forma, ao autor verdadeiro da imagem resta somente a opção de defender seus direitos numa eventual disputa contra o malfeitor. E nessa disputa o grande problema é: como provar sua autoria? Usando a fotografia como exemplo, em tempos passados o negativo do filme poderia ser usado como prova. Hoje esta possibilidade é raramente disponível.

O MACV, assim como outros métodos de Marca d'Água, tem como objetivo prover ferramentas para prova de autoria. A camada de *conteúdo* é definida para este fim, e a camada de *transporte* é usada para proteger tal conteúdo contra eventuais ataques. A segurança fornecida pela camada de *conteúdo* do MACV tem dois pilares:

1. Realizar um cálculo sobre a imagem $I_{Original}$ com as seguintes características:
   - Utiliza um algoritmo de *domínio público* sem a necessidade de chave
   - *Replicável* no ambiente de disputa (no tribunal, por exemplo)
   - Alta *entropia*
   - Baixa probabilidade de *colisão*: imagens com pelo menos 1 bit de diferença apresentam resultados completamente diferentes
   - *Computacionalmente difícil* ao atacante reproduzir tal cálculo sem a imagem $I_{Original}$
2. Criptografar o resultado do cálculo acima com as seguintes características:
   - Utiliza um algoritmo de *domínio público* sem a necessidade de chave

- Apresenta *segurança perfeita*

A aplicação do SHA-256 à $I_{Original}$ é o cálculo adotado no pilar 1 do MACV. Embora o SHA-256 seja somente razoavelmente seguro como algoritmo, veremos a seguir que, para prova de autoria em um tribunal, ele é suficiente. No caso do pilar 2 do MACV foi adotado o $(2,2) - VCS$.

Vamos supor como premissa que autor verdadeiro produz $I_{Original}$ e a submete ao MACV obtendo $M_{Secreta}$ e $I_{Marcada}$; são mantidas em segredo $I_{Original}$ e $M_{Secreta}$ e distribuída somente $I_{Marcada}$. Vamos analisar as possíveis situações:

1. Suponha que o autor descubra que um malfeitor tenha utilizado uma cópia não autorizada de $I_{Marcada}$ (esta cópia será chamada $I_{Litigio}$). Caso se sinta prejudicado, o autor abre um processo judicial contra o malfeitor e apresenta $I_{Original}$ e $M_{Secreta}$. O malfeitor é intimado a provar sua autoria e apresenta $I_{Litigio}$. Aqui temos duas possibilidades:

   (a) Se $I_{Litigio}$ é igual a $I_{Marcada}$, como vimos em 4.4.1.3 e 4.4.2.3, então M_{Autor} é igual a M'_{Autor}, provando que a marca do autor verdadeiro está presente na imagem apresentada pelo malfeitor

   (b) Se $I_{Litigio}$ é diferente de $I_{Marcada}$ significa que houve alguma manipulação de $I_{Marcada}$. Dependendo do ataque realizado, como vimos em 4.5.1.13, M'_{Autor} poderá ser semelhante a M_{Autor}

2. Suponha que o falsário inverta a situação e afirme que seja o autor de $I_{Litigio}$. Ele tem que provar que sua Marca d'Água está em $I_{Original}$; porém, como $I_{Original}$ não é conhecida, é *computacionalmente difícil* ao atacante conseguir seu objetivo

3. Suponha que o falsário analise diversas imagens marcadas pelo autor verdadeiro. É *computacionalmente difícil* ao atacante encontrar um padrão de Marca d'Água do verdadeiro autor pois cada $I_{Original}$ tem seu cálculo SHA-256 protegido pelo $(2,2) - VCS$

Concluímos que a camada de *conteúdo*, caso não tenha sido corrompida, prova *sem ambiguidade* a autoria da imagem pelo verdadeiro autor.

4.6 Análise das propriedades do MACV

Na Seção 3.3 apresentamos um conjunto de propriedades ideais para um sistema de Marca d'Água. Vamos analisar o MACV sob cada uma destas propriedades:

- Armazenada na própria imagem. Por construção, no MACV a imagem marcada $I_{Marcada}$ carrega a informação de segurança $M_{Publica}$ de forma intrínseca. Desta forma, no MACV não existe a necessidade de uma EIC (entidade independente de certificação) para registro de $I_{Original}$, $M_{Publica}$ e/ ou $I_{Marcada}$

- Visualmente imperceptível quando inserida. Esta propriedade é herdada do método de Cox, do qual utilizamos integralmente a camada de *transporte*. Adicionalmente, o fator de escala α utilizado nos experimentos do MACV é somente 10% do utilizado por Cox, reduzindo o impacto visual nos maiores coeficientes DCT da $I_{Marcada}$. Com isso minimizamos o risco de perda de *valor comercial* de $I_{Marcada}$ causado pela inserção da Marca d'Água $M_{Publica}$

- Visualmente significativa quando extraída. A Marca d'Água $M_{Extraida}$ extraída de $I_{Marcada}$ somente revela o segredo M_{Autor} quando sobreposta a $M_{Secreta}$

- Irreproduzível por terceiros não autorizados. Supondo como premissa a execução das práticas de segurança descritas na Seção 3.2.3, impedindo que o atacante obtenha $I_{Original}$, somente o autor verdadeiro consegue encapsular a Marca d'Água correta na imagem original

- Capaz de localizar alterações maliciosas na imagem hospedeira com resolução suficiente. Esta propriedade não se aplica ao nosso contexto de proteção dos direitos autorais e foi desconsiderada

- Publicamente verificável. O método MACV utiliza ferramentas de conhecimento público (SHA, Cox, $(2,2) - VCS$) que podem ser reproduzidas por ambas as partes litigantes em um processo de disputa de direitos autorais

- Indelével por manipulação não autorizada. Vimos na Seção 4.5.1.13 que a camada de *transporte* do MACV não é totalmente robusta

- Resistente a certas operações de processamento de imagem (como mudar o nível de compressão). Vimos na Seção 4.5.1.13 que a camada de *transporte* do MACV não é totalmente robusta

- Aplicável a formatos com e sem perdas, e a imagens binárias, em níveis de cinza e coloridas. O método MACV é aplicável a todas estas situações

- Eficiente em tempo de processamento e espaço de armazenamento. Tempo de processamento não é fator crítico no MACV.g e no MACV.e executados no ambiente protegido do autor. O MACV.v executado no tribunal é aplicado, geralmente, a uma única imagem, não sendo representativo o tempo total de execução. Em termos de espaço de armazenamento, o autor deve manter as cópias de todas as suas $M_{Secretas}$ e $M_{Publicas}$, que são muito menores do que suas respectivas imagens originais (cerca de 4% do tamanho)

- Sem ambiguidade. Se a camada de *transporte* do MACV for robusta então $M_{Extraida}$ (camada de *conteúdo* extraída de $I_{Litigio}$), sobreposta à respectiva $M_{Secreta}$, revela o cálculo SHA-256 sobre $I_{Original}$

- Imperceptível por segurança perfeita. Esta propriedade é obtida pela aplicação do $(2,2) - VCS$ no MACV.e
- Alta entropia. Esta propriedade é obtida pela aplicação do SHA-256 à $I_{Original}$ no MACV.e
- Capacidade de absorção da Marca d'Água. Nesta implementação a capacidade de absorção de uma Marca d'Água MACV é $0,67$ milhões de pixels

4.7 Soluções baseadas em Criptografia Visual e Marca d'Água

O MACV não é o método pioneiro no uso combinado de Criptografia Visual e Marca d'Água para proteção dos direitos autorais de imagens estáticas. No entanto, uma análise de soluções similares revela que o MACV apresenta um conjunto de propriedades únicas até onde nossa pesquisa alcançou. Ora a propriedade não atendida em outros métodos é o *armazenamento na própria imagem*; ora é a *sem ambiguidade*; outras vezes falta *alta entropia*. A seguir apresentamos um breve resumo destas soluções.

A solução apresentada por HASSAN e KHALILI [HK05] usa uma seleção específica de pixels da imagem original como uma informação de verificação. Esta informação é enviada para uma EIC (entidade independente de certificação). Embora apresente a vantagem de não alterar a imagem original, uma EIC é necessária para a verificação de autoria. A técnica apresentada por SLEIT e ABUSITTA [SA08] tem semelhanças com a de HASSAN e KHALILI e a segurança também é baseada na existência de uma EIC. Na solução de ABUSITTA [Abu12] a seleção de pixels não precisa ser específica; como em HASSAN e KHALILI a informação de verificação é registrada em uma EIC.

Em LOE, TSO e LIU [LTL07] encontramos uma técnica que usa DWT para gerar uma chave secreta a partir da imagem original. Porém esta solução também é baseada na presença de uma EIC. SUREKHA, SWAMY, RAO e KUMAR [SSRK09] definem um *Master Share*[3] e o registram em uma EIC. A solução proposta por NAG, SINGH, BISWAS, SARKAR e SARKAR [NSB+11] cria um *Master Share* a partir dos bits mais significativos da imagem original e o mantém em uma EIC.

A proposta de HSU e HOU [HH05] também é dependente de uma EIC. Em artigos recentes encontramos: a técnica de AMIRI e MOGHADDAM [AM15] utiliza DWT e EIC; a solução proposta por DEVI, SINGH, ROY, CHANU e TUITHUNG [DSR+15] adota EIC; FATAHBEYGI e AKHLAGHIAN [FA15] também dependem de EIC.

O método de TIFEDJADJINE, ATAMNA, DIBI e BOURIDANR [TADB05] é interessante pois não usa EIC e não necessita da imagem original no processo de verificação. No

[3]No restante desta dissertação usamos *transparência* como a melhor representação em português para *share*

entanto ele não apresenta a desejável propriedade *sem ambiguidade* pois a verdadeiro autor e o malfeitor podem apresentar, ao mesmo tempo, seus respectivos *shares* secretos para defender a autoria.

A proposta de JITHI e NAIR [JN13] utiliza Criptografia Visual e Marca d'Água porém sem o objetivo de proteção dos direitos autorais; a ideia de JITHI e NAIR é criar um método para transmitir os *shares* gerados pelo $(k, n) - VCS$ de forma segura. Em LUO, PAN e LU [LPL07] os *shares* gerados pelo $(2, 2) - VCS$ são alterados para transportar 2 Marcas d'Água no processo de comunicação. Outros métodos para transmitir informações sigilosas utilizando Criptografia Visual e Marca d'Água foram apresentados recentemente por KAUR e ATTRI [KA15], GAYATHRI e NAGARAJAN [GN15] e ANJIKAR e BAMBODKAR [AB15].

Assim como no MACV, a solução apresentada por TAI e CHANG [TC04] não utiliza EIC. Este método altera a imagem original pelo encapsulamento do *share* público gerado pelo $(2, 2) - VCS$. Porém a propriedade de alta entropia não é atendida por TAI e CHANG pois utiliza o logotipo do autor como chave. O mesmo se aplica às metodologias apresentadas por HOUMANSADR e SHAHROKH [HG05] e por HSU e TU [HT08].

O recente trabalho de HOU, QUAN e LIAO [HQL15] apresenta uma nova técnica de Criptografia Visual baseada em *shares* com significado que pode ser útil (em estudo futuro) como camada de *conteúdo*; porém a segurança do algoritmo de Marca d'Água não é explorada. A nova abordagem proposta por YAN, WANG, SONG e YANG [YWSY15] utiliza um método $(2, 2) - VCS$ estendido; porém não há prova de que esta Marca d'Água seja *Imperceptível por segurança perfeita*.

Uma das propostas mais criativas que encontramos é a de HOU e CHEN [HC00]. É semelhante aos métodos de TAI, HOUMANSADR e HSU porém não existe a necessidade de extração: a Marca d'Água é revelada simplesmente pela sobreposição da imagem marcada pelo *share* secreto. Diferente de todos os métodos citados aqui, HOU e CHEN não utilizam as transformadas no domínio da frequência; também não resolvem a propriedade *sem ambiguidade*. Mas inspiram um caminho interessante para novas propostas de uso combinado de Criptografia Visual e Marca d'Água para fins de proteção dos direitos autorais de imagem.

Como vimos acima, o MACV não é o único método a usar Criptografia Visual combinada com Marca d'Água para proteção dos direitos autorais de imagens estáticas. Mas apresenta um conjunto relevante de propriedades que o torna original em relação aos demais trabalhos observados.

4.8 Conclusões do capítulo

Apresentamos um novo método de Marca d'Água, denominado MACV, no qual a camada de *transporte* segue o modelo de Cox e a camada de *conteúdo* apresenta como novidade uma informação com significado , de alta entropia e com segurança perfeita.

O funcionamento do MACV foi descrito por meio de seus três algoritmos: o MACV.g

- Gerador de Marca d'Água, o MACV.e - Encapsulador de Marca d'Água e o MACV.v - Verificador de Marca d'Água.

Mostramos alguns experimentos com o MACV e avaliamos sua segurança. A camada de *transporte* é razoavelmente robusta contra ataques e pode ser incrementada para maior segurança - não temos notícia de uma solução que seja totalmente segura; a camada de *conteúdo*, caso não tenha sido corrompida, prova *sem ambiguidade* a autoria da imagem pelo verdadeiro autor em um ambiente de disputa.

Revisamos a bibliografia de soluções baseadas em Criptografia Visual e Marca d'Água e percebemos que o MACV apresenta propriedades únicas, especialmente em relação ao *armazenamento na própria imagem*, a prova de autoria *sem ambiguidade* e informação de *alta entropia*.

Capítulo 5

Conclusões

5.1 Considerações Finais

A tecnologia disponível atualmente não impede que uma imagem estática em meio digital seja *copiada*, *adulterada* ou *plagiada* sem autorização. Desta forma, ações maliciosas em uma imagem estática podem ocorrer sem qualquer tipo de controle pelo seu verdadeiro autor. Numa eventual disputa judicial o autor tem poucos recursos para provar seu direito de autoria. Na época da fotografia analógica o negativo poderia ser usado como prova; porém, atualmente, o uso de negativo pelos autores é praticamente nulo.

Neste contexto, para fornecer prova de autoria de uma imagem estática propomos um conjunto de algoritmos denominado MACV. O MACV permite que o autor *marque* suas imagens estáticas no momento de criação. A imagem $I_{Original}$ é marcada e somente sua versão $I_{Marcada}$ é divulgada. Os arquivos $I_{Original}$ e $M_{Secreta}$ devem ser armazenados de forma segura pelo autor e poderão ser utilizados como prova de autoria numa eventual disputa.

A $I_{Marcada}$ é passível de *ataques*, maliciosos ou inconscientes, assim que disponibilizada em ambiente desprotegido. A imagem resultante destes *ataques* é denominada $I_{Litigio}$. O MACV oferece recursos para extrair a Marca d'Água $M_{Extraida}$ de $I_{Litigio}$. Desde que $M_{Extraida}$ não esteja seriamente comprometida, a sobreposição de $M_{Secreta}$ e $M_{Extraida}$ revelará o resultado da função de *hashing* SHA aplicado sobre $I_{Original}$. Somente o autor verdadeiro poderia ter realizado este cálculo a partir de $I_{Original}$.

Na definição do MACV adotamos duas camadas independentes: *transporte* e *conteúdo*. A camada *conteúdo* é de alta entropia e apresenta informações relevantes para a prova de autoria da imagem. Como vimos nos experimentos realizados, a camada de *transporte* é razoavelmente robusta para proteger a integridade da camada de *conteúdo* mesmo após os possíveis ataques à $I_{Marcada}$. Se a camada *transporte* falha em seu objetivo, a camada *conteúdo* possivelmente será corrompida e poderá se tornar inútil como prova de autoria.

Nossa pesquisa não encontrou qualquer camada de *transporte* que seja segura contra todos os ataques possíveis, embora este seja um campo de pesquisa atual e em franca evolução. A modularidade proporcionada pelas camadas independentes de *transporte* e *conteúdo*

permitem que o MACV possa ser beneficiado pela evolução de qualquer uma destas duas frentes.

Observamos que o MACV apresenta as seguintes propriedades desejáveis de Marca d'Água: armazenada na própria imagem; visualmente imperceptível quando inserida; visualmente significativa quando extraída; irreproduzível por terceiros não autorizados; publicamente verificável; indelével por manipulação não autorizada; resistente a certas operações de processamento de imagem; aplicável a formatos com e sem perdas, e a imagens binárias, em níveis de cinza e coloridas; eficiente em tempo de processamento e espaço de armazenamento; sem ambiguidade; imperceptível por segurança perfeita; alta entropia. A vulnerabilidade do MACV aparece na propriedade *resistente a certas operações de processamento de imagem*; vimos nos experimentos realizados que essa resistência é baixa em determinadas operações.

O MACV não é o único método a usar Criptografia Visual combinada com Marca d'Água para proteção dos direitos autorais de imagens estáticas. Mas apresenta um conjunto relevante de propriedades que o torna original em relação aos demais trabalhos observados. Em particular, ressaltamos como diferenciais do MACV as propriedades *Imperceptível por segurança perfeita*, *Alta entropia*, *Armazenamento na própria imagem* e *Sem ambiguidade*.

5.2 Sugestões para Pesquisas Futuras

Podemos elencar algumas frentes de pesquisa futura para evolução do MACV ou criação de novos métodos de Marca d'Água:

1. Encontrar uma camada de *transporte* resistente ao maior número de ataques por processamento de imagem (Seção 4.5)

2. Adotar algum tipo de redundância da camada de *conteúdo* para recuperação de coeficientes eliminados/ adulterados (Seção 4.3.2)

3. Encontrar uma solução que minimize a degradação da Marca d'Água em um ataque de marcações sucessivas (Seção 4.5)

4. Encontrar um método que atenda simultaneamente aos objetivos de proteção dos direitos autorais e de garantia de integridade (Seção 3.2.1)

5. Comparar resultados entre a transformada DWT e a DCT (Seção 4.3.2)

6. No caso da DCT, utilizar algoritmo mais eficiente para ordenação de coeficientes

7. Adotar uma versão do $(2,2) - VCS$ sem expansão de pixels (Seção 4.4)

8. Avaliar a viabilidade de um método onde a Marca d'Água é revelada simplesmente pela sobreposição da imagem marcada pelo *share* secreto (Seção 4.7)

9. Na área de de Armazenamento e Recuperação de Informação, desenvolver um método de busca de imagens marcadas

10. Em aplicações comerciais, convencer fabricantes de equipamento fotográfico e/ ou desenvolvedores de software de imagem a utilizar uma solução para proteção de direitos autorais

Referências Bibliográficas

[AB15] Akhil Anjikar e Rahul Bambodkar. Visual cryptography with color error diffusion and digital watermarking. *channels*, 4(11), 2015. 68

[Abu12] Adel Hammad Abusitta. A visual cryptography based digital image copyright protection. *Journal of Information Security*, 3(2):96, 2012. 67

[AM15] Tayebe Amiri e Mohsen Ebrahimi Moghaddam. A new visual cryptography based watermarking scheme using dwt and sift for multiple cover images. *Multimedia Tools and Applications*, páginas 1–17, 2015. 67

[Bar03] Paulo Sérgio LM Barreto. *Criptografia Robusta e Marcas d'água Frágeis: Construção e Análise de Algoritmos para Localizar Alterações em Imagens Digitais*. Tese de Doutorado, Universidade de São Paulo, 2003. 24, 28, 30

[BDDSS03] Carlo Blundo, Paolo D'Arco, Alfredo De Santis e Douglas R Stinson. Contrast optimal threshold visual cryptography schemes. *SIAM Journal on Discrete Mathematics*, 16(2):224–261, 2003. 15

[BQ95] M BENOfT e Jean-Jacques Quisquater. Cryptology for digital tv broadcasting. *Proceedings of the IEEE*, 83(6), 1995. 23

[CG98] Ken Cabeen e Peter Gent. Image compression and the discrete cosine transform. *College of the Redwoods*, 1998. 32

[CKLS97] Ingemar J Cox, Joe Kilian, F Thomson Leighton e Talal Shamoon. Secure spread spectrum watermarking for multimedia. *Image Processing, IEEE Transactions on*, 6(12):1673–1687, 1997. 21, 30, 31, 34, 35, 36, 38

[DSR+15] B Pushpa Devi, Kh Manglem Singh, Sudipta Roy, Y Jina Chanu e T Tuithung. A watermarking scheme for digital images based on visual cryptography. *International Journal of Contemporary Mathematical Sciences*, 8(32), 2015. 67

[DX13] Fabiano de Moraes Domingues e José Antônio Moreira Xexéo. Prevenção contra fraude em criptografia visual com base na autenticação de múltiplas imagens em transparências circulares. Dissertação de Mestrado, Instituto Militar de Engenharia, 2013. 17

[Exa15] Revista Exame. Mais de 4 bilhões de pessoas no mundo não têm acesso à internet. 2015. http://exame.abril.com.br/tecnologia/noticias/mais-de-4-bilhoes-de-pessoas-no-mundo-nao-tem-acesso-a-internet. 22

[FA15] Ali Fatahbeygi e Fardin Akhlaghian. A new robust semi-blind image watermarking based on block classification and visual cryptography. Em *Pattern*

Recognition and Image Analysis (IPRIA), 2015 2nd International Conference on, páginas 1–6. IEEE, 2015. 67

[Fau93] Olivier Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT press, 1993. 35

[For12] Forbes. The steady rise of fine art photography. 2012. http://www.forbes.com/sites/kathryntully/2012/12/09/the-steady-rise-of-fine-art-photography/. 26

[G115] G1. Logomarca das olimpíadas de 2020 é cancelada após acusação de plágio. 2015. http://g1.globo.com/bom-dia-brasil/noticia/2015/09/logomarca-das-olimpiadas-de-2020-e-cancelada-apos-acusacao-de-plagio.html. 29

[GN15] R Gayathri e V Nagarajan. Secure data hiding using steganographic technique with visual cryptography and watermarking scheme. Em *Communications and Signal Processing (ICCSP), 2015 International Conference on*, páginas 0118–0123. IEEE, 2015. 68

[HC00] Young-Chang Hou e Pei-Min Chen. An asymmetric watermarking scheme based on visual cryptography. Em *Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on*, volume 2, páginas 992–995. IEEE, 2000. 68

[HCT06] Gwoboa Horng, Tzungher Chen e Du-Shiau Tsai. Cheating in visual cryptography. *Designs, Codes and Cryptography*, 38(2):219–236, 2006. 17, 19

[HG05] Amir Houmansadr e Shahrokh Ghaemmaghami. A digital image watermarking scheme based on visual cryptography. Em *International Symposium on Telecommunications*, páginas 1–5, 2005. 68

[HH05] Ching-Sheng Hsu e Young-Chang Hou. Copyright protection scheme for digital images using visual cryptography and sampling methods. *Optical Engineering*, 44(7):077003–077003, 2005. 67

[HK05] Mahmoud A Hassan e Mohammed A Khalili. Self watermarking based on visual cryptography. Em *Proceedings of World Academy of Science, Engineering and Technology*, volume 8, páginas 159–162. Citeseer, 2005. 67

[HQL15] Young-Chang Hou, Zen-Yu Quan e Hsin-Yin Liao. New designs for friendly visual cryptography scheme. *International Journal of Information and Electronics Engineering*, 5(1):15, 2015. 68

[HT08] Ching-Sheng Hsu e Shu-Fen Tu. Digital watermarking scheme with visual cryptography. *IMECS, Hong Kong (March 2008)*, 2008. 68

[Igu07] Fabio Iguchi. *Inserção de marcas d'água digitais usando recorrência de padrões multiescalas*. Tese de Doutorado, Universidade Federal do Rio de Janeiro, 2007. 26, 29

[JN13] PV Jithi e Anitha T Nair. Progressive visual cryptography with watermarking for meaningful shares. Em *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on*, páginas 394–401. IEEE, 2013. 68

[KA15] Manjinder Kaur e Varinder Kaur Attri. Implementation of steganographic method based on interpolation and lsb substitution of digital images with watermarking and visual cryptography. *International Journal of Computer Applications*, 121(21), 2015. 68

[KP99] Martin Kutter e Fabien AP Petitcolas. Fair benchmark for image watermarking systems. Em *Electronic Imaging'99*, páginas 226–239. International Society for Optics and Photonics, 1999. 25

[KP00] Stefan Katzenbeisser e Fabien AP Petitcolas. Information hiding. *Boston, London, Artech House*, 2000. 26

[KT99] Hidenori Kuwakado e Hatsukazu Tanaka. Image size invariant visual cryptography. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 82(10):2172–2177, 1999. 10, 46

[Lop06] Ivan Oliveira Lopes. Marca d'água digital: uma técnica para verificação de autenticidade ou proteção de direitos autorais. Dissertação de Mestrado, Universidade Federal de Uberlândia, 2006. 23, 24

[LPL07] Hao Luo, Jeng-Shyang Pan e Zhe-Ming Lu. Hiding multiple watermarks in transparencies of visual cryptography. Em *Intelligent Information Hiding and Multimedia Signal Processing, 2007. IIHMSP 2007. Third International Conference on*, volume 1, páginas 303–306. IEEE, 2007. 68

[LTL07] Der-Chyuan Lou, Hao-Kuan Tso e Jiang-Lung Liu. A copyright protection scheme for digital images using visual cryptography technique. *Computer Standards & Interfaces*, 29(1):125–131, 2007. 67

[NP97] Moni Naor e Benny Pinkas. Visual authentication and identification. Em *Advances in Cryptology - CRYPTO'97*, páginas 322–336. Springer, 1997. 19

[NS95] Moni Naor e Adi Shamir. Visual cryptography. Em *Advances in Cryptology - EUROCRYPT'94*, páginas 1–12. Springer, 1995. 2, 5, 6, 7, 12, 15, 20

[NSB+11] Amitava Nag, Jyoti Prakash Singh, Sushanta Biswas, Dipankar Sarkar e Partha Pratim Sarkar. A novel copyright protection scheme using visual cryptography. Em *Advances in Computing and Communications*, páginas 612–619. Springer, 2011. 67

[ON99] Paulo Murilo Castro de Oliveira e M Nussenzeveig. Autômatos celulares. *Complexidade & Caos. Rio de Janeiro: Editora UFRJ/COPEA*, página 276, 1999. 13

[SA08] Azzam Sleit e Adel Abusitta. A visual cryptography based watermark technology for individual and group images. *Systemics, Cybernetics And Informatics*, 5(2):24–32, 2008. 67

[Sha49] Claude E Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28(4):656–715, 1949. 10

[Sha79] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979. 6, 11, 12

[SSRK09] B Surekha, Dr GN Swamy, Dr K Srinivasa Rao e A Ravi Kumar. A watermarking technique based on visual cryptography. *Journal of Information Assurance and Security*, 4(6):470–473, 2009. 67

[TADB05] Zohra Tifedjadjine, N Atamna, Zohir Dibi e Ahmed Bouridane. Halftone image watermarking based on visual cryptography. Em *Electronics, Circuits and Systems, 2005. ICECS 2005. 12th IEEE International Conference on*, páginas 1–4. IEEE, 2005. 67

[TC04] Gwo-Chin Tai e Long-Wen Chang. Visual cryptography for digital watermarking in still images. Em *Advances in Multimedia Information Processing-PCM 2004*, páginas 50–57. Springer, 2004. 68

[Ter00] Routo Terada. *Segurança de dados: criptografia em redes de computador*. Edgard Blucher, 2000. 11, 39, 41

[VSTO94] Ron G Van Schyndel, Andrew Z Tirkel e Charles F Osborne. A digital watermark. Em *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 2, páginas 86–90. IEEE, 1994. 30

[WY12] JP Weir e WeiQi Yan. *Visual cryptography and its applications*. Bookboon, 2012. 6, 11, 24, 30

[YWSY15] Bin Yan, Ya-Fei Wang, Ling-Yun Song e Hong-Mei Yang. Size-invariant extended visual cryptography with embedded watermark based on error diffusion. *Multimedia Tools and Applications*, páginas 1–24, 2015. 68