

中图分类号: TP391

单位代号: 10280

密 级:

学 号: 21721313

上海大学



硕士学位论文

SHANGHAI UNIVERSITY
MASTER'S DISSERTATION

题 目	抗屏幕拍摄的频域 文档水印技术研究
--------	----------------------

作 者: 张亚洲
学科专业: 信号与信息处理
导 师: 张新鹏教授
完成日期: 2024 年 5 月

姓 名：张亚洲

学号：21721313

论文题目：抗屏幕拍摄的频域文档水印技术研究

上海大学

本论文经答辩委员会全体委员审查, 确认符合上海大学硕士学位论文质量要求。

答辩委员会签名:

主任:

委员:

导 师:

答辩日期:

姓 名：张亚洲

学号：21721313

论文题目：抗屏幕拍摄的频域文档水印技术研究

原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名：_____日 期：_____

本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

（保密的论文在解密后应遵守此规定）

签 名：_____导师签名：_____日期：_____

上海大学工学硕士学位论文

抗屏幕拍摄的频域文档水印
技术研究

姓 名：张亚洲

导 师：张新鹏教授

学科专业：信号与信息处理

上海大学通信与信息工程学院

二〇二四年五月

A Dissertation Submitted to Shanghai University for the
Degree of Master in Engineering

Screen Shooting Resistant Robust Document Watermarking

Candidate: Yazhou Zhang

Supervisor: Pro. Xinpeng Zhang

Major: Signal and Information Processing

School of Communication and Information Engineering

Shanghai University

May, 2024

摘要

当电子文档，尤其是合同、机密文件和个人档案等敏感信息在显示器上展示时，它们可能在未经授权的情况下遭到偷拍。随着数字化办公和智能手机的发展，使用手机拍摄屏幕的行为不仅时有发生，而且由此引发的隐私侵权、恶意偷拍和非法分发等问题也日益严重。传统的数字水印技术在应对屏幕盗摄这种跨媒体信息传输时性能不佳，难以提供有效的版权保护。因此，为了能够有效抵御屏摄过程引入的复合失真，本文提出了基于 DFT 环形水印的抗屏摄文档底纹水印算法和抗屏摄文档水印的自动提取算法，具体如下：

(1) 针对目前抗屏摄文档水印对屏摄复合失真鲁棒性不足，并且难以平衡水印隐蔽性、文档可编辑性和鲁棒性的问题，本文提出了基于 DFT 环形水印的抗屏摄文档底纹水印算法。具体而言，本文结合 DFT 圆环水印的特性，使用小尺寸水印底纹模板拼接构成完整的水印底纹，将水印信息嵌入在文档底纹中，在保持视觉质量的同时也保留了文档的可编辑性。通过在底纹的中频区域嵌入水印环，可以有效抵御跨媒体传输信道中的各种劣化攻击，从而有效抵抗屏摄复合失真。实验结果表明，与相关方法相比本算法拥有更强的抗屏摄鲁棒性。

(2) 针对现有抗屏摄水印研究中普遍存在的手动定位与矫正问题，本文提出了抗屏摄文档水印自动提取算法，通过将水印提取过程拆分为自动定位、自动矫正和自动提取三个步骤，可以实现对盗摄照片的批量版权认证。具体而言，通过设计一种新型的环形水印模板，利用其频域的高能量特征实现自动定位，然后结合文档图像的文本行特征实现自动矫正，最后再结合水印模板完成水印信息的盲提取。本算法能够在仅拍摄部分文档内容的情况下，依然可以准确提取水印信息。实验结果表明，该算法不仅在复杂环境拍摄下拥有良好的抗屏摄鲁棒性，还可以抵抗常见的图像处理攻击，包括 JPEG 压缩、高斯模糊、中值滤波、尺度缩放、平移和裁剪等。

关键词：数字水印；版权保护；文档水印；屏摄失真；鲁棒盲水印

ABSTRACT

When electronic documents, especially sensitive information such as contracts, confidential documents, and personal records, are displayed on screens, they can be secretly photographed without authorization. With the rise of digital offices and smartphones, the practice of using mobile phones to take pictures of screens is becoming more frequent. This has led to increasingly serious issues such as privacy violations, malicious candid photography, and illegal distribution. Many traditional digital watermarking techniques are not able to provide adequate copyright protection against screen capture, which is a type of cross-media information transmission. To effectively address this issue, this thesis proposes a screen-shooting resistant document underlay watermarking based on DFT circular watermark template and an automatic document watermark extraction algorithm specifically designed for screen-shooting scenarios. The main contributions of this study are as follows:

(1) Addressing the current limitations of screen-shooting resistant document watermarking, which are insufficient in robustness against composite distortions from screen capture and struggle to balance watermark invisibility, document editability, and robustness, this thesis introduces a screen-shooting resistant document underlay watermarking based on DFT circular watermark template. By leveraging the properties of DFT circular watermarks and piecing together small-sized watermark templates to form a complete underlay, the watermark information is embedded within the document underlay. This approach maintains visual quality while preserving the document's editability. By embedding the watermark ring in the mid-frequency region of the underlay, this algorithm effectively resists various degradations in cross-media transmission channels, thereby effectively countering screen-shooting distortions. Experimental results demonstrate that the proposed algorithm has stronger robustness against screen-shooting attack compared to related methods.

(2) To address the prevalent issue of manual positioning and correction in existing screen-shooting resistant watermarking, this thesis presents an automatic document

watermark extraction algorithm specifically designed for screen-shooting scenarios. By dividing the watermark extraction process into three steps: automatic location, automatic correction, and automatic extraction, the algorithm enables batch copyright authentication of covertly photographed documents. A novel ring watermark template is designed, which uses its high-energy frequency domain characteristics for automatic location. Then, by combining the text line features of the document image, automatic correction is achieved. Finally, the watermark information is blindly extracted in conjunction with the watermark template. This algorithm is capable of accurately extracting watermark information even when only a portion of the document content is photographed. Experimental results show that the algorithm not only has good robustness against screen-shooting attack under complex environmental conditions but also resists many common image processing attacks, such as JPEG compression, Gaussian blurring, median filtering, scaling, translation, and cropping.

Keywords: Digital Watermarking, Copyright Protection, Document Watermarking, Screen-shooting Distortions, Robust Blind Watermarking

目 录

摘 要	I
ABSTRACT.....	II
目 录	IV
第一章 绪论	1
1.1 课题来源.....	1
1.2 课题研究背景与意义.....	1
1.3 国内外研究现状.....	2
1.3.1 抗屏摄水印研究现状.....	2
1.3.2 数字文档水印研究现状.....	5
1.4 论文研究内容与结构安排.....	7
1.4.1 主要研究内容.....	7
1.4.2 论文结构安排.....	8
第二章 相关技术基础	10
2.1 数字水印技术.....	10
2.1.1 数字水印系统框架.....	10
2.1.2 盲提取水印.....	11
2.1.3 水印评估指标.....	12
2.2 屏摄失真与分析.....	13
2.2.1 屏幕拍摄组合失真.....	14
2.2.2 屏摄失真分析.....	17
2.3 DFT 圆环水印	18
2.3.1 二维离散傅里叶变换.....	18
2.3.2 DFT 圆环水印	20
2.4 纠错编码与图像预处理算法.....	22
2.4.1 BCH 纠错编码.....	22
2.4.2 自适应直方图均衡化.....	23
2.4.3 维纳滤波与钝化掩蔽.....	24

2.5	本章小结.....	25
第三章	抗屏摄鲁棒文档水印算法	26
3.1	引言.....	26
3.2	算法思想.....	27
3.2.1	底纹预处理.....	27
3.2.2	水印区域定位.....	28
3.2.3	文本失真.....	31
3.3	基于 DFT 环形水印的文档底纹水印方案	31
3.3.1	水印嵌入流程.....	31
3.3.2	水印提取流程.....	34
3.4	实验结果与分析.....	37
3.4.1	实验参数.....	37
3.4.2	文档图像的视觉质量.....	39
3.4.3	水印的嵌入容量及鲁棒性分析.....	40
3.5	本章小结.....	43
第四章	抗屏摄文档水印自动提取方法	45
4.1	引言.....	45
4.2	算法思想.....	46
4.2.1	自动定位与矫正.....	46
4.2.2	水印模板的抗裁剪性.....	48
4.3	抗屏摄文档水印自动提取方案.....	49
4.3.1	水印嵌入流程.....	49
4.3.2	水印提取流程.....	51
4.4	实验结果与分析.....	55
4.4.1	实验参数.....	55
4.4.2	文档图像的视觉质量.....	57
4.4.3	对常见图像处理攻击的鲁棒性分析.....	58
4.4.4	水印的嵌入容量及鲁棒性分析.....	59

4.4.5 算法通用性测试.....	63
4.5 本章小结.....	63
第五章 总结与展望	64
5.1 总结.....	64
5.2 展望.....	65
参考文献	66
作者在攻读硕士学位期间发表的论文	74
作者在攻读硕士学位期间所作的项目	75
致 谢	76

第一章 绪论

1.1 课题来源

本课题来源于“CCF-腾讯犀牛鸟基金项目”，项目编号：CCF-Tencent Open Fund - RAGR20220126。

1.2 课题研究背景与意义

随着互联网和通信技术的日益发展，人们开始习惯使用移动通信设备互相交流，使用便携式摄像设备记录生活，然后分享至网络。人们可以轻松使用手机拍摄下屏幕上显示的文档、图像和视频，进行非法盗用和传播，从而造成严重的隐私泄露和版权侵犯问题。屏幕偷拍造成的机密信息泄露事件层出不穷，如何确保屏幕盗摄下的信息安全和版权认证已成为非常重要的研究课题。数字水印技术是一种新兴的信息隐藏手段，其核心在于将特定的数字信息（水印）嵌入到数字媒体文件中，以实现版权保护、内容认证和信息追踪等功能。这种技术通过在图像^[1-2]、音频^[3-4]、视频^[5-6]或文档^[7-8]等数字载体中嵌入不可见或难以察觉的标记，能够在不影响原始媒体质量的前提下，为媒体内容提供额外的身份和版权信息。数字水印的嵌入和提取过程往往需要精心设计，以确保水印的隐蔽性、鲁棒性和准确性。随着数字媒体的广泛传播，数字水印技术现已成为一门非常重要的数字技术。它不仅可以帮助版权所有者标识和保护自己的作品，防止未经授权的复制和分发，还可以在法律纠纷中提供重要的证据支持。此外，数字水印技术也广泛应用于数字资产管理、广告监测、媒体追踪和内容过滤等领域，为数字内容的安全传输和有效管理提供了强有力的技术保障。

传统的数字水印技术主要针对电子信道传输中的各种失真进行设计，如网络传输中常见的压缩、裁剪和格式转换等攻击。然而，在屏幕拍摄这一跨媒体信息传输过程中，传统的数字水印技术面临着更为复杂的组合失真，包括图像

的几何变形、光照变化、分辨率降低以及设备特有的噪声干扰等。在屏幕拍摄过程中，信息是通过空气信道传输至相机端，屏幕上显示的数字图像经过相机的拍摄后会在相机端形成新的数字图像。显示器的分辨率、相机的参数、拍摄的角度和距离、环境的亮度、摩尔纹的干扰都会导致数字图像在跨媒体信息传输的信息损失。因此，研究一种能够抵抗屏幕拍摄失真的数字水印技术，对于加强数字版权保护、维护网络信息安全具有重要的理论和实践价值。而在电子文档这类媒体文件中，屏幕盗摄所带来的隐私泄露风险尤为显著。在数字化办公和网络通讯日益普及的今天，大量的敏感信息和重要文本都以电子文档的形式存在。这些文档可能包含商业机密、个人身份信息、法律条文、政府资料等，一旦被未经授权的人员通过拍屏的方式获取，就可能导致严重的隐私泄露和信息安全问题。与图像和视频不同，文档信息的泄露风险更高，因为即使只是文档的一小部分内容，也可能包含大量的敏感信息。偷拍者仅需要拍摄文档的部分内容，就能够造成严重的隐私泄露。并且，对于需要频繁更新和修改的文档，如项目方案和报告，水印技术不能影响文档的编辑灵活性。而对于合同、个人档案这类信息重要，内容不可随意修改的文档，水印技术也应满足良好的隐蔽性，不能影响文档的内容显示，从而造成文本信息的误读。

综上所述，抗屏幕拍摄的文档水印技术研究，不仅是数字水印领域的一个技术挑战，也是信息安全领域的一个重要课题。通过研究出能够有效抵御屏幕拍摄失真的文档水印技术，将为个人隐私保护和企业机密安全提供更为坚固的防线，也对网络空间安全和多媒体版权保护具有重要意义。

1.3 国内外研究现状

1.3.1 抗屏摄水印研究现状

抗屏幕拍摄水印的先导研究是抗打印扫描水印和抗打印拍摄水印技术。所谓的打印扫描和打印拍摄，分别指的是将文档打印在纸张上，然后通过扫描或拍摄来转化为数字图像的过程。这两个过程和屏幕拍摄一样都是跨媒体的信息传输过程，水印信息需要经历数字信号到模拟信号的转换，然后再从模拟信号

到数字信号的转换，这种跨媒介传输过程引入的失真非常复杂。现有的研究主要基于两种方法进行：一种是手动调整图像在空间或频率域的特征^[9-13]嵌入鲁棒水印信息，另一种则是构建深度神经网络进行自适应学习^[14-16]。

目前关于抗屏幕拍摄水印的研究主要集中于图像领域。这些研究基于两种主要方法：一种是结合关键点和变换不变域的水印算法，另一种则是基于深度学习的水印算法。在这两种方法的基础上，研究者们努力提高水印的鲁棒性和隐蔽性，使其能够在跨媒介传输的复杂干扰下，仍然能够有效传输并准确提取水印信息。这种研究的核心在于找到一个平衡点，即在不显著影响原始图像质量的前提下，最大化水印的容量和对屏摄失真的鲁棒性。

1) 结合关键点与变换不变域的水印算法：该方法旨在寻找在跨媒体信息传输后仍能保持不变的图形关键点，并在这些关键点周围的变换不变域中嵌入水印信息，以抵抗屏幕拍摄引入的各种复杂组合噪声。该方法的关键在于利用图像的不变特征，确保水印信息在经历屏幕到摄像机这一非传统传播途径时依然能够被准确地定位和恢复。Fang 等^[17]首先使用基于强度的尺度不变特征变换（Scale-Invariant Feature Transform, SIFT）^[18]算法来计算图像关键点，然后使用特定选择的一对离散余弦变换（Discrete Cosine Transform, DCT）系数在关键点周围嵌入水印信息。Chen 等^[19-20]使用相同的架构提出了一种新方法，将 SURF（Speed-up Robust Featyres）^[21]特征方向算子和离散傅立叶变换（Discrete Fourier Transform, DFT）域相结合。在检测到关键点后，构建一个局部特征方形区域以嵌入水印，并使用 SURF 特征方向的计算方法为特征点赋予特征方向以实现水印前后的同步。Li 等^[22]则通过结合特征区域过滤模型与 SuperPoint 神经网络^[23]、四元数离散傅立叶变换（Quaternion Discrete Fourier Transform, QDFT）^[24-25]和张量分解，提出了另一种关键点和不变域组合的抗屏摄水印方案。Deng 等^[26]则提出了一种基于奇异值分解（Singular Value Decomposition, SVD）的屏幕拍摄弹性水印方案，该方案通过在图像的 SVD 奇异值矩阵中嵌入水印并利用 SIFT 算法定位嵌入区域，以增强水印的鲁棒性和不可见性，从而有效抵抗屏幕拍摄过程中产生的失真。然而，上述方法都忽略了水印嵌入对关键点检测的影响。Dong 等^[27]发现了这个问题，并提出了一种不影响水印的关键点增强算法。他们采用凸约束优化框架来统一关键点增强和水印嵌入，从而避免

水印嵌入后对关键点检测的影响。

此类型方法首先利用关键点的局部性和唯一性选择水印的嵌入区域，再利用关键点的不变性抵抗屏摄引起的透视形变，然后通过结合变换域水印对噪声的鲁棒性共同抵御屏摄过程中引入的组合噪声。虽然此类方法在抵御屏幕拍摄失真方面具有可行性和实效性，但在实际应用中仍然面临着一些问题。由于屏摄时光照不均或摩尔纹失真会导致图像的部分区域信息丢失，此类情况会极大影响关键点检测的准确性。此外，现有的屏摄水印算法在提取时往往需要人工预处理照片，需要手动定位并矫正照片中的水印图像，这增加了算法的复杂性和实施成本。因此，未来的研究需要进一步探索如何解决光照和摩尔纹失真对关键点检测的影响，以及如何实现自动化的照片预处理，从而提高抗屏幕拍摄水印技术的实用性和鲁棒性。

2) 基于神经网络的方法：随着深度学习技术的发展，特别是神经网络在图像处理和模式识别领域的应用，研究者开始探索如何利用这些先进的技术来增强抗屏摄水印的性能。Fang 等^[28]在先前的研究基础上，引入了神经网络以增强水印的提取性能。他们设计了一种基于翻转自相关的同步模板，用于定位水印区域，并在提取阶段利用神经网络对图像进行增强和分类，从而提高了水印的检测准确性和鲁棒性。随后，Fang 等^[29]还提出了名为“TERA”的水印算法，利用人眼对高频闪烁的融合现象，在嵌入端生成两帧交替显示的互补水印图像，在提取端则采用注意力引导的神经网络。Jia 等^[30]结合 Hidden^[31]和 StegaStamp^[15]的网络架构，提出了一种端到端的抗屏摄神经网络水印方案。他们对屏幕拍摄过程中引入的组合噪声进行仿真，构建了一个噪声层，并用于在图像空域中训练水印的嵌入和提取，以增强水印对屏摄攻击的鲁棒性。而 Bai 等^[32]则提出了基于频域的抗屏摄神经网络水印方案，在图像频域中训练水印的嵌入和提取。但该方法仅可用于灰度图像，并不适用于彩色图像和具有复杂纹理的图像。Aleksey 等^[33]设计了一种始终显示在屏幕上的水印膜，该水印膜在经过神经网络训练后变得难以察觉，但始终存在的水印膜可能会影响到其他应用程序的显示效果。Ge 等^[34]首次提出了基于神经网络的抗屏摄文档水印方案。然而，水印的不可见性较差，在水印嵌入后，文档图像上的水印图案会与文本重叠，影响文本内容的正常阅读。此外，基于图像的水印训练方式也限制了文档

的后续编辑。

基于深度神经网络的方法通过模拟屏幕拍摄噪声实现端到端的水印嵌入和提取训练。然而，这些噪声模拟层往往难以准确复制屏幕拍摄过程引入的各种复杂噪声。尽管在噪声层中模拟了透视失真，但在水印提取阶段仍然需要人工预处理照片，手动定位并矫正照片中的水印图像，以作为提取网络的输入。同时，神经网络端到端的训练方式在面对文档这一媒体文件时，就很难兼顾水印的隐蔽性和鲁棒性，以及文档的视觉质量和可编辑性，造成诸如文本信息覆盖、阅读观感差、无法编辑等问题。

1.3.2 数字文档水印研究现状

数字文档水印技术是一种将文本创建者、版权所有者或其他相关标识嵌入到电子文档中的技术，目的是在不影响文档正常阅读和编辑的情况下，保护文档内容免受未经授权的复制和分发。这种技术可以实现对数字文档的追踪、认证和保护。目前的文档水印算法主要存在三种方法：基于文档格式的水印算法、基于自然语言处理的水印算法和基于图像处理的水印算法。

基于文档格式的水印算法通过调整文档的文本格式来嵌入水印，例如调整字符的字体^[35]、间距^[36-39]、亮度^[40]或行间距^[41-42]等来嵌入水印。但这种方法可能在隐蔽性、版面一致性、编辑灵活性、鲁棒性以及兼容性方面存在局限性。尤其是面对跨媒体信息传输时，对文本格式的细微调整难以在保持隐蔽性的同时确保水印的鲁棒性。基于自然语言处理的水印算法^[43-49]利用自然语言处理技术，通过改变文本的句法或语义属性来嵌入水印，但这种方法无法保留原始文本内容，并且水印的嵌入容量有限。基于图像处理的水印算法^[50-55]是将文档作为图像处理，将水印信息嵌入到文本图像中，但这种方法可能会在嵌入水印后造成文档的视觉质量损失，影响文本内容的清晰度和可读性。此外，一旦文档被转换为图像格式并嵌入水印，将无法进行二次编辑，因为任何编辑都可能导致水印信息的损坏。这三种方法共同面临的挑战是，水印仅能够在文档编辑完成之后嵌入，这导致之后对文档的任何修改都可能损害水印的完好性。

当面对屏幕拍摄这种跨媒体信息传输时，传统的文档水印算法很难兼顾隐蔽性和鲁棒性。Guglemann 等^[56]提出了一种屏幕文档水印模板，通过在屏幕显

示的文档上叠加一个个亮度细微变化的圆形区域来嵌入水印信息，用模板中圆形区域的明暗来表示二进制比特信息 0 或 1。Fang 等人^[57]提出了基于文档水印底纹的方法，该方法通过在文档底纹中嵌入水印来解决可编辑性和视觉质量问题。通过修改一对选定的 DCT 系数，将水印信息嵌入到文档的背景底纹中，并在提取阶段利用翻转模板的对称性来准确定位水印区域。然而，该算法中每一比特水印信息的嵌入需要 16×16 像素的图像块，且在应用翻转模板后，128 比特水印信息将占用 512×512 像素的区域。这导致如果偷拍者仅拍摄文档的部分区域，水印提取可能会失败。此外，该方法在水印定位前需手动对拍摄的文档图像进行透视矫正，未能彻底解决人工预处理问题，并且 DCT 水印算法对照明和摩尔纹失真的鲁棒性较差。Yang 等人^[58]则提出了一种基于水印字体的文本水印算法，通过修改字形的质心来生成新的水印字体，并通过字体分割和字体的相对质心位置来提取水印信号。但这种方法会导致字体的较大变化，并且为不同的水印信号生成不同的水印字体库需要大量的计算资源和时间。

屏摄文档水印算法相较于屏摄图像水印算法的难点在于如何保持文档特性的同时保证水印对屏摄攻击的鲁棒性。水印算法不应影响文档内容的可读性和可编辑性。同时，考虑到算法的实用性，当偷拍者拍摄屏幕上显示的文档时，偷拍者更关注的是文本内容，不会像拍摄图像一样将整个图像完整拍摄下来，因此盗摄照片中的文档图像很少有清晰的边缘特征。文档在编辑过程中的尺寸大小也不是固定不变的，这将导致在水印提取前的预处理阶段，很难对盗摄照片中缺少边缘特征和尺寸不明的文档内容进行透视矫正。除此之外，屏摄失真中的亮度失真和摩尔纹失真会导致文档图像部分区域的内容丢失，这种信息损失是不可逆的，如何解决这种不可逆的信息损失也是屏摄文档水印算法的一大难点。而且，在面对互联网中大量盗摄文档照片的版权认证问题，消除人工预处理是十分实用且必要的，实现文档水印的自动化定位和矫正将是屏摄文档水印技术实用性发展的关键。

1.4 论文研究内容与结构安排

1.4.1 主要研究内容

本文对屏幕拍摄引起的失真、抗屏摄图像水印算法和抗屏摄文档水印算法进行了调研，分析了各领域的国内外研究现状。在抗屏摄文档水印技术的研究中，主要面临的难点集中在如何平衡水印的隐蔽性与文档内容的完整性，确保在不影响阅读观感和可编辑的前提下，能够有效抵抗屏幕拍摄过程中的组合噪声。在文档编辑过程中水印的稳定性也是一个关键问题，需要确保即使在文档内容发生变化后，水印信息依然可以被准确地识别和提取。除了需要抵抗屏幕拍摄攻击，文档水印还应该能够在数字信道传输过程中保持其完整性，这意味着水印需要具备对常见电子信道攻击的鲁棒性，以确保屏摄后的水印文档在不同设备和平台的传播中依然可以提取水印信息。此外，当面对批量盗摄照片进行版权认证时，如何实现水印的自动盲提取这个难题也必须得以解决。

本文针对上述问题提出了基于 DFT 环形水印模板的文档底纹水印算法，在文档底纹中嵌入 DFT 环形水印，可以在不影响文档可编辑性的情况下实现良好的视觉质量和对屏摄攻击的较强鲁棒性。此外，为了解决水印提取过程中的人工预处理问题，本文提出了抗屏摄文档水印自动提取算法，设计了一种新型的 DFT 环形水印模板。论文的具体内容如下：

1) 提出了一种基于 DFT 环形水印的抗屏摄文档底纹水印算法。通过在小尺寸的水印底纹中嵌入 DFT 环形水印模板，然后拼接平铺组成完整的水印底纹。本算法借助于 DFT 频谱的线性与叠加特性，可以确保只要截取的文档片段尺寸等于或大于水印模板的最小尺寸，就能够准确地提取嵌入的水印信息。因此，该算法能有效抵御光照失真、摩尔纹失真以及裁剪攻击。通过在在文档底纹中嵌入频域水印，可以同时兼顾文档的视觉质量和可编辑性，并且 DFT 水印的几何不变形和对常见图像处理攻击的鲁棒性，能够有效抵抗透视形变和拍摄过程引入的多种噪声。实验结果表明，本算法在抗屏摄鲁棒性上领先于已有的抗屏摄文档水印算法。

2) 提出了抗屏摄文档水印自动提取算法。当前抗屏摄水印研究在提取阶段

无法避免对照片做人工预处理操作，需要手动定位和矫正照片中变形的水印图像。并且当前多数研究普遍采用四点矫正算法，通过标注图像的四个角点，并定义矫正后的目标角点来计算透视变换矩阵将水印图像矫正回原平面。本算法将水印提取过程拆分为自动定位、自动矫正和自动提取三个步骤。通过设计一种新型的 DFT 环形水印模板，在两个水印环中间加入定位环，既可以提升水印容量，还能通过定位环的高能特征实现水印图像在照片中的定位。在完成对照片中水印文档的自动定位后，利用文档特有的文本行特征将文档矫正回原平面。但此时矫正后的图像尺寸是变形的，定位圆环也变形为了椭圆环，此时可以再次利用定位环的高能特征检测定位环并将其拉伸回正圆环，最后结合水印提取算法完成水印的盲提取。实验结果表明，本算法可以有效解决抗屏摄文档水印的人工预处理问题，实现了抗屏摄文档水印的自动提取。在保留良好视觉质量和抗屏摄失真鲁棒性的同时，也拥有对常见的信号处理和几何攻击的良好鲁棒性，包括 JPEG 压缩、高斯模糊、中值滤波、缩放、平移和裁剪攻击。

1.4.2 论文结构安排

本文基于作者在硕士研究生期间承担的课题工作为基础展开，共有五章内容，其中各章节内容安排如下：

第一章，首先阐明了课题的研究背景及意义，明确了本文研究的重点和方向。然后介绍了国内外在抗屏摄水印技术和文档水印技术领域的研究进展，并深入分析了现有研究的局限性。最后，简要介绍了本文在课题研究方面的创新成果。

第二章，首先介绍了数字水印技术，讲解了数字水印系统的基本架构、盲提取水印技术的分类，以及屏摄水印算法性能评价的关键指标。然后，详细介绍并分析了屏摄过程中可能遇到的失真类型，如相机失真、拍摄失真、屏幕显示失真、光照失真和摩尔纹失真。此外，本章还介绍了图像的二维离散傅里叶变换的基础性质，DFT 圆环水印技术，以及 BCH 纠错编码和图像预处理技术，包括自适应直方图均衡化、维纳滤波和钝化掩蔽处理。

第三章，针对现有屏摄文档水印算法鲁棒性不足的问题，提出了一种基于 DFT 环形水印的文档底纹水印算法。首先结合第二章中对屏摄失真的具体分

析，探讨了底纹预处理、水印区域定位和文字失真的算法思想。随后利用 DFT 频谱和环形水印模板的特性，通过拼接小尺寸水印底纹构建完整的水印底纹，可以同时满足良好的视觉质量和文档的可编辑性。最后，经实验验证，该算法不仅有效抵抗了屏摄组合失真，而且相较于现有方法，展现了更优的鲁棒性。

第四章，针对现有抗屏摄水印研究中普遍存在的手动定位与矫正问题，提出了一种自动盲提取算法。本章首先探讨了如何实现对照片中水印文档的自动定位与矫正，以及如何处理仅拍摄部分文档内容的情况。随后，在水印嵌入端设计了一种新型的 DFT 环形水印模板，并将水印提取端拆分为三个步骤：自动定位、自动矫正和自动提取。通过结合环形水印模板的定位特征和文档文本行特征，实现了水印的自动盲提取。最终，经实验验证，该算法对屏摄失真及常见图像处理攻击均具有良好鲁棒性。

第五章，对本文的研究内容进行总结，并结合现有方案的不足，对抗屏摄水印研究的未来发展做出展望。

第二章 相关技术基础

2.1 数字水印技术

2.1.1 数字水印系统框架

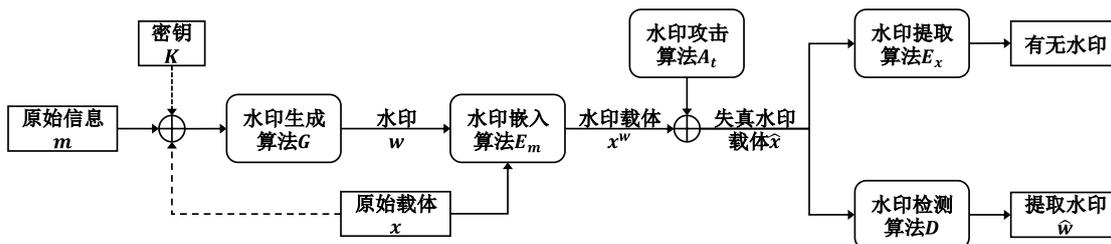


图 2.1 数字水印系统的基本框架

数字水印系统包含嵌入器和检测器两大部分。嵌入器至少具有两个输入量：一个是原始信息，它通过适当变换后作为待嵌入的水印信号；另一个是要在其中嵌入水印的载体对象。水印嵌入器的输出结果为含水印的载体对象，通常用于传输和转录。水印检测器的输入对象则是含水印的这个对象或未含水印的其他对象，水印检测器会判断输入对象是否存在水印，若存在，则利用水印提取算法输出嵌入的水印信号。图 2.1 详细展示了数字水印系统的基本框架，它可以定义为九元体 $(M, X, W, K, G, E_m, A_t, D, E_x)$ ，分别定义如下：

- 1) M 代表所有可能原始信息 m 的集合；
- 2) X 代表所要保护的数字对象 x 的集合；
- 3) W 代表所有可能水印信号 w 的集合；
- 4) K 代表水印密钥 K 的集合；
- 5) G 表示用原始信息 m ，密钥 K 和原载体 x 共同生成水印的算法，即

$$G: M \times X \times K \rightarrow W, w = G(m, x, K) \quad (2.1)$$

其中，原载体对象不一定参与水印的生成过程。

- 6) E_m 表示将生成水印 w 嵌入原始载体对象 x 中的嵌入算法，即

$$E_m: X \times W \rightarrow X, x^w = E_m(x, w) \quad (2.2)$$

其中， x 和 x^w 分别代表水印嵌入前后的载体对象。

7) A_t 表示在电子信道中存在的各种攻击算法, 即

$$A_t: \mathbf{X} \times \mathbf{K} \rightarrow \mathbf{X}, \hat{\mathbf{x}} = A_t(\mathbf{x}^w, K') \quad (2.3)$$

其中, K' 代表攻击者伪造的密钥, $\hat{\mathbf{x}}$ 代表被攻击后的失真的含水印对象。

8) D 表示水印检测算法, 即

$$D: \mathbf{X} \times \mathbf{K} \rightarrow \{0,1\}, D(\hat{\mathbf{x}}, K) = \begin{cases} 1 & \text{如果}\hat{\mathbf{x}}\text{中存在}\mathbf{w} \\ 0 & \text{若}\hat{\mathbf{x}}\text{中不存在}\mathbf{w} \end{cases} \quad (2.4)$$

9) E_x 表示水印提取算法, 即

$$E_x: \mathbf{X} \times \mathbf{K} \rightarrow \mathbf{W}, \hat{\mathbf{w}} = E_x(\hat{\mathbf{x}}, K) \quad (2.5)$$

其中 $\hat{\mathbf{w}}$ 即为最终提取的水印信号, 与原生成水印信号 \mathbf{w} 进行对比后就能评价水印提取的准确率。

2.1.2 盲提取水印

基于不同的水印嵌入方式, 可以将水印算法划分为三个类别: 盲提取水印、非盲提取水印和半盲提取水印, 它们在提取水印时对先验信息的需求程度不同。

1) 盲提取水印: 在提取水印时不需要任何关于原始载体的先验信息。这种技术允许在没有原始媒体或参考信号的情况下, 可以直接从经过各种信号处理的含水印载体中提取水印。盲提取水印的设计旨在水印的提取过程中无需原始载体的参与, 从而提高水印的实用性和鲁棒性;

2) 非盲提取水印: 在提取水印时需要原始载体的完整信息或参考信号。这种技术通常依赖于原始载体的精确匹配来检测和提取水印, 因此对原始载体的可用性有很高的要求。非盲水印的优点是它可以提供非常高的检测准确性和鲁棒性, 因为在提取过程可以利用原始载体的所有信息;

3) 半盲提取水印: 位于盲提取水印和非盲提取水印之间, 它在提取水印时需要一些关于原始载体的先验信息, 但不像非盲水印那样需要完整的原始载体。这些先验信息可能包括原始载体的某些统计特性、部分内容或特定的结构特征。半盲水印技术需要利用这些有限的信息来辅助水印的提取过程, 从而提高提取的准确性和鲁棒性。

当前大多数抗屏摄水印算法在水印提取阶段都依赖于手动定位和矫正照片中的水印图像。此外，部分方法还需要预先知晓图像的尺寸信息，因此这些方法均被归类为非盲提取水印。而本文第四章提出的抗屏摄文档水印自动提取算法，则解决了手动定位和矫正的问题。该算法无需原始载体的任何先验信息，能够直接将屏摄照片作为水印提取阶段的输入，从而实现了水印的盲提取。

2.1.3 水印评估指标

1) **视觉质量评估**：在数字水印领域，视觉质量是衡量水印效果的重要指标之一，它直接关系到水印的隐蔽性和对图像质量的影响。视觉质量的评价通常包括客观和主观两个方面。客观评价方法通过数学模型来量化水印对图像质量的影响，而主观评价则依赖于人眼观察和评价。本文采用了两种广泛认可的客观质量评价方法：峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）和结构相似性指数（Structural Similarity Index Measure, SSIM）^[59]，来评估水印的视觉质量。PSNR 是一种基于均方误差（Mean Squared Error, MSE）的指标，用于衡量原始图像与处理后图像之间的相似度。PSNR 的计算公式如下：

$$\text{PSNR} = 10 \times \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (2.6)$$

其中， MAX_I 是图像可能的最大像素值， MSE 是原始图像与处理后图像之间的均方误差。PSNR 的值越高，则表示两幅图像之间的差异越小，从而意味着水印的隐蔽性更好，对图像质量的影响也越小。然而，PSNR 主要关注图像整体的误差水平，可能无法充分反映图像中特定区域的质量变化。为了更全面地评估图像的视觉质量，本文还引入了 SSIM 指标。SSIM 是基于人类视觉感知特性设计的，它考虑了亮度、对比度和结构三个不同维度的相似性。SSIM 能够更好地捕捉到图像中的局部变化和结构信息，从而提供更为准确的视觉质量评价。SSIM 的计算公式如下：

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2.7)$$

其中， x 和 y 分别代表原始图像和处理后图像， μ_x 和 μ_y 对应两个图像的均值，

σ_x^2 和 σ_y^2 对应两个图像的方差， σ_{xy} 则是两个图像的协方差。常数 c_1 和 c_2 是为了避免分母为零会造成不稳定性而加入的小常数，通常 $c_1 = (k_1 L)^2$ ， $c_2 = (k_2 L)^2$ ，其中 L 是图像像素值的动态范围，而 $k_1 = 0.01$ 和 $k_2 = 0.03$ 为默认值。高 SSIM 值表明水印图像在结构上与原始图像保持高度一致，这对于确保水印的隐蔽性和图像质量的保护至关重要。通过结合 PSNR 和 SSIM 两种客观评价方法，可以从不同角度对水印的视觉质量进行综合评估。这两种评价指标不仅能够量化水印对整体图像质量的影响，还能够评估水印对图像局部结构和视觉感知的影响，从而为水印算法的设计和优化提供重要的参考依据和客观评价依据。

2) 鲁棒性评估：鲁棒性是衡量数字水印算法在面对各种信号处理攻击和噪声干扰时稳定性的关键指标。它反映了水印在图像经过处理后仍能被准确检测和提取的能力。本文采用了比特提取准确率 (Bit Accuracy Rate, BAR) 作为鲁棒性的衡量指标，以提供一个公平且一致的评价标准。BAR 是一种量化水印鲁棒性的指标，它计算了在水印提取过程中正确的水印比特数与总比特数的比例。比特提取准确率实际是误码率 (Bit Error Rate, BER)^[60] 的补数，BER 是用来衡量传输或处理过程中各种错误的比特数占总比特数的比例。所以 BAR 的值越高，表示水印的鲁棒性越好，即在经历攻击或噪声干扰后，水印信息的保持越完整。通过使用 BAR 作为评价指标，能够直观地了解水印算法在不同攻击和干扰条件下的性能表现。高 BAR 值意味着即使在恶劣的条件下，水印信息也能被准确地提取出来，从而证明了水印算法的鲁棒性。这种评价方法为水印技术的比较和优化提供了一个有效的量化基准。

2.2 屏摄失真与分析

屏幕拍摄是用相机拍摄屏幕上显示的图像的过程。在拍摄时，信号是通过空气信道从屏幕传输到相机，水印在此过程中分别进行了数模转换和模数转换，其中也经历了各种失真。本文结合方^[61]和陈等^[62]对屏摄失真的已有分析，将屏摄失真进一步细分为五个方面：相机失真、拍摄失真、屏幕显示失真、照明失真和摩尔纹失真。

2.2.1 屏幕拍摄组合失真

1) **相机失真**: 相机失真主要涉及相机硬件和软件处理过程中引入的图像质量变化, 不同品牌手机的相机硬件不同, 图像处理算法不同, 可设置的图形处理参数也各不相同, 会引起诸如色彩处理失真、动态范围压缩失真、噪点抑制失真、图像压缩失真等。

- a) **色彩处理失真**: 相机在将捕获的图像数据转换为可查看的照片时, 通常会应用色彩校正算法。这些算法旨在模拟特定的色彩特性或增强图像的视觉效果, 但可能会导致色彩偏差, 使得图像的颜色与实际屏幕显示的颜色不一致。例如, 相机可能会过度饱和或过度锐化图像, 造成颜色失真。
- b) **动态范围压缩失真**: 为了适应相机传感器的动态范围限制, 相机在处理高对比度场景时可能会自动应用动态范围优化 (如 HDR 处理)。这种处理可能会导致图像的亮部细节丢失或暗部细节被过度增亮, 从而影响水印的可见性和准确性。
- c) **噪点抑制失真**: 在低光照条件下拍摄或使用高感光度 ISO 设置时, 图像传感器会产生噪点。为了减少噪点, 相机可能会应用噪声抑制算法。然而, 这些算法可能会平滑图像细节, 导致水印信息的损失。
- d) **图像压缩失真**: 为了节省存储空间, 相机通常会将图像压缩为 JPEG 格式。JPEG 压缩是有损的, 意味着在压缩过程中会丢弃一部分图像数据。这种压缩可能会导致图像细节的丢失, 尤其是在高对比度区域, 从而影响水印的完整性。

2) **拍摄失真**: 是指在将三维世界中的物体通过摄像机捕捉并转换为二维图像的过程中, 由于摄像机的角度、距离、镜头特性以及其他因素导致的图像变形和质量变化。如图 2.2 所示, 在拍摄屏幕上显示的图像时, 这种失真尤为明显, 因为屏幕本身就是一个二维平面, 而摄像机需要将其再次转换为另一个二维平面的图像。所以照片中的水印图像实际是屏幕平面到相机平面的投影, 并且由于摄像机镜头与屏幕之间的相对位置和角度, 会造成图像局部区域的拉伸或压缩。例如, 当摄像机从斜角度拍摄屏幕时, 离摄像机较近的图像区域会比

距离较远的边缘区域更加清晰和放大，这是因为透视作用导致的中心区域放大效果。同样因为摄像机的角度、距离和位置的变化会产生 RST 失真，即旋转 (Rotation)、缩放 (Scaling) 和平移 (Translation) 变换对图像造成的影响。



图 2.2 拍摄引起的透视失真示意图

3) 屏幕显示失真: 随着技术的发展，液晶显示屏 (Liquid Crystal Display, LCD) 已经完全取代阴极射线管 (Cathode Ray Tube, CRT) 显示器成为市场主流。液晶显示器是由大量的像素点组成，这些像素点按照矩阵的形式排列。每个像素点由红、绿、蓝 (RGB) 三个子像素组成，通过调整这些子像素的亮度来混合出不同的颜色。显示器的显示质量很大程度上取决于像素点的密度和排列方式。显示器的分辨率指的是屏幕上显示的像素点的总数，所以分辨率越高，单位面积内的像素点越多，理论上图像的细节表现也就越丰富。然而，显示器的物理构造限制了其能够展示的细节程度，由于液晶显示器的点矩阵结构，每个像素点实际是由更小的液晶单元组成，这些单元在物理上存在一定间隙，这些间隙是不发光的，其灰度值较低，与液晶点的颜色之间会形成周期性的颜色梯度变化，如图 2.3 (a) 所示。当水印文档在显示器上显示时，高分辨率有助于更精确地呈现水印的细节。但是，由于液晶显示器的物理结构，即使分辨率很高，也会因为液晶点的间隙和灰度值较低而丢失一些图像细节。此外，显示器在显示图像时会经过降采样处理，进一步减少了图像的细节层次。

这种降采样和物理结构的限制，会导致在屏幕上展示的水印文档的显示效果不如原始数字文件，并且部分图像细节的丢失也会影响水印信息的完整表达。

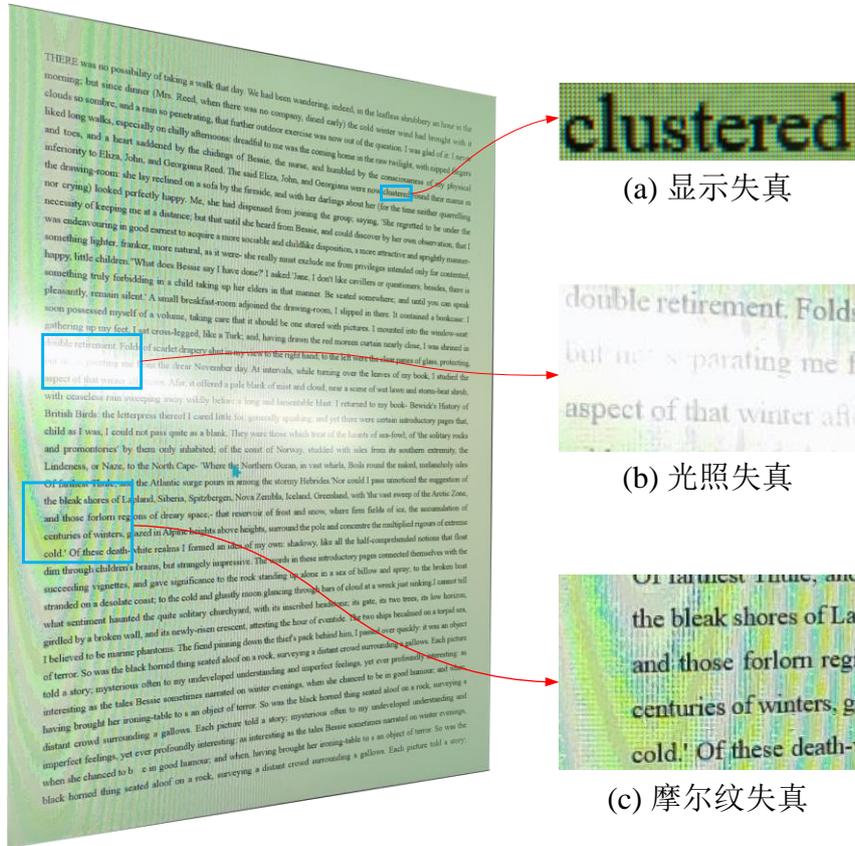


图 2.3 屏摄失真引起图像质量变化的三种常见失真

4) **光照失真**: 是指在屏幕拍摄过程中，由于屏幕的自发光特性和外部光照条件的影响，导致图像质量下降的现象。这种失真涉及到多个方面，包括亮度不均、过曝、光斑等。首先，屏幕的 LCD 点矩阵结构会导致图像显示时的亮度分布不均。LCD 屏幕通过控制背光层和液晶层的交互来调节像素的亮度，但由于背光层的发光特性和液晶层的调制效果，屏幕不同区域的亮度可能存在差异。这种现象在屏幕边缘和角落尤为明显，可能导致屏幕显示的水印文档图像在亮度上出现不均匀的条纹或斑点。其次，拍摄场景的照明条件对屏幕拍摄的图像质量有显著影响。在强光照射下，屏幕表面可能会产生反射和眩光，这些反射光会干扰屏幕显示的图像，导致照片中出现光斑或过曝区域，如图 2.3 (b) 所示。过曝现象是指图像中的亮度超出了相机传感器的动态范围，导致细节丢失，特别是在高光区域。

5) **摩尔纹失真**: 摩尔纹是一种视觉干扰现象，通常在数字成像和显示技术

中出现，当两个或多个具有相近频率的周期性图案重叠时，会产生这种失真。在相机内的电荷耦合器件（Charge-Coupled Device, CCD）传感器和屏幕的 LCD 结构中，这种失真尤为常见，因为这两种设备都以虚线或点阵的形式来分布显示图像信息。摩尔纹的产生源于差拍原理，该原理指出，当具有相似空间频率的两个条纹叠加时，复合信号的幅度会根据它们空间频率之间的差异而变化。这种因叠加而产生的视觉效果，通常以无规律的条纹或波纹的形式出现。摩尔纹在视觉艺术、纺织设计、以及数字成像等领域中是一个常见的问题，因为它可能会影响最终图像的质量和美观度，如图 2.3 (c) 所示。

2.2.2 屏摄失真分析

在屏幕拍摄过程中，图像的质量和完整性会受到上述五种失真的影响，导致照片中水印图像的非线性退化，退化程度取决于屏幕分辨率、相机像素数、拍摄角度和距离、光照变化、摩尔纹占比和相机后处理等。

其中，屏幕分辨率和相机像素数是影响拍摄照片质量的关键设备参数。高分辨率的屏幕能够显示更多的图像细节，而使用低像素的相机拍摄时，这些细节将无法被完全捕捉，导致图像质量下降。此外，相机的像素质量和传感器大小也会影响其捕捉细节的能力，从而影响最终成像的清晰度和真实性。本文提出的方法经实验验证，在使用较低分辨率和像素数较低的手机摄像头时也可以实现水印的准确提取。

拍摄角度和距离会引起图像的透视变形和缩放，这些变化会影响图像中水印的空间占比和尺寸。当相机从非正面角度拍摄时，图像中的直线可能会出现弯曲，导致水印的形状和位置也发生扭曲。此外，随着拍摄距离的增加，水印图像在照片中的占比会减小，水印的细节也会变得模糊不清，从而增加了水印检测和提取的难度。因此，为了准确提取水印，必须在水印提取之前对图像进行定位和矫正，以消除由拍摄角度和距离引起的形变和缩放，进而恢复水印图形的原始平面和大小。

光照变化和摩尔纹是在屏幕拍摄过程中常见的两种失真现象，它们对图像质量有着显著的影响。光照失真主要是由于拍摄环境中光源的位置和强度不一致造成的，这可能导致图像中的局部区域出现过曝，进而导致重水印信息的不

可逆损失。为了降低局部过曝对水印提取的干扰，本文提出了一种策略，即在文档中使用由多个小尺寸水印底纹组成的完整水印底纹。通过这种设计，多个水印以平铺拼接的方式分布在整个文档上。这样的布局使得即使文档的某个区域受到光照失真的影响，也不会影响整体水印的检测和提取。摩尔纹是由于液晶显示屏与相机采样之间的空间频率上的混叠而产生的，会在照片中形成形状不一的条纹，有时会覆盖图像的大部分区域。摩尔纹的存在同样对水印的准确提取造成了很大影响。通过实验验证，DFT 圆环水印可以有效抵抗摩尔纹失真。通过在频域的中频带嵌入水印信息，能够在一定程度上避免摩尔纹的干扰，从而保护水印信息的完整性和可提取性。

相机拍摄照片的后处理过程，如色彩校正、锐化、降噪和压缩，也会对最终成像质量产生影响。这些处理步骤旨在优化图像的视觉效果，但不当的处理算法可能会损害图像中的水印信息。特别是在图像压缩过程中，算法通常会对图像的高频成分进行修改，以减少数据量。而高频成分通常包含图像的细节和边缘信息，压缩可能会导致这些重要信息的丢失，进而影响水印的可检测性。为了提高水印的鲁棒性，本文选择在中频带嵌入水印。中频区域相较于高频区域，更能保持图像在压缩前后的一致性，因此对压缩算法更为稳健，从而提高水印算法对抗压缩的能力。这种策略使得水印可以在相机后处理中得以留存，并且在面对常见的图像处理攻击，如 JPEG 压缩时也可以表现出一定的鲁棒性。

2.3 DFT 圆环水印

2.3.1 二维离散傅里叶变换

离散傅里叶变换在数字水印领域占有重要地位，其中二维离散傅里叶变换可以将图像从空间域转换到频率域，然后在图像的频域中完成水印的嵌入和提取。通过修改图像的频域成分，可以实现在不影响图像质量的前提下，嵌入隐蔽不可见的水印信息。二维离散傅里叶变换的公式为：

$$F(u, v) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left(\frac{xu}{M} + \frac{yv}{N} \right)} \quad (2.8)$$

其中, $f(x, y)$ 是图像在空间域的像素值, $F(u, v)$ 是对应的频率域表示, M 和 N 分别是图像的宽度和高度, u 和 v 为频率域中的坐标。 $F(u, v)$ 是图像在频率域的复数表示, 可以拆分为幅度谱 $AM(u, v)$ 和相位谱 $\phi(u, v)$, 其公式分别为:

$$AM(u, v) = |F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)} \quad (2.9)$$

$$\phi(u, v) = \angle F(u, v) = \arctan\left(\frac{I(u, v)}{R(u, v)}\right) \quad (2.10)$$

其中, $R(u, v)$ 和 $I(u, v)$ 分别表示 $F(u, v)$ 的实部和虚部。 $AM(u, v)$ 表示各个频率分量的强度, 即图像中各个频率分量的能量分布。 $\phi(u, v)$ 表示各个频率分量的相位信息, 描述了不同频率分量相对于基频的相位差, 其在图像重建中非常重要, 细微的调整也会导致图像的严重失真。所以在数字水印应用中, 经常将水印信息嵌入到幅度谱 $AM(u, v)$, 然后再结合原相位谱 $\phi(u, v)$ 进行二维傅里叶反变换 (Inverse Discrete Fourier Transform, IDFT) 得到水印图像, 其公式为:

$$f(x, y) = \frac{1}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi\left(\frac{xu}{M} + \frac{yv}{N}\right)} \quad (2.11)$$

二维离散傅里叶变换从图像处理角度看, 具有很多特殊的性质:

1) 线性

$$af_1(x, y) + bf_2(x, y) \Leftrightarrow aF_1(u, v) + bF_2(u, v) \quad (2.12)$$

2) 周期性

$$f(x, y) = f(x + k_1M, y) = f(x, y + k_2N) = f(x + k_1M, y + k_2N) \quad (2.13)$$

$$F(u, v) = F(u + k_1M, v) = F(u, v + k_2N) = F(u + k_1M, v + k_2N) \quad (2.14)$$

3) 平移性

$$f(x, y) e^{j2\pi\left(\frac{xu_0}{M} + \frac{yv_0}{N}\right)} \Leftrightarrow F(u - u_0, v - v_0) \quad (2.15)$$

$$f(x - x_0, y - y_0) = F(u, v) e^{-j2\pi\left(\frac{x_0u}{M} + \frac{y_0v}{N}\right)} \quad (2.16)$$

4) 对称性

$$f(x, y) \text{实函数} \Leftrightarrow F^*(u, v) = F(-u, -v) \quad (2.17)$$

在对幅度谱 $AM(u, v)$ 分析和处理时，通常会对频谱图进行零频移位，也即将频谱坐标系的原点从 $(0, 0)$ 平移至 $(M/2, N/2)$ ，由公式(2.14)可得：

$$f(x, y)(-1)^{x+y} \Leftrightarrow F\left(u - \frac{M}{2}, v - \frac{N}{2}\right) \quad (2.18)$$

零频移位可以使直流系数位于频谱中心，低频信息围绕频谱中心，高频信息分散至频谱角点。此时幅度谱 $AM(u, v)$ 的值越大，表示该频率成分的能量越强，其位置和数值反映了图像不同频率成分的能量分布和大小。以频谱中心为圆心，靠近中心的区域对应于低频成分，远离中心的区域对应于高频成分，而介于两者之间的区域对应于中频成分。低频区域对应于图像中的大尺度结构和平滑变化；中频区域对应于图像的细节和纹理信息；高频区域对应于图像的边缘、复杂结构、突变部位和噪声等信息。图像的低频区域通常包含了图像的主要视觉信息，而高频区域则包含了图像的细节和噪声信息，中频介于两者之间，它既不像低频区域那样对视觉质量有显著影响，也不像高频区域那样容易受到噪声和其他信号处理的影响，例如图像的压缩处理会优先删去高频信息。所以频域水印通常会选取中频作为嵌入区域，可以同时兼顾隐蔽性和鲁棒性。

2.3.2 DFT 圆环水印

DFT 水印技术通过在频率域中调整图像的频率成分来嵌入隐蔽的水印信息。该技术存在多种嵌入策略，包括基于幅度的变化、相位的调整、能量的重新分配、复数系数的直接修改以及自适应嵌入等。特别地，基于幅度调制的方法可以进一步细分为加性幅度调制^[63-65]和乘性幅度调制^[66-67]两种方式。DFT 环形水印便是基于加性幅度调制，通过在 DFT 频谱内的特定环形区域嵌入水印，实现对水印信息的隐藏。DFT 水印的主要优点在于其能够在不显著降低图像空间域视觉质量的同时，为水印提供良好的隐蔽性，并且对常见的信号处理攻击，如几何变换、压缩、滤波和噪声添加等，都有着较强的鲁棒性。此外，DFT 水印技术还能够利用频率域的特性，如平移和缩放不变性、抗裁剪性，来设计抵抗几何变换的水印方案。因此，为了增强水印对几何攻击的鲁棒性，研究者们设计了特殊形状的水印模板，引入了圆环形水印^[68-71]。这种圆环形水印

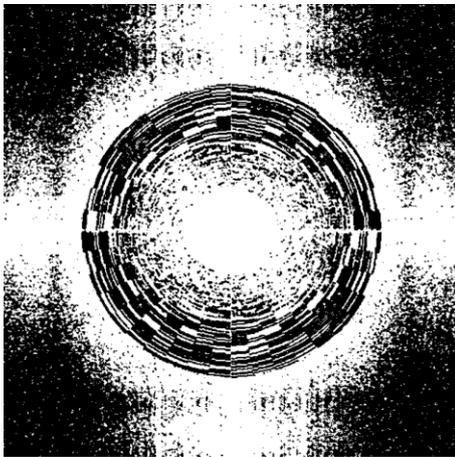
模板可以抵御 JPEG 压缩、滤波、加噪、尺度变换、平移、剪切、旋转、打印、再次扫描等攻击。文献[69]首先提出了多重环形水印：

$$W(r, \theta) = \begin{cases} 0 & \text{若 } r < R_1 \text{ 且 } r > R_2 \\ \pm 1 & \text{若 } R_1 < r < R_2 \end{cases} \quad (2.19)$$

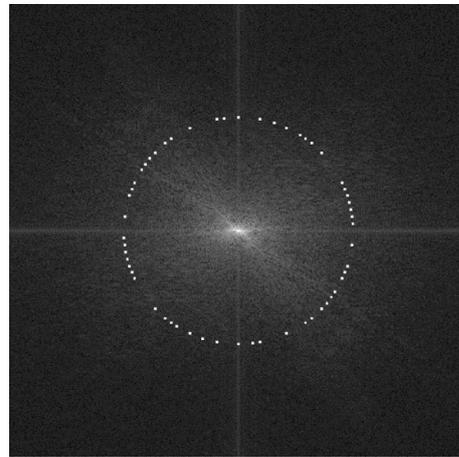
其中， $r = \sqrt{k_1^2 + k_2^2}$ ， $\theta = \arctan(k_2 / k_1)$ ， k_1 和 k_2 表示频谱图的直角坐标， r 和 θ 为极坐标。数目为 $R_2 - R_1$ 的每个同心圆划分了 S 个扇区，每个扇区分配 1 或 -1 的水印信息。然后使用基于加性幅度调制的嵌入方法得到水印频谱图 M' ：

$$M'(k_1, k_2) = M(k_1, k_2) + \alpha \cdot W(k_1, k_2) \quad (2.20)$$

其中， α 是水印嵌入强度因子， $W(k_1, k_2)$ 为 $W(r, \theta)$ 对应的直角坐标频谱系数（1 或 -1）。图 2.4 (a) 展示了 512×512 大小的 Lena 在嵌入多重圆环水印后的二值化频谱图。如前文所述，在水印嵌入前需要先对频谱图进行零频移位，将原点坐标移至图像中央，并且多重圆环水印嵌入在中频区域，从而有效平衡水印的隐蔽性和鲁棒性。



(a) 多重圆环水印二值化频谱图



(b) 单圆环水印频谱图

图 2.4 圆环水印频谱示意图

文献[68]和[72]则提出了在频域中嵌入单圆环水印：

$$W(k_1, k_2) = \begin{cases} m & \text{若 } k_1^2 + k_2^2 = R^2 \\ 0 & \text{若 } k_1^2 + k_2^2 \neq R^2 \end{cases} \quad (2.21)$$

其中， m 为长度为 L 的伪随机二进制序列，通过将半径为 R 的圆环划分为 L 份，在每份对应的 (k_1, k_2) 坐标上嵌入伪随机序列对应值比特（0 或 1）。单圆环

水印嵌入后的 Lena 频谱图如图 2.4 (b) 所示, 在半径 R 上的圆环嵌入伪随机二进制序列, 并且由于 DFT 的对称性, 水印圆环也关于图像中心对称。但以上研究仅可以通过相关检测法检测水印是否存在, 无法准确提取完整的水印信息。

本文基于现有的 DFT 圆环水印和文档水印研究, 在第三章提出了在文档底纹中嵌入 DFT 环形水印的方案。原始水印底纹的尺寸设为 256×256 , 通过平铺拼接的方式构成完整的水印文档底纹。由于 DFT 的线性、周期性和叠加原理, 拼接后的底纹在频域中表现为强度叠加的等比放大水印环, 水印环的半径和强度都会随着拼接尺寸的放大而等比增加。这一性质使得当任意截取超过模板尺寸的水印底纹时, 也能够保持水印的完整性和可检测性, 从而提高了水印的对裁剪攻击的鲁棒性。为了解决手动定位、矫正和非盲提取的问题, 本文在第四章又设计了一种新的 DFT 圆环水印模板, 利用 DFT 频域的高能量特征可以在跨媒体传输后留存的特性, 在两个水印环中间加入一个高能定位环, 并基于此新模板提出了抗屏摄文档水印的自动提取框架, 从而实现了全盲自动提取水印。

2.4 纠错编码与图像预处理算法

2.4.1 BCH 纠错编码

为了增强数字水印的鲁棒性, 通信中的纠错编码技术在数字水印中被广泛使用。通过对数字产品版权信息进行纠错编码生成带有一定自主纠错能力的数字水印信息, 可以提高数字水印在信道中传输的可靠性。在经历信道传输的各种信号处理攻击后, 受干扰的水印信号只要错误数量小于纠错编码的纠错能力, 就能实现对错误信号的纠正, 最终恢复初始的水印信号。BCH 纠错编码^[73]就是一种被广泛使用的线性纠错编码技术, 它是循环码的一个重要子类, 具有纠正数据传输或存储过程中出现的多个错误的的能力。该编码技术是由博斯 (Bose)、查德胡里 (Chaudhuri) 和霍坤格姆 (Hocquenghem) 在 1959 年独立发现的二元线性循环码。BCH 编码的码长 n 与纠错能力 t 之间的关系是: 对于特定的正整数 $m(m \geq 3)$, 存在一个码长为 $n = 2^m - 1$ 的 BCH 码, 能够纠正 t 个随机错误, 其中 t 通常小于或等于 $m/2$ 。而 BCH 编码的生成多项式为:

$$g(x) = m_1(x) \cdot m_3(x) \cdots m_{2t-1}(x) \quad (2.22)$$

其中 t 即为纠错个数， $m_i(x)$ 是最小多项式。为了抵抗屏摄失真造成的水印信号干扰，本文先后采用了 BCH(63,18) 和 BCH(127,36) 两种编码方式，分别可以纠错 10 和 15 位，以实现水印在错误一定比特数的情况下可以纠错为准确的原始水印信息。纠错编码虽然具有纠错能力，但也牺牲了水印容量，在信道传输中的有效信息码长 k 小于 BCH 编码码长 n ，其余的 $n-k$ 个比特位是校验位，其作用是为了确保有效信息的准确传达。但在评价水印的鲁棒性时，在信道传输的实际信息位是码长为 n 的 BCH 编码，所以不应将纠错能力纳入评价指标，而应只考虑屏摄前后 BCH 编码传输和提取的准确率。

2.4.2 自适应直方图均衡化

在数字图像处理中，直方图是一个重要的工具，它展示了图像中每个像素亮度值的分布情况。直方图均衡化是通过改变图像的直方图来改变图像中各像素的灰度，从而可以增强动态范围偏小的图像的对比度的技术。这种方法通过拉伸那些在直方图中高度集中的亮度级别，并压缩那些较少出现的亮度级别让图像的直方图更接近均匀分布，从而增强图像整体的细节和纹理。在传统的直方图均衡化中，整个图像被视为一个整体，通过一个全局的映射函数来调整所有像素的亮度值。输入的直方图用 $H(p)$ 表示，输入的灰度级范围为 $[p_0, p_k]$ ，直方图均衡化就是找一个单调的像素亮度变换 $q = T(p)$ 使得输出直方图 $G(p)$ 在整个输出亮度范围 $[q_0, q_k]$ 是均匀的。其中像素亮度变换 T 的单调性表示为：

$$\sum_{i=0}^k G(q_i) = \sum_{i=0}^k H(p_i) \quad (2.23)$$

其中均衡化的直方图 $G(q)$ 对应于均衡的概率密度函数 f ：

$$f = \frac{N^2}{q_k - q_0} \quad (2.24)$$

对于 $M \times N$ 大小的图片来说，当连续概率密度为理想值时，结合上述两公式可以推导得到：

$$MN \int_{q_0}^q \frac{1}{q_k - q_0} ds = \frac{MN(q - q_0)}{q_k - q_0} = \int_{p_0}^p H(s) ds \quad (2.25)$$

$$q = T(p) = \frac{q_k - q_0}{MN} \int_{p_0}^p H(s) ds + q_0 \quad (2.26)$$

其中公式(2.26)中的积分被称为累积的直方图，在数字图像中用求和来近似，因此实际输出直方图无法满足理想情况。直方图均衡化增强了靠近直方图极大值附近亮度的对比度，减小了极小值附近的对比度。然而，这种方法可能会导致某些区域的过度增强或其他区域的细节丢失。为了解决这个问题，研究者们提出了自适应局部直方图均衡化。此方法将图像分割成多个小区域，然后对每个区域独立地应用直方图均衡化算法。这种图像处理算法可以根据图像每个局部区域的独特亮度分布进行调整，从而更精细地控制对比度的增强，保留图像的局部细节和纹理信息。自适应局部直方图均衡化的核心思想是在每个小区域内计算直方图，并将其变换为累积分布函数（Cumulative Distribution Function, CDF），然后将 CDF 拉伸以增加对比度。通过这种方式，算法能够针对图像中的不同特征和结构进行优化，保留重要的视觉信息，同时提高图像的整体质量。这种方法特别适合处理那些具有不均匀亮度分布的图像，它能够有效地提升图像的局部对比度，使得之前可能被忽略的细节得到突出。

在处理屏摄照片中的水印图像时，自适应直方图均衡化能够有效增强水印图像空域的细节和纹理特征。因为基于幅度调制的 DFT 频谱水印，其频域中频谱系数的修改在空域中表现为相应的纹理特征。在零频移位后的频谱中，大半径环上的高能量系数在空域表现为精细纹理特征，小半径环上的高能量系数在空域表现为粗糙纹理特征。反之，通过增强水印图像的空域纹理特征，也可以增强频域的水印特征，从而有助于提高水印的可检测性和准确性。

2.4.3 维纳滤波与钝化掩蔽

维纳滤波是一种基于统计决策理论的自适应滤波技术，由诺伯特·维纳（Norbert Wiener）在 1942 年提出。它主要用于图像和信号处理中的噪声抑制和信号恢复。维纳滤波的核心思想是最小化原始信号和估计信号之间的均方误差，从而得到最佳的信号估计。维纳滤波考虑了图像的局部噪声和信号的统计

特性，从而在保持图像重要特征的同时实现平滑处理。

钝化掩蔽技术是 20 世纪 30 年代以来印刷和出版业一直使用的图像锐化技术，它是通过从原图像中减去一幅钝化的图像来实现的，可以简单地分为三个步骤：对原始图像进行平滑处理，得到平滑图像；从原始图像中减去平滑图像，得到掩蔽模板；将原始图像与掩蔽模板加权相加，实现钝化掩蔽。令 $\bar{f}(x,y)$ 表示模糊后的图像，则掩蔽模板 $g_{mask}(x,y)$ 为：

$$g_{mask}(x,y) = f(x,y) - \bar{f}(x,y) \quad (2.27)$$

锐化图像 $g(x,y)$ 则为：

$$g(x,y) = f(x,y) + k \cdot g_{mask}(x,y) \quad (2.28)$$

其中，当 $k > 1$ 时，表示的是高提升滤波；当 $k = 1$ 时，表示的便是钝化掩蔽；当 $k < 1$ 时，可以减少钝化模板的贡献。本文在预处理阶段结合了维纳滤波和钝化掩蔽两种技术，将维纳滤波用作平滑处理，然后减去滤波结果，并加上原始图像以完成钝化掩蔽处理，进一步增强图像的空域细节和纹理特征，从而加强频域水印的可检测性。

2.5 本章小结

本章详细介绍了本文使用的相关技术，主要包括数字水印技术、屏幕拍摄失真、DFT 圆环水印、纠错编码与图像预处理算法。数字水印技术部分介绍了数字水印系统框架、盲提取水印和屏摄水印的主要评估指标；屏幕拍摄失真部分详细介绍了屏摄过程中的各种失真，包括相机失真、拍摄失真、屏幕显示失真、光照失真和摩尔纹失真，并仔细分析了各种失真对水印鲁棒性的影响；DFT 圆环水印部分主要介绍了二维离散傅里叶变换及其特殊性质和 DFT 环形水印技术；纠错编码与图像预处理算法介绍了 BCH 纠错编码和两种图像预处理算法：自适应直方图均衡化、维纳滤波和钝化掩蔽。

第三章 抗屏摄鲁棒文档水印算法

3.1 引言

随着数字化时代的到来，数字多媒体的信息安全问题日益受到人们的广泛关注。技术的飞速发展使得人们越来越依赖电脑办公，同时，智能手机的普及也让人们习惯了通过拍照记录信息，并迅速分享到网络中。然而，这种便捷的信息传播方式也为版权保护带来了新的挑战，偷拍屏幕已成为一种简单而普遍的侵犯数字媒体（如图像、文档、视频）版权的手段。为了解决这一问题，研究者们提出了多种数字水印技术，旨在为数字内容提供有效的版权保护和隐私安全。电子文档作为数字多媒体的一种重要形式，在屏摄后尤其容易引起隐私泄露风险。例如，合同、机密报告、个人档案等敏感文档一旦被非法拍摄并传播，将导致严重的信息安全问题。电子文档的数字水印不仅要保证水印的不可见性，确保文档的阅读体验不受影响，还要保持文档的可编辑性，不能简单地将其视为图像来处理。

目前，抗屏摄文档水印算法有基于深度神经网络的文档图像水印、基于质心修改的文档字库水印和基于 DCT 翻转模板的文档底纹水印等，这些方法虽然都有一定的抗屏摄鲁棒性，但也都存在局限性。基于深度神经网络的文档图像水印方案将电子文档视为图像进行处理，其处理的对象是文档转换而成的图像，这种方法将无法保证文档的可编辑性。此外，水印的嵌入是经过神经网络训练完成后直接覆盖在由文档导出的图像之上，水印的隐蔽性较差，同时也可能影响部分文本的正常阅读。基于质心修改的文档字库水印方案通过创建具有特殊字形特征的水印字库来提高对屏摄失真的抵抗能力。为了嵌入不同的水印信息，需要构建多个完整的水印字库，而完整水印字库的构建过程耗时耗力。并且当应用于简单结构的语言（英语）时，其字形变化程度较大，文档的阅读观感也会变差。基于 DCT 翻转模板的文档底纹水印方案选择在文档的底纹中嵌入 DCT 水印模板。这种方法的优势在于，它能够在保持水印文档视觉质量的同时，解决文档的可编辑性问题。然而，由于 DCT 水印的提取依赖于对 DCT 分

块的准确定位，而该方法是利用翻转模板的对称性来进行定位，但这也导致了模板尺寸较大，难以有效应对仅拍摄部分文档的情况。

相较于 DCT 水印模板，DFT 环形水印模板在抵抗几何攻击和裁剪攻击方面表现更为出色。DFT 环形水印的独特优势在于其利用了频域中的独特结构，有效地保护了水印信息免受各种失真的影响。在跨媒体信息传输的研究领域，例如抗打印扫描水印和抗打印拍摄水印技术，DFT 环形水印已证明其卓越的性能。此外，基于 DFT 频域的线性和叠加特性，当多个圆环水印模板通过平铺拼接组合成一个完整的水印底纹时，水印圆环在频域中的尺寸将按照相同的比例放大，这一特性进一步增强了抵抗裁剪攻击的能力。因此，在深入分析现有研究成果、屏摄组合失真及文档水印的必要特性后，本章提出了一种基于 DFT 环形水印的抗屏摄文档底纹水印算法。通过将 DFT 圆环水印模板嵌入在小尺寸的文档底纹中，并使用平铺拼接后的完整水印底纹作为文档的底纹，不仅保证了文档良好的视觉质量和灵活的可编辑性，而且在抗屏摄鲁棒性方面表现出色。并且经实验证明，本方法实现了水印隐蔽性和抗屏摄鲁棒性的良好平衡。

3.2 算法思想

3.2.1 底纹预处理

在数字图像水印技术的研究中，处理彩色图像时，通常会将图像从 RGB 颜色空间转换到 YCbCr 颜色空间，并选取 Y 分量作为水印的嵌入对象。在 YCbCr 模型中，Y 分量代表图像的亮度信息，而 Cb 和 Cr 分量则分别代表蓝色和红色色度信息。由于人眼对亮度变化的敏感度高于对色度变化的敏感度，选择 Y 分量作为水印嵌入的目标，可以在不显著影响图像视觉质量的同时，增强水印的隐蔽性。此外，亮度分量包含了图像的大部分能量信息，对于常见的图像处理攻击，如滤波和压缩等具有较好的鲁棒性，有助于在图像经历各种信号处理后，仍能保持水印的完整性和可提取性。在本章研究中，为了与算法[57]进行对比实验，本算法同样选取了浅绿色 (RGB[199,237,204])、浅黑色 (RGB[40,40,35]) 和白色 (RGB[255,255,255]) 三种纯色背景作为文档底纹。

底纹颜色的选择旨在符合人眼视觉观感，避免使用对比度过高的颜色，以确保使用水印底纹后的文档阅读过程中不会引起人眼视觉上的不适。

基于以上考虑，本文选择上述纯色底纹的 Y 分量作为水印的嵌入对象，以进一步减少水印嵌入对底纹色彩的影响，防止由于色彩变化而导致人眼阅读观感下降。然而，纯色图像在频谱中不包含任何频率成分，这意味着无法直接在其 Y 分量上进行圆环水印的加性嵌入。为了解决这一问题，在水印嵌入前需要对纯色底纹进行预处理，即在 Y 分量中加入轻微的高斯白噪声，从而使得纯色底纹的 Y 分量具有可操作的频率成分，为水印的嵌入提供必要条件。但同时，高斯白噪声的方差取值会影响到水印底纹的视觉质量，因此仅需要添加轻微的高斯白噪声使得 DFT 圆环水印可以有效嵌入。

3.2.2 水印区域定位

现有的抗屏摄水印研究在屏摄照片中对水印区域的定位大都是基于关键点的匹配，但对于布满文字的文档图像，其并没有明确的关键点可以用于检测和匹配。算法^[57]通过构造翻转对称水印模板，在提取时检测对称性来实现定位操作。本文使用小尺寸的 DFT 水印模板平铺拼接为完整的水印底纹，而 DFT 频谱水印的特性可以从另一种角度解决水印区域的定位问题。

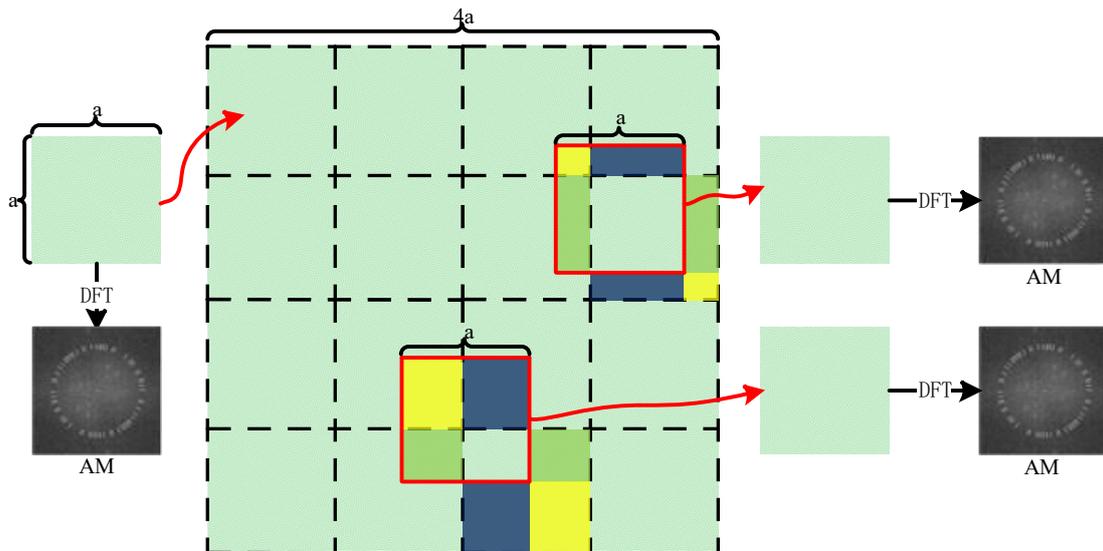


图 3.1 DFT 频谱的平移不变形示意图

1) **频谱的平移不变性：**如图 3.1 所示，当使用 $a \times a$ 大小的水印模板平铺拼接为 $4a \times 4a$ 大小的水印底纹后，因为在 $a \times a$ 大小的水印底纹中嵌入了半径为 r

的圆环水印，此时如果在 $4a \times 4a$ 大小的水印底纹中任意位置截取 $a \times a$ 大小的水印底纹，其可以拆解重组为相邻的 $a \times a$ 大小水印底纹分块。由 2.3.1 节中的公式 (2.9)、(2.10) 和 (2.16) 可知，空域中的平移等同于频域中的相位变化，与幅度无关。因此 DFT 频谱具有平移不变性。所以重组后的水印底纹的频谱并不会发生变化，即频谱中的水印圆环保持不变。

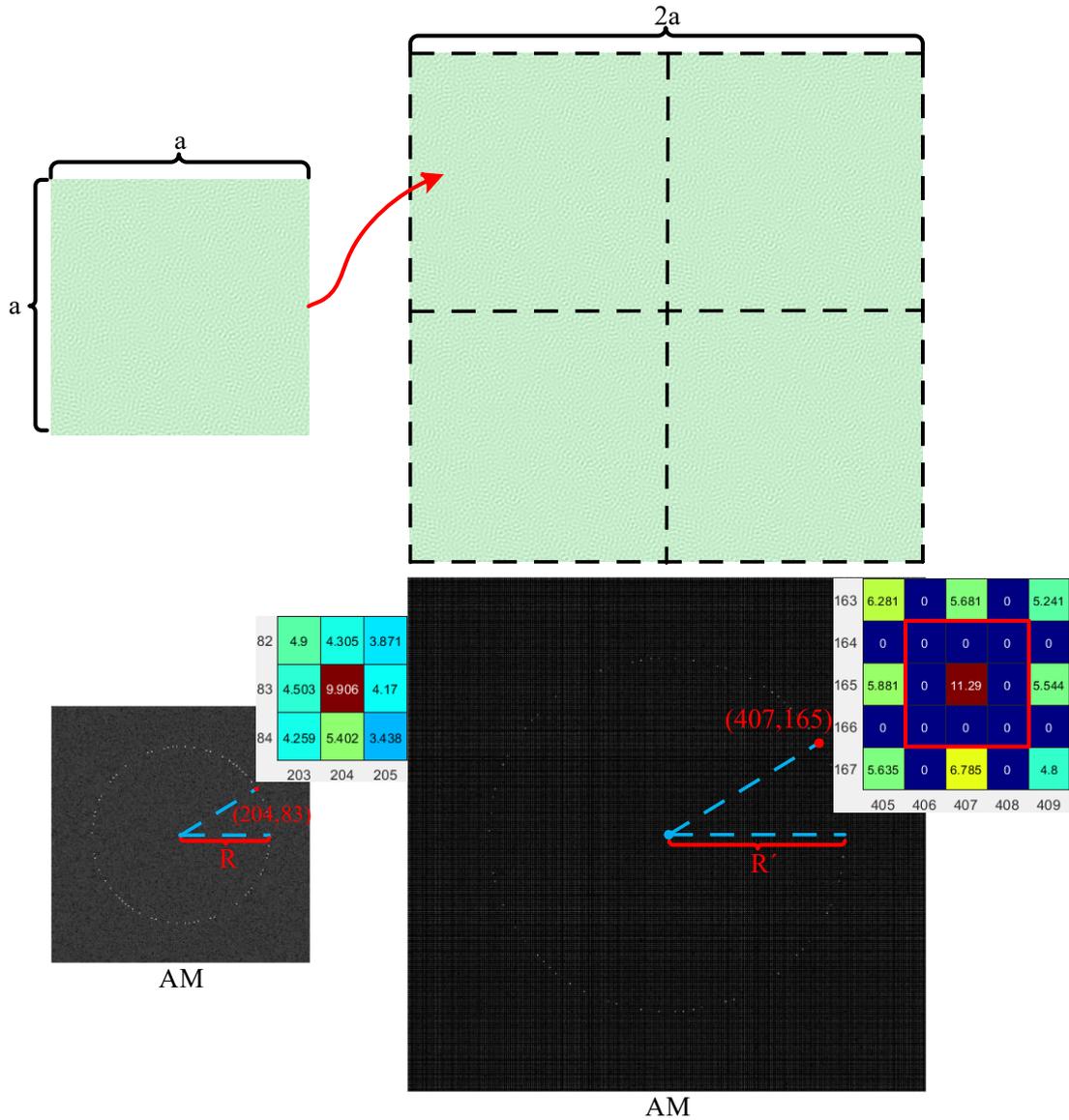


图 3.2 DFT 频谱的线性和叠加特性示意图

2) 频谱的线性和叠加特性: 如图 3.2 所示, 使用 $a \times a$ 大小的水印模板平铺拼接为 $2a \times 2a$ 大小的水印底纹后, 分别对两个水印底纹进行 DFT 变换, 得到:

$$\begin{cases} G(u, v) = F[f(x, y)] \\ G(u', v') = F[f(x', y')] \end{cases} \quad (3.1)$$

其中, (x, y) 和 (x', y') 分别为 $a \times a$ 大小和 $2a \times 2a$ 大小的水印底纹的空域坐标, (u, v) 和 (u', v') 分别为 $a \times a$ 大小和 $2a \times 2a$ 大小的水印底纹对应的频域坐标, 则 $f(x', y')$ 和 $f(x, y)$ 的关系为:

$$f(x', y') = f(x, y) + f(x+a, y) + f(x, y+a) + f(x+a, y+a) \quad (3.2)$$

而由 2.3.1 节中 DFT 变换的周期性公式(2.13)可知:

$$f(x', y') = 4f(x, y) \quad (3.3)$$

再由 2.3.1 节中的线性公式(2.12)可以推导出:

$$G(u', v') = 4G(u, v) \quad (3.4)$$

其中 $u' = 2u$, $v' = 2v$ 。而由于平铺拼接操作没有改变图像的内容, 只是将其复制和扩展, 所以在新图像的幅度谱中, 除了复制的幅度谱点之外, 其他位置的幅度仍然是 0。这是因为这些位置在原始图像中并没有对应的频率成分。而复制的幅度谱点其坐标随着图像的尺寸变大而倍增, 幅度谱系数也随着平铺图像的数量而倍增。原水印圆环半径 R 为:

$$R = \sqrt{\left(u - \frac{a}{2}\right)^2 + \left(v - \frac{a}{2}\right)^2} \quad (3.5)$$

所以拼接后的水印圆环半径 R' 与原水印圆环半径 R 的关系为:

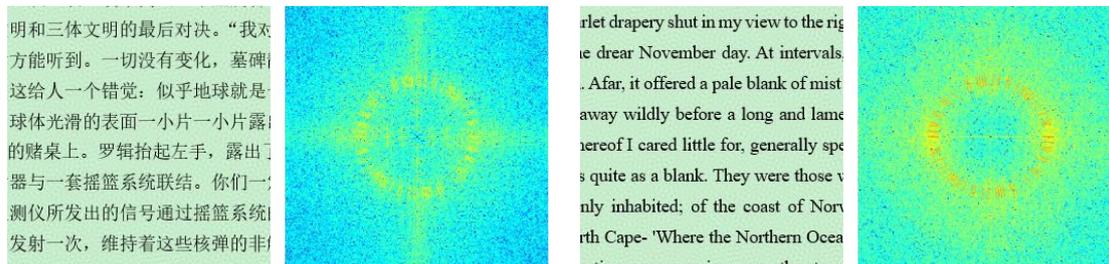
$$R' = \sqrt{(2u - a)^2 + (2v - a)^2} = 2R \quad (3.6)$$

如图 3.2 所示, $a \times a$ 大小的水印底纹的 AM 幅度图中坐标(204,83)的幅度系数的自然对数值为 9.906, $2a \times 2a$ 大小的水印底纹的 AM 幅度图中对应坐标(407,165)的幅度系数的自然对数值为 11.29, 其自然对数值大小刚好增加了 $\log 4$, 即对应原始幅度增加了四倍:

$$\log G(u', v') = \log 4G(u, v) = \log G(u, v) + \log 4 \quad (3.7)$$

结合上述两种性质, 本方法便无需定位具体的水印区域, 因为当截取任意方形尺寸 $b \times b (b > a)$ 的含有水印底纹的文档图像后, 其频域中都含有相同水印信息的水印圆环, 并且其圆环半径为 $R' = R \times b / a$ 。

3.2.3 文本失真



(a) 中文水印文档及其频谱热力图

(b) 英文水印文档及其频谱热力图

图 3.3 文字覆盖后的频谱水印特征示意图

由于在水印底纹上会进行文档内容的编辑，文字是直接覆盖在底纹之上的。文字的覆盖会引入新的信息失真，被文字覆盖的水印底纹区域等同于经历了裁剪攻击。算法[57]通过利用 DCT 水印纹理的对称性对文字区域进行翻转补偿，从而在一定程度上消除文字失真引起的信息损失。而在本算法中，由于 DFT 水印的抗裁剪性，文字覆盖造成的失真对频域中圆环水印的检测和提取影响较小，因此无需针对文字失真进行额外处理。如图 3.3 所示，图 (a) 和图 (b) 分别展示了中英文水印文档和预处理后对应的频谱热力图，可以发现其频谱中的圆环水印特征清晰可见，并且在文字覆盖后的水印提取准确率均为 100%，说明在本算法中文字失真对水印的准确提取并无影响。

3.3 基于 DFT 环形水印的文档底纹水印方案

3.3.1 水印嵌入流程

水印的嵌入流程分为两个步骤：水印模板的生成和水印模板的嵌入。水印模板的生成首先需要对二进制消息序列做预处理，使用 BCH 纠错码对原始消息序列进行编码得到最终嵌入的水印序列 \mathbf{W} ，然后再根据消息序列生成圆环水印模板 \mathbf{C} 。水印模板的嵌入过程是将 α 倍的水印模板 \mathbf{C} 与底纹 \mathbf{Y} 分量的频谱图 \mathbf{AM} 做加性嵌入，其中 α 表示嵌入强度。然后经历 DFT 反变换和 YCbCr 到 RGB 的色彩空间变换，得到最终的水印底纹。

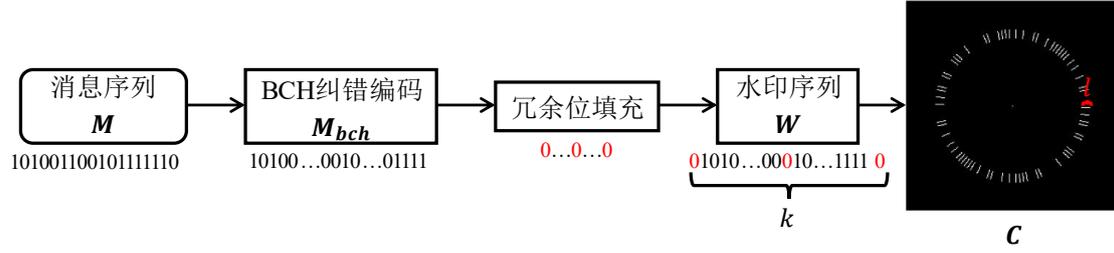


图 3.4 水印模板生成流程图

1) **水印模板生成**: 水印模板的生成流程如图 3.4 所示。首先, 对二进制消息序列 M 使用 BCH 纠错编码, 以生成具有纠错能力的二进制序列 M_{bch} 。为了避免由于底纹图像拼接时边界的水平和垂直不连续性所引起的干扰, 需要对序列做冗余位填充。在周期性拼接水印底纹的过程中, 底纹边界处的不连续性会在中心化频谱的中心区域形成明亮的十字型条纹。本文选择不在中心处的 x 轴和 y 轴邻近区域嵌入水印信息, 即在 x 轴和 y 轴的邻近区域嵌入 0 比特的冗余信息, 以此来规避图像边界不连续性造成的干扰。

本方案选用 $a \times a$ 尺寸的纯色底纹用于水印的嵌入, 因此需构建一个相同大小的全零矩阵 C 作为水印模板。鉴于 DFT 频域的中心对称性, 水印信息实际是嵌入在中心化频域的第一和第二象限中, 即圆环水印的水印信息实际是在上半圆环中。通过修改半径为 R 的圆环上特定坐标的值为消息序列 M 中相应的 0 或 1, 便可以将二进制消息序列 M 以圆环的形式嵌入到水印模板 C 中。圆环的坐标计算如下:

$$\begin{cases} x_1 = (x_{center}) + R \cdot \sin(\theta) \\ y_1 = (y_{center}) - R \cdot \cos(\theta) \end{cases} \quad (3.8)$$

对于一二象限的圆环坐标 (x_1, y_1) , 则可以通过以下关系得到三四象限的对称坐标 (x_2, y_2) :

$$\begin{cases} x_2 = 2 \cdot (x_{center}) - x_1 \\ y_2 = 2 \cdot (y_{center}) - y_1 \end{cases} \quad (3.9)$$

其中 (x_{center}, y_{center}) 是水印模板的中心坐标, 而 θ 表示水印信息嵌入时的圆环角度。通过修改嵌入角度 θ , 将坐标设置为对应水印序列 W 中的水印比特:

$$C_{(x,y)} = W_i, \quad i \in [1, k] \quad (3.10)$$

其中 W_i 代表水印序列 W 第 i 位比特信息，其值为 0 或 1， k 表示水印序列 W 的长度。因此，每一比特水印信息的间隔角度为 π/k 。为了增强水印的鲁棒性，本算法还对水印圆环进行了扩频处理，即通过嵌入多个水印圆环来重复嵌入水印信息。每比特水印信息由一段线段表示，扩频线段的长度设为 l ，则圆环水印的半径 $R \in [R_c - l/2, R_c - 1 + l/2]$ ，其中 R_c 为扩频水印圆环的中心半径。通过这种冗余性设计，便可以得到一个多重圆环水印模板 C 。

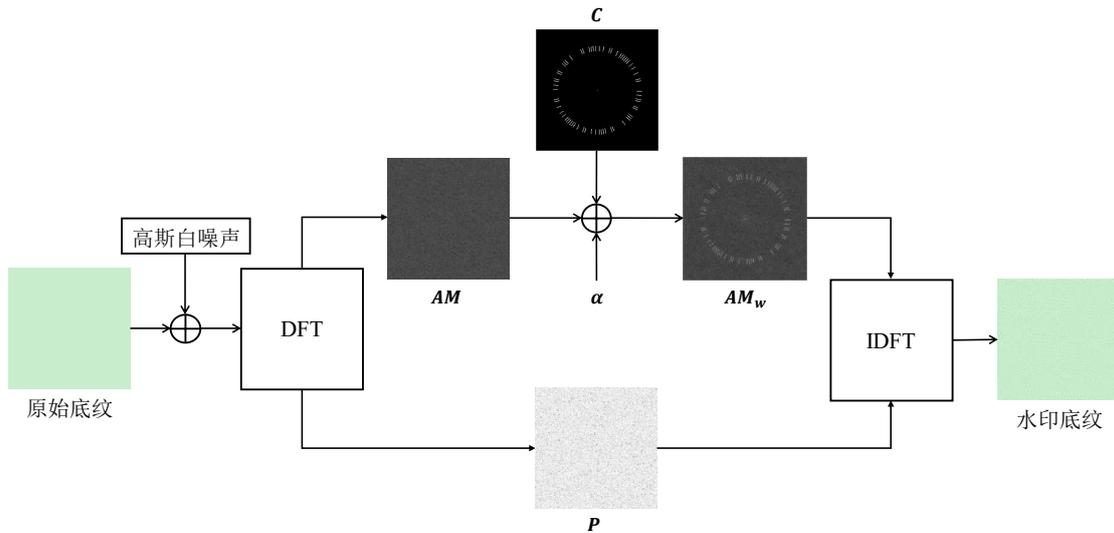


图 3.5 水印模板嵌入流程图

2) 水印模板嵌入：水印模板的嵌入流程如图 3.5 所示。首先，将原始底纹图像从 RGB 色彩空间转换为 YCbCr 色彩空间，并提 Y 分量作为水印嵌入对象。由于纯色底纹的频谱不具备可修改的频率成分，所以需要在原始底纹的 Y 分量上添加方差为 0.0005 的轻微高斯白噪声，不会过于影响原始底纹的视觉质量，并同时满足水印加性嵌入的必要条件。随后对加噪后的 Y 分量执行 DFT 变换，得到频谱图 AM 和相位图 P 。对频谱图 AM 进行零频移位处理，确保直流分量位于频谱图的几何中心。随后，将水印模板 C 以加性调制的方式嵌入到频谱图 AM 中，通过控制嵌入强度 α ，便可以得到水印频谱图 AM_w ：

$$AM_w = AM + \alpha \cdot C \quad (3.11)$$

其中， α 表示为水印嵌入的强度参数。将水印频谱图 AM_w 与原始相位图 P 重新结合，并执行 IDFT 变换，得到含水印的亮度分量 Y_w 。然后将含水印的 Y_wCbCr 图像重新转换为 RGB，便得到最终的彩色水印底纹。

3.3.2 水印提取流程

水印的提取流程分为三个步骤：透视矫正、图像预处理、水印提取和纠错解码。如图 3.6 所示，首先需要将屏摄照片中已透视变形的文档图像矫正回原平面。然后对矫正后的文档图像进行图像预处理，增强水印底纹的纹理特征并滤除低频噪声。最后使用水印提取算法提取图像的频域水印信息。

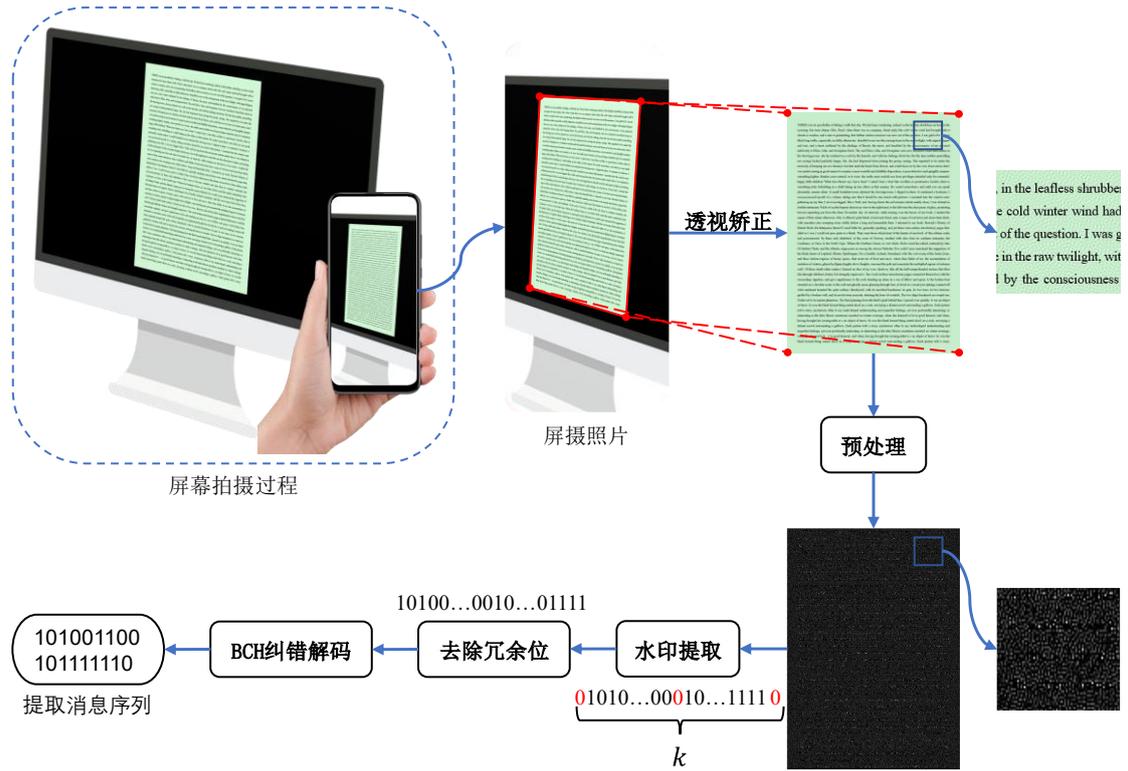


图 3.6 水印提取流程图

1) **透视矫正**：在屏摄过程中，相机拍摄的是屏幕上的电子文档，其自身保持平整状态未发生形变，但由于拍摄角度导致相机成像平面与屏幕平面不平行而产生了透视倾斜变形。在得到屏摄照片后，首先便需要对照片中透视变形的文档图像进行线性化矫正处理。本方法采用的是基于轮廓提取的矫正算法，通过获取目标轮廓的最小外接矩形，即手动标定照片中文档图像的四个顶点，并定义矫正后图像的四个顶点，通过对应的单应矩阵对图像进行透视变换以达到矫正目的。若 (x, y) 是失真文档图像上的其中一个顶点的坐标， (x', y') 为矫正后图像中对应点的坐标，则可以使用一个双线性方程来表达：

$$\begin{cases} x' = c_1x + c_2y + c_3xy + c_4 \\ y' = c_5x + c_6y + c_7xy + c_8 \end{cases} \quad (3.12)$$

已知变形文档图像的四个顶点坐标，便可以得到如上的 8 个方程，从而可以解出对应的 8 个变量 $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ ，得到矫正模型。本方案在定义矫正后图像的四个顶点时，需要得知原始文档图像的先验尺寸，这是因为频谱中的水印圆环会随着尺寸的变化而形变为椭圆环，会影响后续的水印提取。

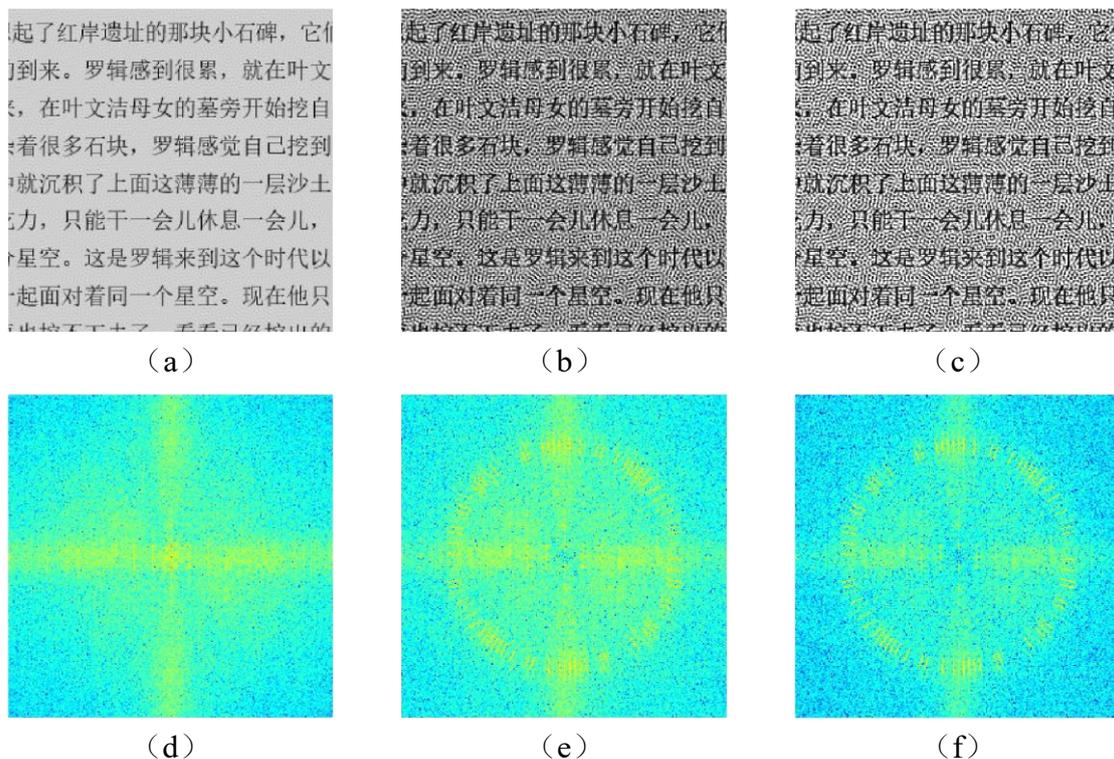


图 3.7 图像预处理前后的 Y 分量和频谱图热力图

2) 预处理: 在屏摄过程中，除透视和几何失真外，相机失真和屏幕显示失真也会造成图像的显示失真。为了减少这种显示失真以及增强水印的频谱特征，需要对矫正后的屏摄图像进行预处理。在得到矫正后的文档图像后，首先需要提取其 Y 分量作为预处理对象，图 3.7 (a) 便展示了预处理前的 Y 分量。图 3.7 (d) 则为对应的频谱热力图，此时频谱中的无明显水印环特征。预处理分为两步，第一步是自适应直方图均衡，可以增强文档空域的底纹纹理特征，进而加强其频域对应的水印特征。图 3.7 (b) 即为自适应直方图均衡处理后的 Y 分量，可以发现其底纹纹理特征得到了充分强化。图 3.7 (e) 则为对应的频谱热力图，此时可以清晰地观察到频谱中的水印圆环特征，但低频区域同时也包含了大量高能噪声。因此预处理的第二步是使用维纳滤波与钝化掩蔽相结合的方法进行低频滤波和锐化处理。与抗屏摄视频水印算法^[74]的预处理相同，本

算法同样使用原始维纳滤波器，并从 Y 分量中减去滤波结果，得到掩蔽模板。不同的是，本算法需再加上初始的 Y 分量，以实现对其钝化掩蔽处理，便可得到如图 3.7 (c) 所示的 Y 分量以及图 3.7 (f) 所示的频谱热力图。此时可以发现频谱中的低频噪声已被滤除，水印圆环特征也变得更为清晰。

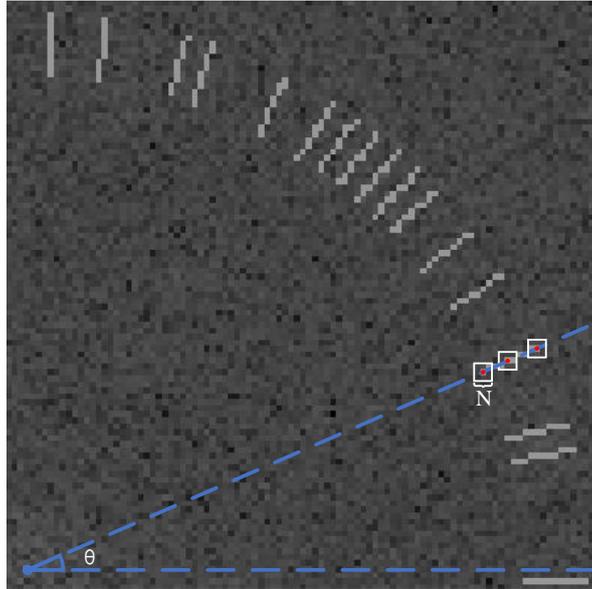


图 3.8 水印序列提取原理示意图

3) 水印提取: 预处理后的 Y 分量便可以应用水印提取算法提取水印信息。因为文档底纹是由小尺寸的水印底纹模板拼接而成，根据 DFT 频谱的线性和叠加特性，文档图像在透视矫正恢复至原尺寸后，其频谱中的水印圆环可能会形变为椭圆环。通过将 Y 分量缩放拉伸为 $b \times b$ 尺寸的正方形，便可以将频谱中的椭圆环水印拉伸回正圆环。此时，水印圆环的半径 r 可表示为：

$$r \in \left[r_c - \frac{l}{2} \cdot t, r_c - 1 + \frac{l}{2} \cdot t \right], t = \frac{b}{a} \quad (3.13)$$

其中， r_c 代表扩频水印圆环的中心半径，且 $r_c = R_c \cdot t$ 。图 3.8 展示了水印序列的提取原理，即按照相同的间隔角度 π/k 提取半径 r 在 N 范围内的最大值 max_θ ，并计算半径 r 上所有幅值的均值 $mean_r$ 。则水印信息 w_θ 的提取公式为：

$$w_\theta = \begin{cases} 1 & \text{若 } max_\theta \geq mean_r \\ 0 & \text{若 } max_\theta \leq mean_r \end{cases} \quad (3.14)$$

其中， N 范围内的最大值 max_θ 是指以角度为 θ 半径为 r 的坐标为中心， $N \times N$ 范围内的最大幅值。 N 的值至少为 3，但当半径倍增时应相应增加，以增加水

印比特信息提取时的鲁棒性，因此其应设定为当半径为 r 时，相邻的两比特水印之间弧长的 $2/3$ ：

$$N = \begin{cases} \frac{2\pi r}{3k} & \text{若 } \frac{2\pi r}{3k} > 3 \\ 3 & \text{否则} \end{cases} \quad (3.15)$$

其中， k 为水印圆环嵌入的水印比特数。在水印提取过程中，实际遍历并提取了 $l \cdot t$ 个圆环水印序列。最终，通过对比不同水印序列，选择每个信息位中出现频率最高的水印比特，并将选定的比特组合成最终的提取水印序列 w 。

4) 纠错解码：在得到提取水印序列 w 后，首先需要去除嵌入时添加的冗余位，即将 x 轴和 y 轴的 0 比特信息位除去，便可以得到待解码的消息序列 m_{bch} 。最后使用 BCH 纠错解码便可得到最终的提取消息序列 m 。

3.4 实验结果与分析

3.4.1 实验参数

本小节主要介绍本方法的实验参数，包括实验场景的设备参数设置，和算法中所使用的参数，并且通过设计定量统计实验探讨参数的最优解。

1) 设备参数：为了验证此方法对硬件设备的普适性，本文使用分辨率为 1080p 的显示器 ‘DELL-E2213H’ 显示水印文档，后置摄像设为 12MP 的 ‘Redmi K30s Ultra’ 的手机来拍摄屏幕。直接使用 WORD 软件载入水印底纹，纸张大小设为 A4 纸大小，字体大小设为默认值 10.5 磅，行间距也设为默认值 1.15 倍。分别使用中文文档和英文文档进行实验，其中中文文档节选自《三体》的某一段落，英文文档节选自《简爱》的某一段落。在后面的实验中统一使用 DocCN 来表示中文文档，DocEN 来表示英文文档。

2) 底纹模板参数：为了保证局部文档在拍摄后仍然可以提取水印信息，本方法将底纹模板的尺寸设为 256×256 ，即 $a = 256$ ，此尺寸可以保证水印信息的顺利提取，并且仅包含少量的文字信息，不会造成过大的隐私泄露风险。纠错码选择 BCH(63,18) 编码，纠错能力为 10 位，实际嵌入消息长度为 18 比特，纠

错编码序列长度为 63 比特，添加冗余位后水印序列长度 k 为 66 比特。

本方案选择在频谱的中频区域嵌入水印圆环，水印圆环的半径参数是本算法的关键参数，对水印文档的视觉质量、抗屏摄鲁棒性和抗压缩鲁棒性都会产生影响。因此，为了选择最优的半径参数，本节将进行定量分析实验，综合考虑以上三个指标来选择最优的圆环半径。因为每比特水印信息的提取范围至少为 3×3 ，也即两水印比特之间的间隔距离至少大于 2。经实验验证，本算法的水印半径最小可设为 48，可以满足两比特水印之间留有足够的空间用以提取最大幅值，并且能够准确提取所有的水印比特信息。水印圆环的扩频长度 l 设为 10，通过增加水印圆环的数量，增加冗余量，牺牲一定程度的水印文档的视觉质量，以提高水印提取时的准确率。因此扩频水印圆环的中心半径 R_c 分别选取 55、65、75 和 85 作为实验参数，则多重水印圆环的半径 $R \in [R_c - 5, R_c + 4]$ 。嵌入强度 α 统一设为 7000，可以使得水印文档图像的 PSNR 值在 40dB 左右。使用浅绿色底纹生成水印底纹模板，并分别检测使用相应水印底纹的中文文档的 PSNR 值，在正面不同拍摄距离时的水印比特错误数，以及应对不同质量因子的 JPEG 压缩后的水印比特错误数。

实验结果如表 3.1 所示，可以发现在使用相同的嵌入强度时，水印圆环半径的大小对水印文档的视觉质量的影响较小，但对抗屏摄鲁棒性和抗压缩鲁棒性影响较大。由表 3.1 的数据可知，使用不同水印半径生成的水印底纹文档经过正面不同距离拍摄后，在正面 35 厘米和 55 厘米处拍摄时的水印提取错误比特数都较高，这是因为在这两个距离拍摄时均出现了摩尔纹干扰。而在正面 95 厘米处拍摄时，可以发现随着水印半径的增大，水印提取错误比特数随之增加，因为水印提取前的预处理步骤会使用维纳滤波对屏摄图像进行平滑处理，将滤除部分中高频信息，而水印圆环半径的大小对应底纹频谱中的低频、中频和高频区域，因此水印半径的增大会降低本方案的抗屏摄鲁棒性。同理，由表 3.1 中使用不同质量因子的 JPEG 压缩后的水印文档图像的水印提取错误数可知，随着水印半径的增大，水印圆环由中频区域向高频区域移动，其对 JPEG 压缩的抵抗性也相应下降。综合考虑以上三个指标，当选取水印中心半径 $R_c = 55$ 时，才能同时兼顾水印图像的视觉质量、抗屏摄鲁棒性和抗压缩鲁棒性。

表 3.1 不同圆环半径的中文文档的视觉质量和鲁棒性评估

不同指标 \ 圆环半径	$r_c = 55$	$r_c = 65$	$r_c = 75$	$r_c = 85$
PSNR	39.77	39.67	39.66	39.72
SSIM	0.991	0.989	0.989	0.989
正面 35cm	4	4	5	5
正面 55cm	4	6	6	8
正面 75cm	1	5	1	2
正面 95cm	2	3	6	11
JPEG 压缩 (Q=60)	0	0	1	3
JPEG 压缩 (Q=50)	0	1	5	7
JPEG 压缩 (Q=40)	0	1	8	8

据此，本章将水印圆环的中心半径 R_c 设为 55，则多重圆环水印的半径 $R \in [50, 59]$ 。因为水印的嵌入强度 α 与水印底纹的视觉质量成反比，但与屏摄鲁棒性成正比，也即嵌入强度越大，视觉质量越差，屏摄鲁棒性越强。经实验验证，最终选取 $\alpha = 7000$ 作为水印嵌入强度，此时水印文档图像的 PSNR 值在 40dB 左右，并拥有良好的抗屏摄鲁棒性，在两者之间取得了良好的平衡。

3.4.2 文档图像的视觉质量

图 3.9 所示为本章算法与抗屏摄神经网络文档水印算法[34]和抗屏摄 DCT 底纹文档水印算法[57]的视觉质量对比。图 3.9 (a) 展示了三种纯色底纹下的原始文档；图 3.9 (b) 和 (c) 分别展示了使用方法[34]和[57]生成的三种水印底纹的文档图像。其中[34]将文档作为图像训练水印的嵌入，水印直接覆盖在文档图像上，隐蔽性较差，视觉质量也较差。而[57]提出的 DCT 水印底纹和本章提出的 DFT 水印底纹，文本覆盖在水印底纹之上，不会影响文本内容的正常阅读和编辑。其中本章设计的水印底纹几乎不可见，水印的隐蔽性也最强，文档图像呈现的视觉质量也最好。为了客观评估水印文档的视觉质量，本章选择使用峰值信噪比 (PSNR) 和结构相似度指数 (SSIM) 作为度量标准。表 3.2 详细展示了不同方法下中英文文档具体 PSNR 和 SSIM 值。由于 Ge 等人提出的方法仅使用了白底黑字的文档图像，所以表中仅展示其白色底纹的文档图像的 PSNR 和 SSIM 值。从表 3.2 中可以看出本文算法的 PSNR 值和 SSIM 值均优于其余两种算法。其中三种颜色的水印底纹的 PSNR 值均在 40dB 左右，SSIM 值也均在 0.9 以上，表明水印文档图像与原始文档图像有着相当高的相似度。



图 3.9 不同水印算法的中文文档图像视觉质量对比

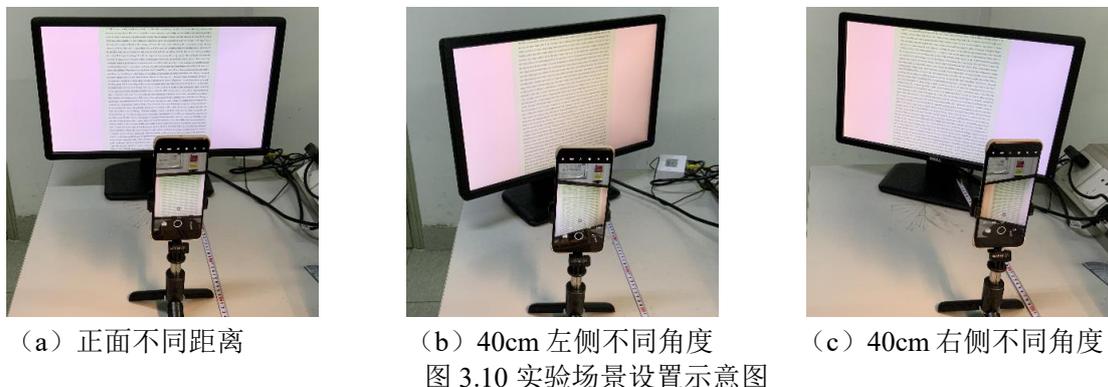
表 3.2 不同水印方案的中英文文档的视觉质量评估

方法		底纹颜色	绿色底纹		黑色底纹		白色底纹	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DocCN	Ge 等 ^[34]	/	/	/	/	25.62	0.817	
	Fang 等 ^[57]	30.26	0.927	30.29	0.802	33.11	0.885	
	本章方法	39.77	0.991	39.96	0.958	42.79	0.983	
DocEN	Ge 等 ^[34]	/	/	/	/	22.88	0.724	
	Fang 等 ^[57]	30.32	0.921	30.30	0.753	33.18	0.859	
	本章方法	39.85	0.990	39.95	0.946	42.89	0.979	

3.4.3 水印的嵌入容量及鲁棒性分析

本节主要评估本章提出的水印方案的嵌入容量以及对屏幕拍摄攻击的鲁棒性。本章使用的是 BCH(63,18)纠错编码，纠错能力为 10 比特，实际传输的原始信息比特数为 18，也即本章方案的水印嵌入容量为 18 比特，牺牲了部分嵌入容量以换取更好的鲁棒性。为了确保鲁棒性实验的公平性，实验中使用的硬件设备是统一不变的，并且拍摄条件也保持一致。如图 3.10 所示，在室内亮度不变的情况下使用支架固定手机，并使用直尺标定拍摄距离，拍摄角度的调整则通过反向等角度旋转显示器来实现。因为纯白色底纹的 Y 分量为 255，相同嵌入

强度的情况下其空域的修改程度会减弱，从而导致屏摄鲁棒性也会减弱。因此，本实验仅选取浅绿色（RGB[199,237,204]）和浅黑色（RGB[40,40,35]）作为实验的文档底纹。屏摄鲁棒性则使用比特提取准确率 BAR 作为衡量指标。



1) 距离鲁棒性实验：图 3.11 展示了浅绿色水印底纹的文档图像在正面不同距离拍摄后的屏摄照片和对应的矫正图像。表 3.3 则展示了正面不同距离拍摄后水印比特提取准确率。本章使用的是 BCH(63,18)纠错编码，加上冗余位后实际传输水印比特数为 66，因此当比特提取准确率高于 84.8%时就可以解码恢复原始消息序列。相较 Fang 等人^[57]的方法，本章提出的方法在视觉质量更为优秀的同时比特提取准确率更高，可以保证在距离 1m 远处拍摄时仍能准确提取水印信息。Ge 等人^[34]的方案并不稳定，在相同的实验条件下，其在正面 45 厘米处拍摄时会出现摩尔纹干扰，水印提取成功率会严重下降，说明此方法无法有效抵抗摩尔纹干扰。而本方案对摩尔纹失真具有一定的鲁棒性。如图 3.11 所示，正面 35 厘米处拍摄的屏摄照片中出现了大量摩尔纹，此时本算法的比特提取准确率也相应下降，但仍在纠错能力范围内，可以保证原始消息的准确提取。

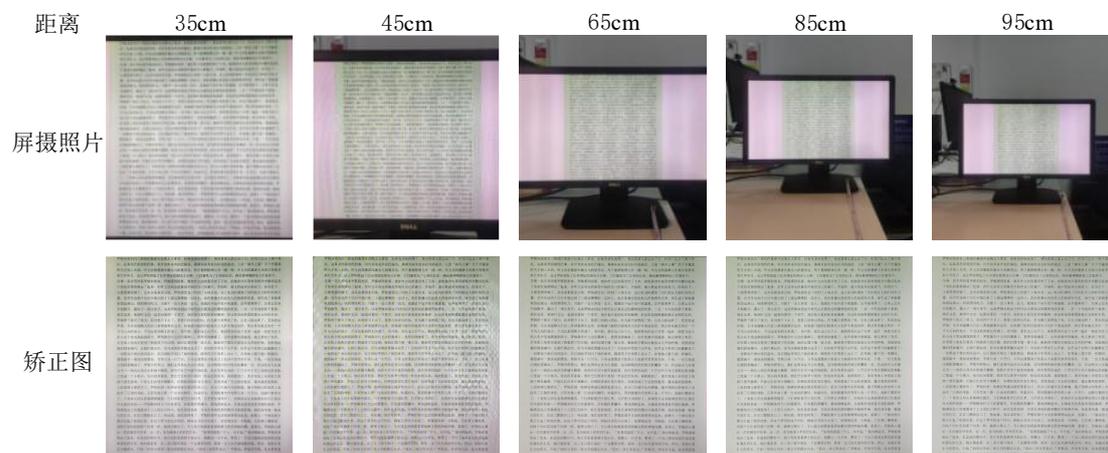


图 3.11 正面不同距离拍摄实验展示

表 3.3 正面不同距离拍摄的水印比特提取准确率 (%)

距离 (cm)		35	45	55	65	75	85	95
DocCN	Ge 等 ^[34]	90.2	60.9	88.5	96.1	95.2	91.1	93.3
	Fang 等 ^[57]	90.6	75.6	85.9	93.7	92.8	92.5	82.5
	绿色底纹	94.5	99.2	97.6	99.9	99.6	97.6	99.2
	黑色底纹	99.2	98.4	97.6	97.6	98.4	96.9	96.1
DocEN	Ge 等 ^[34]	99.3	52.3	97.2	99.4	99.4	99.1	99.3
	Fang 等 ^[57]	92.2	92.9	94.0	92.8	92.2	92.8	89.5
	绿色底纹	93.7	98.4	97.3	97.7	98.4	97.6	98.1
	黑色底纹	99.2	98.4	97.6	98.4	97.2	98.6	97.8

2) 水平角度鲁棒性实验: 图 3.12 展示了浅绿色水印底纹的文档图像在 40 厘米处, 水平部分角度拍摄后的屏摄照片和对应的矫正图像。表 3.4 展示了 40 厘米处水平不同角度拍摄后的水印比特提取准确率。由表中数据可知, 本方法在不同角度拍摄后的中英文文档的平均提取准确率分别达到了 98.3%和 97.6%。并且在左 65° 或右 65° 的大偏转角度下, 提取准确率也都保持在 95%以上。因为即使在大偏转角度下, 远离相机一侧的底层纹理特征丢失, 但由于靠近相机一侧的底层纹理特征得到保留, 所以仍然可以准确解码恢复嵌入的原始信息。

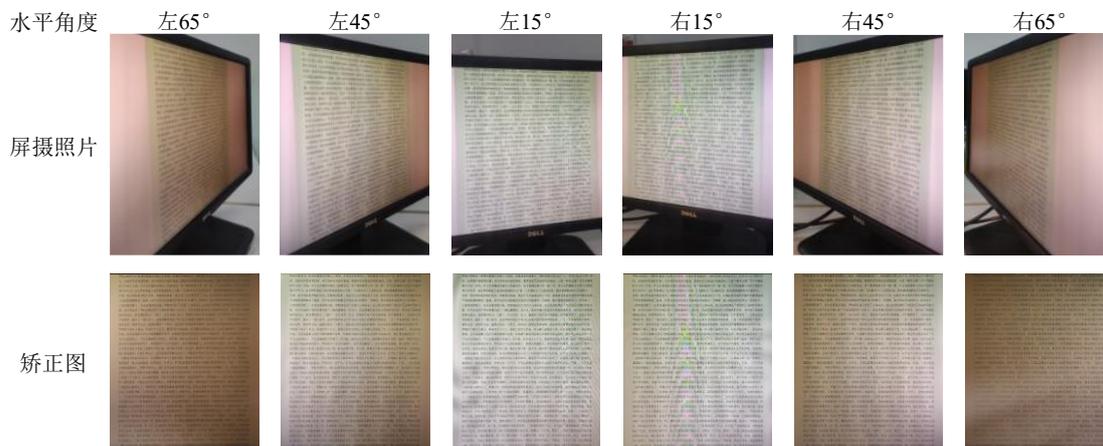


图 3.12 40cm 处水平不同角度拍摄实验展示

表 3.4 40cm 处水平不同角度拍摄的水印比特提取准确率 (%)

水平角度 (°)		左 65	左 60	左 45	左 30	左 15	0	右 15	右 30	右 45	右 60	右 65
Doc CN	Ge 等 ^[34]	83.2	54.9	72.8	64.5	71.7	75.3	84.2	89.7	96.1	96.4	96.1
	Fang 等 ^[57]	87.5	75.1	90.2	85.2	88.2	91.4	90.8	88.2	87.2	91.4	80.5
	绿色底纹	96.1	96.1	98.4	99.2	96.1	98.4	98.4	98.4	99.2	98.4	97.6
	黑色底纹	98.4	98.4	99.2	99.2	99.2	98.4	99.2	99.2	98.4	98.4	98.4
Doc EN	Ge 等 ^[34]	98.4	69.1	54.2	98.1	92.1	99.2	98.1	97.1	54.1	99.2	97.5
	Fang 等 ^[57]	88.1	95.0	92.7	92.4	93.4	93.2	92.4	92.6	92.6	92.2	86.5
	绿色底纹	96.9	96.9	96.9	99.2	94.5	96.1	97.6	96.1	97.7	96.1	96.9
	黑色底纹	98.4	99.2	98.4	99.2	98.4	99.2	97.6	98.4	96.9	97.6	99.2

3) 垂直角度鲁棒性实验: 图 3.13 展示了浅绿色水印底纹的文档图像在 40 厘米处, 垂直部分角度拍摄后的屏摄照片和对应的矫正图像。表 3.5 展示了 40 厘米处垂直不同角度拍摄后的水印比特提取准确率。由表中数据可知, 本方法在不同垂直角度拍摄后的中英文文档的平均提取精度分别为 97.8%和 98.3%, 并且在上 65° 或下 65° 的大偏转角度下的提取精度也都在 95%以上。本方法在应用于中文文档时, 鲁棒性显著优于其余两种方法, 这是因为它们对摩尔纹干扰的抵抗性较差。如图 3.13 中偏上 15° 拍摄时出现了明显的摩尔纹, 此时比特提取准确率略有下降, 但仍然可以准确解码恢复嵌入的原始消息序列。

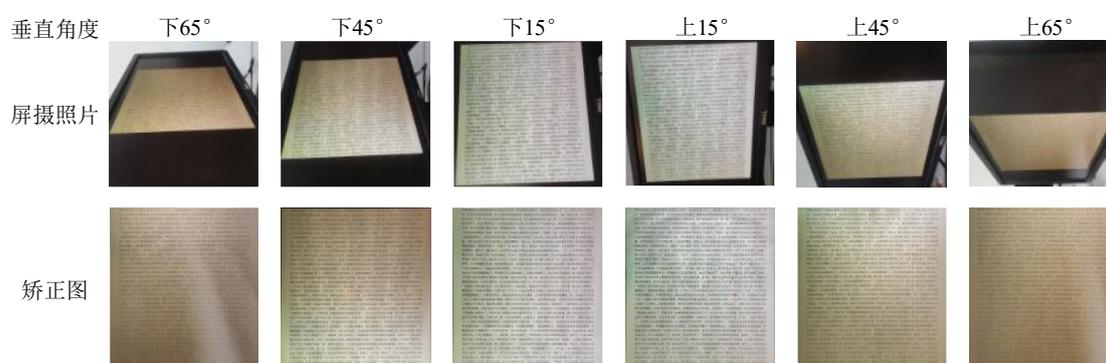


图 3.13 40cm 处垂直不同角度拍摄实验展示

表 3.5 40cm 处垂直不同角度拍摄的水印比特提取准确率 (%)

垂直角度 (°)		下 65	下 60	下 45	下 30	下 15	0	上 15	上 30	上 45	上 60	上 65
Doc CN	Ge 等 ^[34]	94.6	93.1	79.3	64.8	95.7	75.2	74	92.7	97.5	97.8	96.2
	Fang 等 ^[57]	84.6	91.0	92.5	82.1	91.1	88.2	80.4	91.4	88.2	88.3	85.2
	绿色底纹	96.1	97.6	97.6	94.5	97.6	98.4	94.5	95.3	98.4	97.6	97.6
	黑色底纹	95.3	98.4	99.2	99.2	99.2	98.4	99.1	99.2	99.5	99.3	98.4
Doc EN	Ge 等 ^[34]	98.2	99.2	98.1	99.2	98.2	99.7	97.6	95.5	99.3	98.2	79.1
	Fang 等 ^[57]	86.1	92.1	93.2	92.6	92.4	93.9	93.2	92.5	93.5	93.0	87.3
	绿色底纹	95.3	96.1	96.9	98.4	98.4	98.4	97.6	99.2	97.6	97.6	98.4
	黑色底纹	99.2	99.7	99.2	99.2	99.2	98.4	97.6	99.2	98.4	99.2	99.2

3.5 本章小结

本章针对屏幕拍摄过程中可能出现的五种主要失真: 相机失真、拍摄失真、屏幕显示失真、光照失真和摩尔纹失真, 提出了基于 DFT 环形水印的抗屏摄文档底纹水印算法。该算法不仅可以保持良好的视觉质量, 而且对屏摄组合失真也拥有显著的鲁棒性。在深入分析了 DFT 频谱的特性后, 本章采用小尺寸

的水印模板拼接组成完整的水印底纹。该方法并非简单地重复嵌入水印单元，而是利用 DFT 频谱的线性和叠加特性，使频谱中的多个水印单元叠加增强为一个更大的水印单元，从而在水印提取时不再需要定位操作。因此，本方法具有卓越的抗裁剪性能，可以有效抵抗文本失真、光照失真和摩尔纹失真。此外，将水印信息嵌入至文档底纹中，可以保留文档的可编辑性，同时对文档的阅读观感也影响较小。经过大量多距离多角度的拍摄实验验证，本章提出的算法在抗屏摄失真方面具有良好的鲁棒性，且性能优于现有的抗屏摄文档水印算法。

第四章 抗屏摄文档水印自动提取方法

4.1 引言

数字水印技术根据提取过程中对先验信息依赖程度的不同，可以分为盲提取水印、半盲提取水印和全盲提取水印。在屏幕拍摄这一跨媒介传输过程中，含有水印的原始数字媒体文件会被转换成照片这种新形式的数字媒体文件。这些照片中不仅包含水印载体，还可能掺杂着其他背景元素，例如显示器本身、桌面上的杂物和房间陈设等。照片中记录的水印载体实际上是屏幕显示的内容在相机镜头下的几何投影，这转换一过程会受到拍摄距离和角度的影响，从而引起不同程度的透视变形。目前，多数的抗屏摄水印方案在进行水印提取阶段，普遍需要人工介入以确定照片中水印载体的准确位置并进行必要的透视矫正操作。同时，有的方案在这矫正时还需依赖于图像的先验尺寸信息。鉴于这些依赖性，这些水印方案不属于盲提取水印。针对上述挑战，本章提出了一种抗屏摄全盲文档水印自动提取系统，该系统能够在无需人工干预的情况下，直接从照片中自动提取水印信息。

本文在第二章中详细分析了屏摄过程中可能引入的五种失真：相机失真、拍摄失真、屏幕显示失真、照明失真和摩尔纹失真。拍摄者在记录数字文档时，通常会有选择性地拍摄文档中的关键部分，这导致文档内容在屏摄时就有可能没有被完整记录。此外，照明失真和摩尔纹失真也会在文档的某些区域造成严重的信息损失，这种损失在效果上与裁剪攻击引起的信息缺失相似。因此，可以认为这些失真对文档完整性的影响与裁剪攻击有同等的破坏性。鉴于此，一个旨在具有高抗屏摄鲁棒性的文档水印算法，必须具备对裁剪攻击以及类似效果失真的高度抵抗能力。本章对第三章提出的 DFT 圆环水印模板进行了重新设计，将单一扩频水印圆环扩展为两个，以提升水印容量，并在这两个圆环之间嵌入了一个高能定位圆环。该定位圆环因其高能量特性，不仅能够实现对照片中水印文档的自动定位，还支持水印的盲提取。在水印嵌入端，依照第三章的方法，将 DFT 圆环水印模板嵌入至小尺寸的文档底纹中，然后通过平铺

拼接以构成完整的水印底纹，这样做既保留了文档的可编辑性，也增强了其对屏摄攻击的鲁棒性。而在水印提取端，本章将流程细化为三个步骤：自动定位、自动矫正和自动提取。通过在 DFT 频谱的中频段嵌入水印圆环，构造出高能量的频带特征，此特征在屏摄后也能被有效地识别，进而实现水印文档在照片中的自动定位。随后，利用文档中垂直或水平对齐的文本行和线段的几何特征，对定位后的文档图像执行自动透视矫正。最终，通过检测并拉伸矫正图像中变形的定位椭圆环，调整图像频谱尺寸，并比较定位环与水印环在相同角度下的幅值大小，从而精确提取出水印信息。

4.2 算法思想

4.2.1 自动定位与矫正

文档的几何变形可以根据文档自身是否发生形变分为两类：一类是扭曲文档，其变形由文档自身的形变和拍摄角度共同引起，包括折痕、弯曲和褶皱变形等；另一类是平面文档，尽管文档自身平整无形变，但由于拍摄角度的影响，导致相机成像平面与文档平面不平行，从而产生透视倾斜变形，屏摄照片中的文档图像经历的便是此类形变。平面文档的矫正算法通常采用四点法或者消失点法，随后通过投影变换实现文档图像的矫正。四点法即为第三章使用的基于轮廓提取的矫正算法，通过手动标定或自动检测变形文档的四个角点，并定义矫正后的目标角点，求解坐标组的单应性矩阵以完成矫正。消失点法则利用文档中原本相互平行的文本基线在变形后会相交于消失点的特性，通过计算投影变换矩阵来实现矫正。

在屏摄过程中，若拍摄距离较近，则可能仅捕捉到文档的部分内容，导致摄屏照片中的文档图像缺乏完整的边缘信息。相反，若拍摄距离较远，虽然能够完整捕捉文档图像，但同时也会包含无关的背景环境。因此，在自然拍摄场景下解决文本图像变形问题时，应将任务划分为定位和矫正两个独立步骤。现有研究中，基于 PageNet^[75]和 U-net^[76]网络架构的语义分割算法^[77]被用于从照片中分割出文档区域，随后通过后处理和投影变换矫正文档区域的倾斜变形。然

而，第三章提出的抗屏摄水印算法能够在 1 米距离拍摄条件下准确提取水印信息。在这个拍摄距离下，照片中的干扰元素很多，若仅依据文档特征进行语义分割，可能会错误地将桌面上的其他纸质文档或显示器中显示的其余文本信息一并分割出来。为避免此问题，本章选择直接利用频域水印的能量特征来实现水印文档区域的自动定位。通过在频谱图的中频区域嵌入一个具有高幅值的定位圆环，该圆环由全 1 比特的消息序列生成，嵌入强度高于水印圆环。由于 DFT 频谱的平移不变性、线性和叠加性，含水印区域相较于非水印区域，在中频区域展现出显著的高能量特征。如图 4.1 所示，图 (a) 展示了经处理后含水印底纹分块的频谱图，而图 (b) 展示了相同处理条件下不含水印底纹分块的频谱图，两个分块图像均取自同一张屏摄照片。

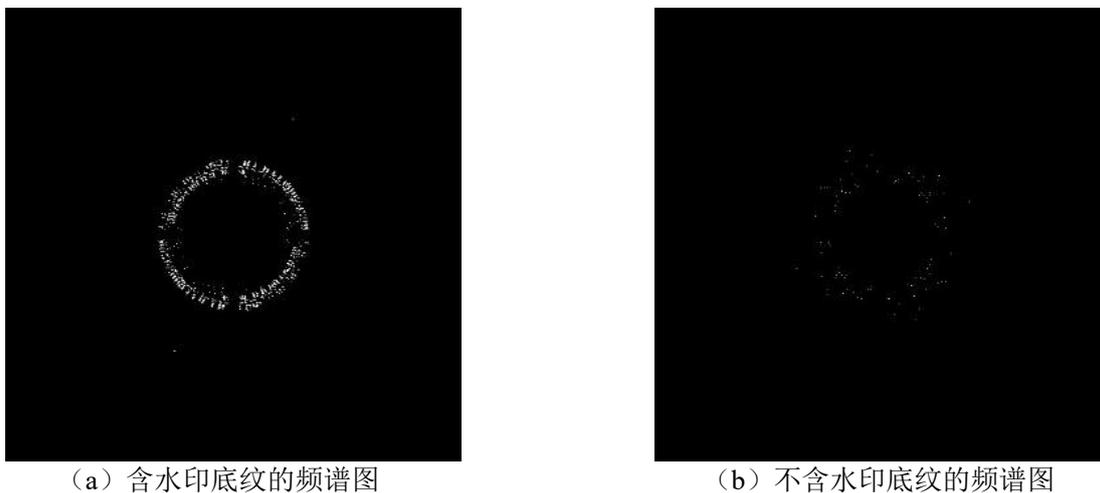


图 4.1 屏摄照片中不同分块处理后的频谱图

在文档图像矫正领域，基于文本行的方法是常见且有效的。文本行是文档图像中普遍存在且具有规律性结构的特征。在文献^[77-82]中，研究者通过文本行特征形成的水平线和换行符或字符垂直笔画特征形成的垂直线来估计两个消失点，以完成透视矫正。其中文献[82]中提出的算法在处理复杂布局或文本行稀少的情况下表现出色，适用于相机拍摄的文档图像或扫描图像，可以有效矫正由相机视角引起的透视变形。因此，本研究决定采用文献[82]中提出的算法对定位后的水印文档进行矫正。该算法通过结合文本行和线段的特性，并引入迭代优化方案以处理异常值，实现了对文档中文本和非文本区域的鲁棒矫正。实验结果表明，该算法在仅有部分文档内容的情况下矫正变形文档。

4.2.2 水印模板的抗裁剪性

在第三章中分析了平铺拼接后的圆环水印模板具有平移不变形、线性和叠加特性。这些特性对于仅获取部分水印文档时仍能完整提取水印信息至关重要，且无需预先定位图像中的水印区域。在第三章的研究中，是将变形的文档图像手动矫正回原尺寸后再提取水印，此时仅需将文档图像尺寸拉伸为正方形，便能确保其频谱中的水印环呈现为正圆形。然而，屏摄照片中的水印区域会随着拍摄距离的距离而缩小，其在照片中的尺寸是不定的，并且本章提出的自动定位算法并不能精准定位水印文档，文档矫正矫正算法也仅能将变形的文档图像矫正回原平面，因此矫正后的文档图像的尺寸也是不定的，导致其频谱中的水印环会形变为样式不一的椭圆环。

随后，再次利用定位环的高能量特征，准确检测出定位椭圆环并将其拉伸回正圆环，然后便可以顺利提取内外水印环的水印信息。因此，本方案也同样可以有效抵抗截屏攻击。当以大于模板尺寸的大小任意截取水印文档内容时，本方案都能实现对水印信息的盲提取。如图 4.2 所示，左侧文档图像的截取尺寸为 898×294 ，而右侧文档图像的截取尺寸为 396×840 。尽管两者频谱图中的水印环均呈现为椭圆环形态，但通过定位环的检测，便可以将其重新拉伸为正圆环，从而实现水印的准确提取。

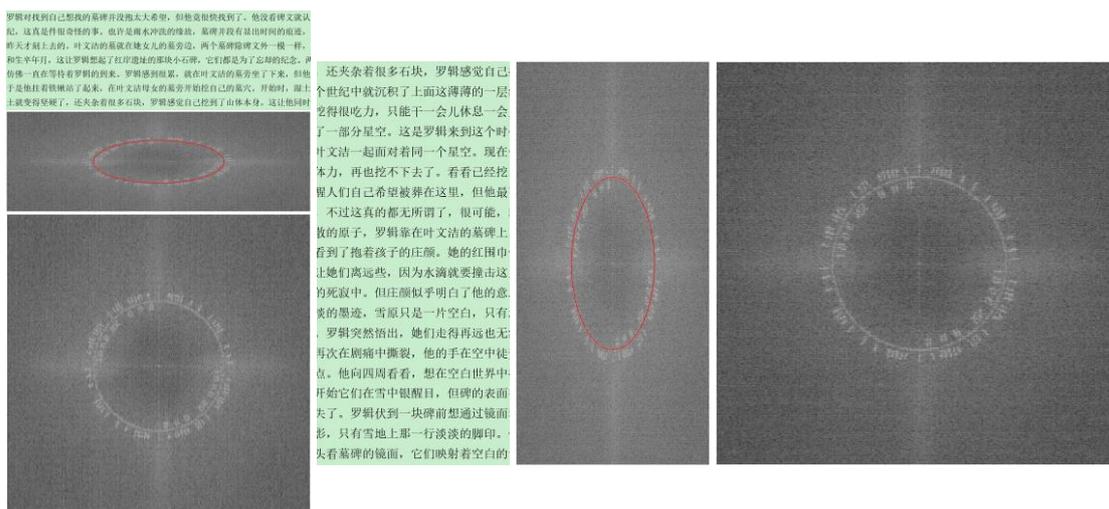


图 4.2 任意尺寸截取后的水印文档图像和频谱图

4.3 抗屏摄文档水印自动提取方案

4.3.1 水印嵌入流程

与上一章的方案相同，水印的嵌入流程分为两个步骤：水印模板的生成和水印模板的嵌入。首先对二进制消息序列 M 做预处理，使用 BCH 纠错码对原始消息序列进行编码。为了增加水印容量，本方案选择将编码序列拆分为两个序列，然后分别生成两个不同半径的圆环水印模板。并且为了能够实现水印的盲提取以及提取时对水印环的准确定位，本方案在两个圆环水印中间加入了一个定位圆环，最终组成一个完整的水印模板。然后依然选择原始底纹的亮度分量 Y 来嵌入水印模板。水印模板的嵌入过程仍然采取加性嵌入，但水印环和定位环的嵌入强度并不保持一致。通过对原始底纹的 Y 分量进行 DFT 变换后，取出频谱图 AM 与水印模板做加性嵌入，其中定位环的嵌入强度为 α_c ，水印环的嵌入强度为 α 。最终依次经历 IDFT 和色域空间变换得到最终的水印底纹。

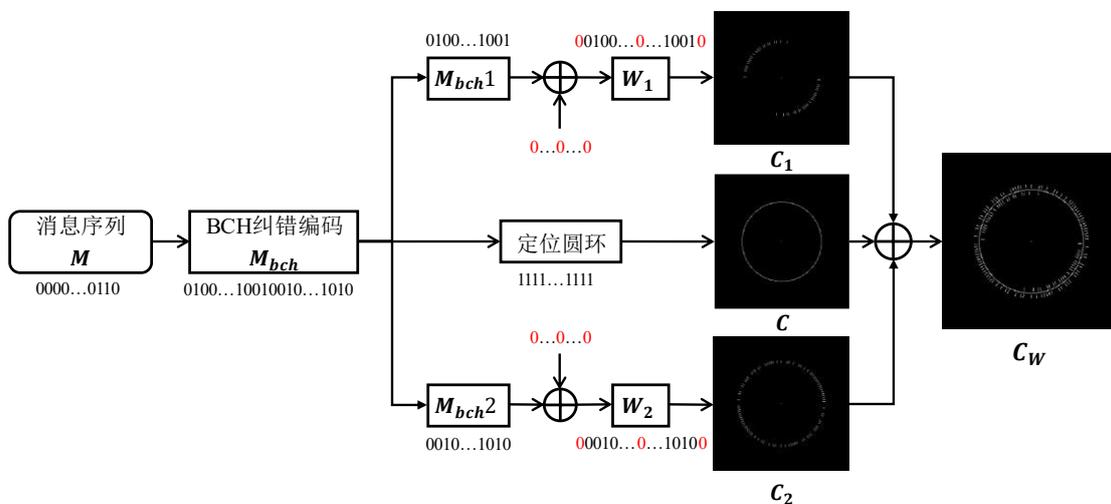


图 4.3 水印模板生成流程图

1) **水印模板生成**：生成水印模板的流程如图 4.3 所示。首先，对二进制消息序列 M 使用 BCH 纠错编码，以生成具有纠错能力的二进制序列 M_{bch} 。为了增加水印容量，本方案选择在中频区域嵌入两个水印环，分为内环 C_1 和外环 C_2 ，所以需要将 M_{bch} 拆分为 M_{bch1} 和 M_{bch2} 。同样为了避免因频域周期性引发的图像边界不连续性，而产生的中心化频谱区域的中心十字型亮条干扰，对拆

分后的序列分别进行冗余位填充，在序列的首尾和中间加入 0 比特信息位。最终得到内环 C_1 的水印序列为 W_1 ，外环 C_2 的水印序列为 W_2 。同时，为了解决水印的盲提取和水印文档的自动定位问题，本方案在内外两个水印环之间嵌入了一个定位环 C ，由全 1 的二进制序列组成。水印模板的尺寸与底纹图像的尺寸保持一致，均为 $a \times a$ 。为了抵抗压缩攻击和文字编辑带来的低频扰动，本方案选择底纹的中频区域来嵌入圆环水印，其中定位圆环的半径为 R_c ，内环水印的半径为 R_1 ，外环水印的半径为 R_2 。水印模板 C_1 、 C_2 和 C 的生成过程与第三章相同，分别使用公式(3.8)和公式(3.9)计算圆环坐标，然后使用公式(3.10)将水印序列代入相应坐标中。为了增加鲁棒性，本方案仍然对水印圆环采取扩频处理，每比特的水印信息由同角度不同半径形成的线段来表示，扩频线段的长度设为 l 。内环水印半径 R_1 和外环水印半径 R_2 与定位环半径 R_c 的关系为：

$$\begin{cases} R_1 \in [R_c - 3, R_c - (2+l)] \\ R_2 \in [R_c + 3, R_c + (2+l)] \end{cases} \quad (4.1)$$

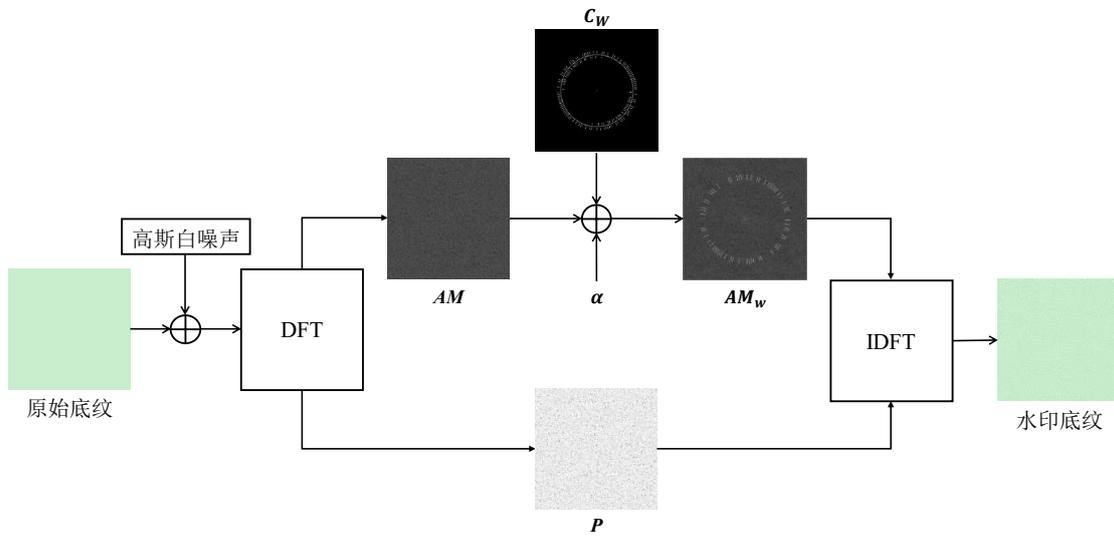


图 4.4 水印模板嵌入流程图

2) **水印模板嵌入**：嵌入水印模板的流程如图 4.4 所示。水印模板的嵌入步骤与第三章相同，选取原始底纹的 Y 分量作为嵌入对象，并添加方差为 0.0005 的轻微高斯白噪声。对加噪后的 Y 分量进行 DFT 变换得到幅度谱 AM 和相位谱 P ，对中心移位后的 AM 进行水印模板的加性嵌入得到水印幅度谱 AM_w ：

$$AM_w = AM + \alpha \cdot (C_1 + C_2) + \alpha_c \cdot C, \quad (\alpha > 1.5\alpha_c) \quad (4.2)$$

其中，定位环的嵌入强度 α_c 需大于 1.5 倍的水印环嵌入强度 α 。接着将水印幅度谱 AM_w 和相位谱 P 重新组合并经过 IDFT 变换得到含水印的 Y 分量 Y_w 。最后使用色彩空间转换得到最终的含水印底纹。

4.3.2 水印提取流程

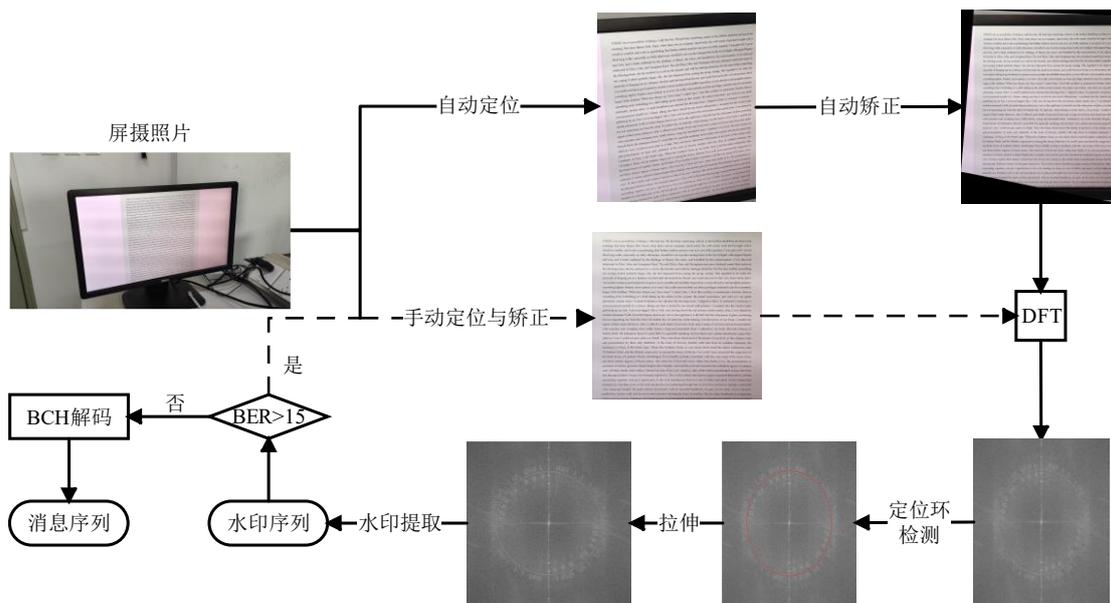


图 4.5 水印提取流程图 水印提取流程图

图 4.5 详细展示了水印的提取过程。首先，获取到拍摄的水印文档照片后，关键步骤是自动定位照片中的水印文档区域。随后，依据定位区域内的文本行特征，自动执行透视矫正，完成水印提取前的预处理工作。将矫正后的文档图像进行 DFT 变换，然后根据高能量频谱特征检测并定位到因未矫正回原尺寸而变形的椭圆定位环。将其拉伸回圆环后，分别提取嵌入在内外环中的水印信息，对提取水印信息进行拼接和 BCH 解码处理得到提取的消息序列。本方案采用 BCH(127,36)编码，具备 15 比特的纠错能力。自动提取过程中，若水印的错误比特数超出 15 位，说明自动定位和矫正出现了偏差。在这种情况下，建议使用手动透视矫正并重新提取水印信息。

1) 水印区域的自动定位：为了在不受到照片中其他文本信息干扰的情况下，准确识别并定位照片中的水印文档，本方案采用了一种基于中频区域高能量特征的自动定位算法，其算法流程如算法 4-1 所示。自动定位的原理是基于频域的高能量特征在屏摄后依然能够被检测到。如图 4.6 所示，图 4.6 (a) 展示了

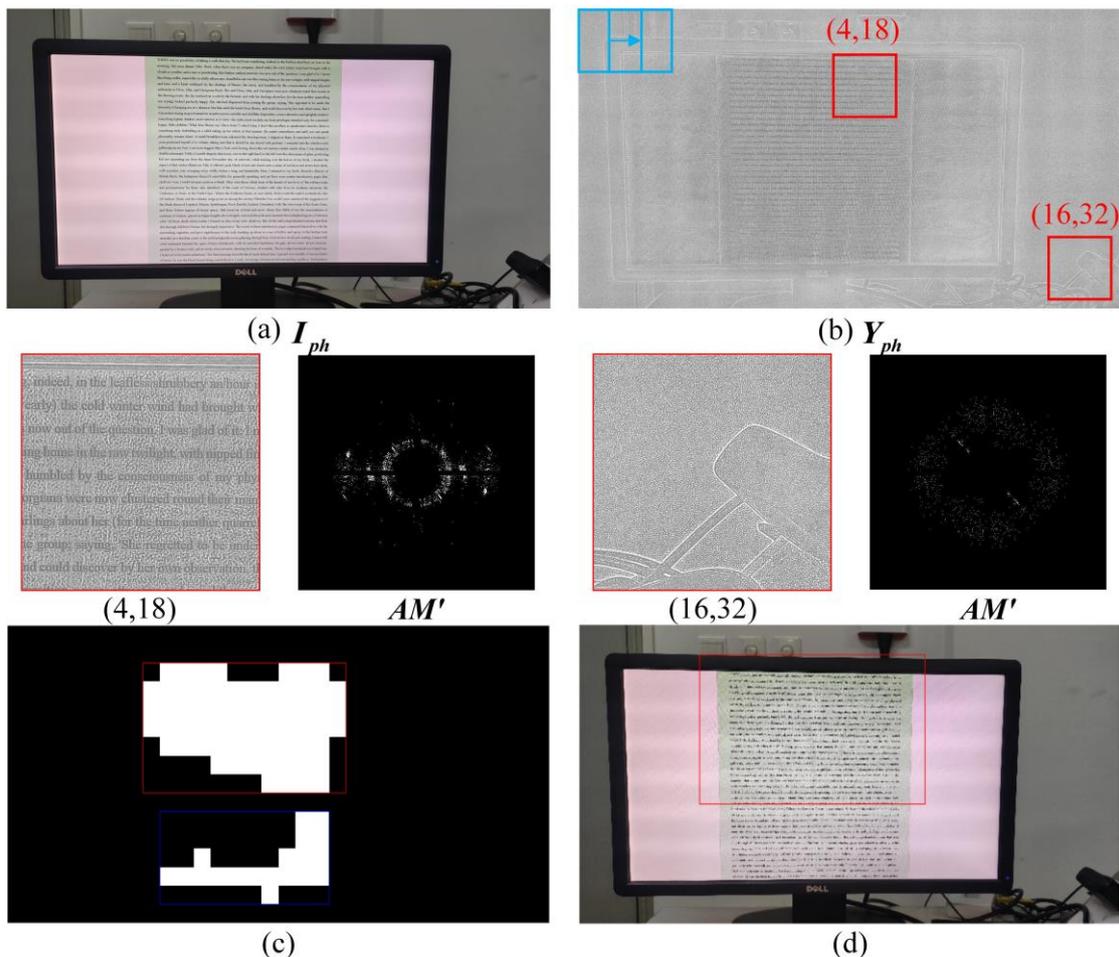


图 4.6 水印区域的自动定位示意图

从正面 55cm 处拍摄得到的照片 I_{ph} 。首先，提取 I_{ph} 的亮度分量 Y 并进行预处理，包括自适应直方图均衡化和钝化掩蔽处理，旨在增强水印底纹的纹理特征，从而加强其频域特征。然后，如图 4.6 (b) 所示，对预处理得到的 Y_{ph} 图像执行基于滑动窗口的分割操作，以固定步长移动窗口并截取图像的分块区域。在图 4.6 中，图像 (4,18) 和 (16,32) 便是通过滑动窗口分割得到的分块图像，其中 (4,18) 分块为含有水印的文档图像块，而 (16,32) 分块为不含有水印的背景图像块。对每个截取的分块图像执行 DFT 变换，得到相应的频谱图 AM 。随后对频谱图 AM 进行二值化处理，阈值设定为 β 倍的频谱直流系数 $F(0,0)$ 。为了便于比较，对二值化后的频谱图执行中心频移，便可以得到如图 4.6 所示的 AM' 。通过比较含水印和不含水印的分块图像的二值化频谱图 AM' ，可以观察到它们的中频能量值存在显著差异。本方案采用统计二值化频谱图 AM' 中值为 1 的像素数量的方法来量化能量数值。最后，如图 4.6 (c) 所示，标记出二值化

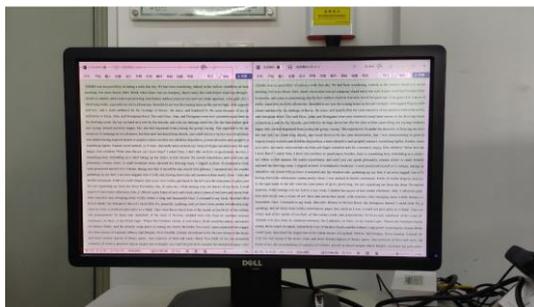
频谱图中能量强度高于所有图像块平均强度的图像块，并通过划分最大连通域来完成水印文档的自动定位。最终的自动定位结果如图 4.6 (d) 所示。

算法 4-1: 水印区域自动定位算法流程

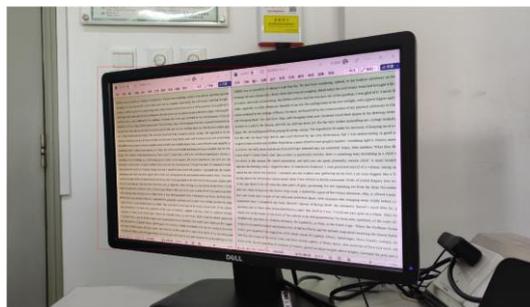
输入: 屏摄照片 I_{ph} , 步长 $step$, 分块尺寸 $block_{size}$

输出: 定位区域图像 I_{loc}

- 1: 将 I_{ph} 从 RGB 转换为 YCbCr 并提取 Y
 - 2: 对 Y 进行预处理 (自适应直方图均衡和钝化掩蔽), 得到 Y_{ph}
 - 3: 使用 $step = 128$ 和 $block_{size} = 512$ 的滑动窗口从 Y_{ph} 中提取分块图像
 - 4: **for** 每个分块 **do**
 - 5: DFT 变换并提取幅度谱 AM
 - 6: $threshold = \beta \times F(0,0)$
 - 7: 用 $threshold$ 对 AM 二值化, 得到 AM'
 - 8: 分块强度 $strength$ 表示为 AM' 中非零值的数量
 - 9: **end for**
 - 10: 标记强度大于平均值的分块
 - 11: 在标记块中获取最大连通域得到 I_{loc} 。
 - 12: **返回:** 定位区域图像 I_{loc}
-



(a) 60cm 正面拍摄定位结果



(b) 60cm 右偏 30° 斜拍定位结果

图 4.7 复杂场景下的水印自动定位示意图

此外, 为了验证所提方法在复杂场景下的鲁棒性, 进行了一组对比实验。在 WORD 中打开两个相同内容的文档进行并排对比, 其中左侧文档采用浅绿色的水印底纹, 而右侧文档则使用浅绿色的纯色底纹。如图 4.7 所示, 即使在两种不同的拍摄角度下, 所提出的自动定位算法均能准确地识别水印文档区域, 从而有效地排除了照片中不含水印内容的干扰。

2) 文档的自动矫正: 应用文档矫正算法[82]对自动定位得到的水印文档图像执行自动透视校正。该算法依托于文档中存在的对齐文本行或线段, 包括照片、图形和表格等元素, 以实现文档的透视矫正。该算法的鲁棒性可以保证

即使自动定位算法仅定位了文档的部分区域，只要该部分包含可识别的文本行特征，即可成功完成透视矫正。

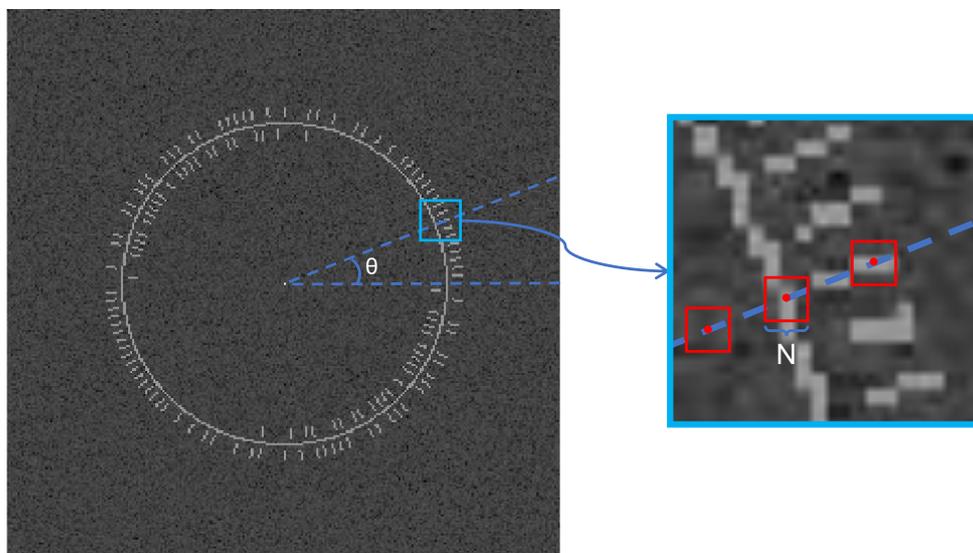


图 4.8 水印提取原理示意图

3) 水印的盲提取：提取矫正后的文档图像的亮度分量 Y_{oval} 作为水印的预处理对象，然后使用自适应直方图均衡强化水印底纹的纹理特征，再使用钝化掩蔽消除低频噪声。将纹理增强和去噪滤波后的 Y_{oval} 执行 DFT 变换得到其幅度谱 AM_{oval} 。因为文档矫正算法仅会将文档图像矫正回原平面，文档图像的尺寸会发生改变。由于 DFT 频谱的线性和叠加特性，频谱中的圆环水印也会相应变形为椭圆环水印，而在水印提取前，需要将其重新拉伸回正圆环。结合定位环的高能量特征，通过在不同长半轴 a 和短半轴 b 的搜索遍历，寻找能量值累加最高的椭圆环，最终得到定位环的长半轴 a_1 和短半轴 b_1 。再根据 a_1 和 b_1 将 AM_{oval} 中的椭圆环水印拉伸位圆环水印，并确定定位环的半径 r_c ，因此内环水印半径 r_1 和外环水印半径 r_2 为：

$$\begin{cases} r_1 = r_c - 3 \cdot t \cdot r_c - (2+l) \cdot t \\ r_2 = r_c + 3 \cdot t \cdot r_c + (2+l) \cdot t \end{cases}, t = \frac{r_c}{R_c} \quad (4.3)$$

其中 t 为提取的定位环半径 r_c 与定位环嵌入半径 R_c 的比值。然后分别在相同角度 θ 下提取半径 r_c 、 r_1 和 r_2 在 N 范围内的最大值 $max_c(\theta)$ 、 $max_1(\theta)$ 和 $max_2(\theta)$ ，则内环水印 w_1 和外环水印 w_2 为：

$$w(\theta) = \begin{cases} 1 & \text{若 } \max_r(\theta) \geq \max_{r_c}(\theta) \\ 0 & \text{若 } \max_r(\theta) \leq \max_{r_c}(\theta) \end{cases} \quad (4.4)$$

图 4.8 展示了水印提取的原理，其中最大值是在以相应半径和角度的坐标为中心的 $N \times N$ 区域内选择的。 N 的大小依然设定为半径为 r 时，相邻的两比特水印之间弧长的 $2/3$ ，如公式(3.15)所示。 k 仍为圆环水印的水印比特数，在定位到水印圆环后，遍历提取 $l \cdot t$ 个圆环水印序列。通过分析不同水印序列，选择每个信息位中出现频率最高的水印比特，并将这些选定的比特组合以构成最终提取水印序列 w_1 和 w_2 。

4) 纠错解码：在得到水印序列 w_1 和 w_2 后，首先需要分别去除添加的冗余位，得到待解码的消息序列 $m_{bch}1$ 和 $m_{bch}2$ 。将 $m_{bch}1$ 和 $m_{bch}2$ 合并为完整的待解码消息序列 m_{bch} ，最后再通过 BCH 解码便可以得到最初嵌入的消息序列 m 。

4.4 实验结果与分析

4.4.1 实验参数

本小节主要介绍本方法的实验参数，包括实验场景的设备参数设置，和水印底纹模板参数，并结合第三章的定量统计实验结果来选定具体参数。

1) 设备参数：为了验证此方法对硬件设备的普适性，本文使用分辨率为 1080p 的显示器 ‘DELL-E2213H’ 显示水印文档，后置摄像设为 12MP 的 ‘Redmi K30s Ultra’ 的手机来拍摄屏幕。直接使用 WORD 软件载入水印底纹，纸张大小设为 A4 纸大小，字体大小设为默认值 10.5 磅，行间距也设为默认值 1.15 倍。分别使用中文文档和英文文档进行实验，其中中文文档节选自《三体》的某一段落，英文文档节选自《简爱》的某一段落。在后面的实验中统一使用 DocCN 来表示中文文档，DocEN 来表示英文文档。

2) 底纹模板参数：为了保证局部文档在拍摄后仍然可以提取水印信息，本方法将底纹模板的尺寸设为 256×256 ，即 $a = 256$ ，此尺寸可以保证水印信息的顺利提取，并且仅包含少量的文字信息，不会造成过大的隐私泄露风险。纠错码选择 BCH(127,36)编码，纠错能力为 15 位，实际嵌入消息长度为 36 比特，纠

错编码序列长度为 127 比特，拆分为两个序列并添加冗余位后的两个水印序列长度 k 均设为 66 比特。

本方案依然选择在频谱的中频区域内嵌入水印圆环，内外环水印的扩频长度 l 设为 5，并且为了可以在 N 范围内顺利提取最大幅值，均与定位环 C 有 3 个单位的间隔距离。根据第 3.4.1 节中对圆环半径和嵌入强度的定量统计实验数据，在 256×256 的频谱图中，当水印序列长度为 66 时，水印半径最小可设为 48，可以满足两比特水印之间留有足够的空间用以提取最大幅值，并且能够准确提取所有的水印比特信息。并且由表 3.1 的数据可知，水印圆环的半径越大，水印越偏向向高频区域，其抗屏摄性能和抗裁剪性能越低。因此根据公式 (4.3) 可以推导出定位环 C 的嵌入半径 R_c 应设为 55，内环水印半径 $R_1 \in [48, 52]$ ，外环水印半径 $R_2 \in [58, 62]$ 。

而嵌入强度可视应用场景的不同而灵活设置。经实验证明，嵌入强度在一定范围内与水印比特提取准确率成正比，但与水印文档的视觉质量成反比。本章应对三种不同的应用场景设置了三种嵌入强度方案。表 4.1 展示了三种嵌入器强度方案，表 4.2 展示了三种方案与其他两种方法对抗不同图像处理攻击的鲁棒性。其中三种方案的应用场景分别为：

- a) **方案 I**：可以抵抗常规图像处理攻击。本方案可以使用较低的嵌入强度，获得较高的视觉质量，并且足以抵抗多种常规的图像处理攻击，例如：JPEG 压缩、高斯模糊、中值滤波、尺度缩放和裁剪攻击。
- b) **方案 II**：可以抵抗屏摄攻击但无法实现自动定位。此方案与第三章相同，需要手动选取照片中水印文档的四个角点完成透视校正。不同的是本方案不再需要原文档的先验尺寸，因为水印提取时会检测定位环并自动调整，所以矫正时图像的尺寸可以自行设置。
- c) **方案 III**：可以抵抗屏摄攻击并且可以实现水印的自动盲提取功能。嵌入强度的提升有助于抵抗自动定位的偏差，但文档的视觉质量也会因此下降。

表 4.1 三种不同的嵌入强度方案

嵌入方案	I	II	III
嵌入强度	$\alpha = 7000$ $\alpha_c = 4200$	$\alpha = 12000$ $\alpha_c = 7000$	$\alpha = 18000$ $\alpha_c = 10000$

表 4.2 不同方案抵抗不同图像处理攻击的鲁棒性

攻击类型	JPEG 压缩	高斯噪声	中值滤波	尺度缩放	裁剪攻击	屏摄攻击
Ge 等 ^[34]	✓	✓	✗	✗	✗	✓
Fang 等 ^[57]	✗	✓	✗	✗	✗	✓
方案 I	✓	✓	✓	✓	✓	✗
方案 II	✓	✓	✓	✓	✓	✓
方案 III	✓	✓	✓	✓	✓	✓

4.4.2 文档图像的视觉质量



(a) 原始文档 (b) 方案 I (c) 方案 II (d) 方案 III

图 4.9 不同水印算法的中文文档图像视觉质量对比

图 4.9 展示了三种不同颜色的底纹下，原始文档图像和本章提出的三种嵌入强度方案的中文水印文档图像。可以发现，随着嵌入强度的增强，水印底纹的纹理特征越清晰，水印的隐蔽性则随之变差，但仍然可以提供一个好的阅读观感。为了客观评估水印文档的视觉质量，本章选择使用峰值信噪比（PSNR）和结构相似度指数（SSIM）作为度量标准。表 4.3 详细展示了不同方法下中英

文文档的具体 PSNR 和 SSIM 值。可以发现，本算法在使用嵌入强度最大的方案 III 时，文档图像的视觉质量仍然是三种算法中最优秀的。方案 I 在不同颜色的底纹中的 PSNR 平均值可以达到 42.0dB。随着嵌入强度的增大，方案 II 和方案 III 的 PSNR 平均值分别为 37.5dB 和 34.2dB。而 Fang 等人的方法的 PSNR 平均值为 31.2dB，并且即便是嵌入强度最大的方案 III 也比 Fang 等人的方法平均高出 3dB。除了使用方案 III 的黑色底纹，本算法在其余情况下的 SSIM 值均高于 0.9。并且与其余两种方法相比，本算法与原文档图像的相似度也同样最高。

表 4.3 不同水印方案的中英文文档的视觉质量评估

方法 \ 底纹颜色		绿色底纹		黑色底纹		白色底纹	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DocCN	Ge 等 ^[34]	/	/	/	/	25.62	0.817
	Fang 等 ^[57]	30.26	0.927	30.29	0.802	33.11	0.885
	方案 I	41.35	0.994	41.61	0.971	43.11	0.984
	方案 II	36.94	0.983	37.05	0.931	38.59	0.959
	方案 III	33.56	0.964	33.63	0.875	35.19	0.922
DocEN	Ge 等 ^[34]	/	/	/	/	22.88	0.724
	Fang 等 ^[57]	30.32	0.921	30.30	0.753	33.18	0.859
	方案 I	41.42	0.993	41.59	0.962	43.18	0.980
	方案 II	36.98	0.981	37.03	0.912	38.63	0.949
	方案 III	33.62	0.961	33.64	0.845	35.28	0.905

4.4.3 对常见图像处理攻击的鲁棒性分析

为了证明本算法对多种常见信号处理攻击具有良好的鲁棒性，本章进行了相应实验。实验结果如表 4.4 所示，实验使用的水印文档图像大小为 990×1401 ，是在 WORD 中使用 A4 纸尺寸下并且页面缩放为 100% 时一页文档的截图尺寸。因为纯白色底纹的 Y 分量为 255，相同嵌入强度的情况下其空域的修改程度会减弱，经实验证明，方案 I 下的白色水印底纹无法有效抵抗 JPEG 压缩攻击和尺度缩放攻击，因此本章测试了方案 II 下的白色水印底纹对常见信号处理攻击的鲁棒性。由表 4.4 的实验数据可知，本算法对 JPEG 压缩、高斯模糊、中值滤波、尺度缩放和裁剪攻击都有相当的鲁棒性。其中三种颜色的水印底纹均可以抵抗质量因子为 50 的 JPEG 压缩，而绿色水印底纹可以抵抗质量因子为 40 的 JPEG 压缩。三种颜色的水印底纹也都可以抵抗标准差为 1.2 的高斯模糊攻击和窗口大小为 3×3 的中值滤波。因为水印底纹是由 256×256 大小的水印模板

拼接而成，并且由于 DFT 频谱的特性，其具有良好的抗裁剪性能。即便在文档图像遭受高达 95% 的裁剪时，依然能够准确地提取出水印信息。然而，同样由于 DFT 频谱的缩放特性，当文档图像经历缩小处理时，频谱中的水印圆环尺寸保持不变。随着缩放尺度的增加，频谱中的水印环可能会超出图像的尺寸范围，导致无法被有效检测。因此，本算法仅能抵御相当于原图像尺寸 55% 的缩放攻击。

表 4.4 常见信号处理攻击的水印比特提取准确率 (%)

攻击类型	绿色底纹 (方案 I)		黑色底纹 (方案 I)		白色底纹 (方案 II)	
	DocEN	DocCN	DocEN	DocCN	DocEN	DocCN
高斯模糊 ($\sigma = 1$)	94.49	91.34	98.43	97.64	96.06	95.28
高斯模糊 ($\sigma = 1.1$)	91.34	91.34	96.85	95.28	94.49	93.70
高斯模糊 ($\sigma = 1.2$)	91.34	90.55	95.28	92.13	94.49	92.13
JPEG 压缩 (Q = 60)	95.28	94.49	95.28	90.55	92.13	94.49
JPEG 压缩 (Q = 50)	93.70	95.28	90.55	91.34	89.76	88.98
JPEG 压缩 (Q = 40)	88.19	90.55	87.40	88.19	88.98	82.68
缩放至 75%	96.85	98.43	93.70	91.34	88.98	96.85
缩放至 65%	98.43	100.00	98.43	96.85	88.98	99.21
缩放至 55%	96.85	98.43	91.34	94.49	88.19	97.64
裁剪 75%	99.21	98.43	97.64	97.64	96.06	95.28
裁剪 85%	97.64	99.21	97.64	96.85	99.21	94.49
裁剪 95%	98.43	99.21	98.43	97.64	98.43	93.70
中值滤波 [3×3]	97.64	98.43	99.21	98.43	96.06	97.64

4.4.4 水印的嵌入容量及鲁棒性分析

本节旨在评估本章提出水印算法的嵌入容量以及对屏幕拍摄攻击的鲁棒性。本章使用的是 BCH(127,36) 编码，纠错能力为 15 位，实际嵌入消息长度为 36 比特，也即本章方案的水印嵌入容量为 36 比特。相较第三章提出方案的水印容量提升了两倍，但也为了抵消自动定位和矫正带来的误差而提升了纠错能力，导致未能进一步扩充水印容量。实验设备和拍摄场景的选取与第三章保持一致，并且为了确保能够准确评估算法的屏摄鲁棒性，必须首先排除自动定位和自动矫正可能引入的偏差。因此，实验中分别对手动定位与矫正后的屏摄鲁棒性，以及采用自动水印提取算法的屏摄鲁棒性进行了验证。在手动定位与矫正条件下，通过比较方案 II 与其他两种现有屏摄文档水印算法在水印比特提取准确率的表现，来验证本章算法的屏摄鲁棒性。在采用自动水印提取算法时，使用方案 III 来检验算法的有效性，以及对屏摄攻击的鲁棒性。

1) 手动定位与矫正下算法的鲁棒性: 图 4.10 展示了方案 II 与其他两种屏摄文档水印算法对距离、水平角度和垂直角度的鲁棒性。可以发现, 本章算法很好地保留了第三章算法的屏摄鲁棒性, 能够在接近 1 米的拍摄距离下准确提取水印信息。此外, 在 40cm 的距离下, 无论是面对水平还是垂直的透视角度变化, 该算法均能有效地抵御高达 65°的大角度透视失真。相比之下, Ge 等人^[34]提出的方法在稳定性上表现较差, 对摩尔纹失真较为敏感, 这导致其在比特提取准确率上会有显著下降。在同等条件下, 本章提出的算法展现出了对摩尔纹失真的优越鲁棒性。并且算法中的冗余性设计也有效提升了水印比特的提取准确率。尽管与第三章相同, 都需要手动定位和矫正透视变形的文档图像, 但本章提出的算法在采用四点法进行矫正时, 不再需要原档图像的先验尺寸信息。在矫正过程完成后, 会通过检测和拉伸定位环来实现尺寸的自行调整。

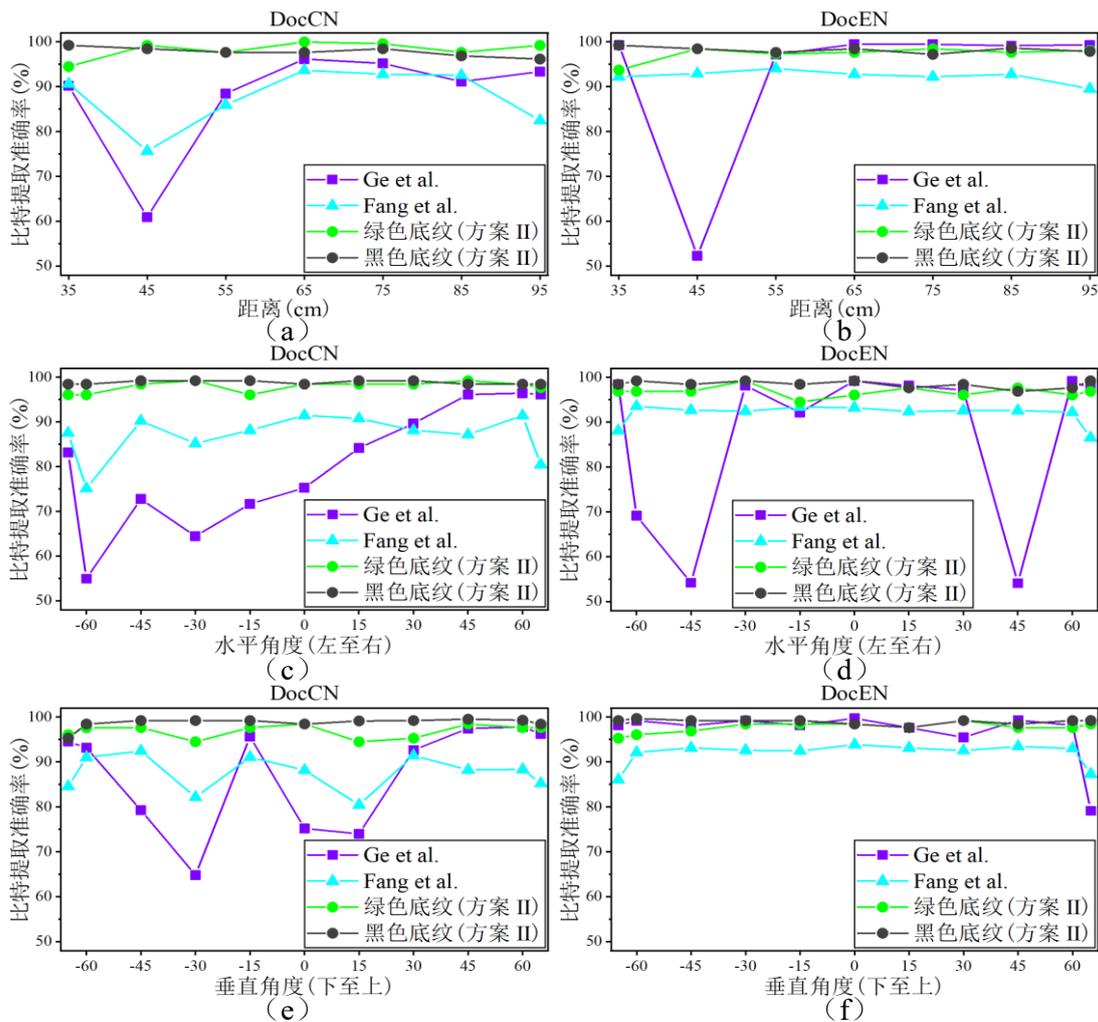


图 4.10 手动定位与矫正下的屏摄鲁棒性

2) 自动定位与矫正下算法的鲁棒性: 鉴于所选文档校正算法^[82]的固有局限性, 其在矫正较大的垂直透视角度偏转时存在局限。因此, 本实验仅针对方案 III 对距离和水平角度的鲁棒性进行了评估。在正面拍摄条件下, 由于产生的偏转角较小, 算法的性能主要取决于自动定位的准确性。如图 4.11 所示, 本章提出的自动定位算法能够有效地在近距离 (45 厘米) 和远距离 (95 厘米) 条件下准确定位照片中的水印区域。表 4.5 展示了在采用自动水印提取算法后, 中英文档在正面不同距离拍摄条件下的水印比特提取准确率。由表中数据可知, 在正面拍摄的情况下, 无论文档底纹颜色为绿色还是黑色, 水印的平均比特提取准确率均达到了 97.4%, 这得益于准确的自动定位算法和高冗余度的水印盲提取算法。与手动定位与矫正相比, 自动水印提取算法在水印比特提取准确率上略有下降。这主要是由于自动定位算法在较长距离拍摄时, 缺乏一定的精度, 并且定位结果可能包含了照片中的无关背景物品, 从而影响了水印提取的准确率。尽管如此, 在 95 厘米的拍摄距离下, 本章提出的自动水印提取算法仍能保持超过 90% 的比特提取准确率。

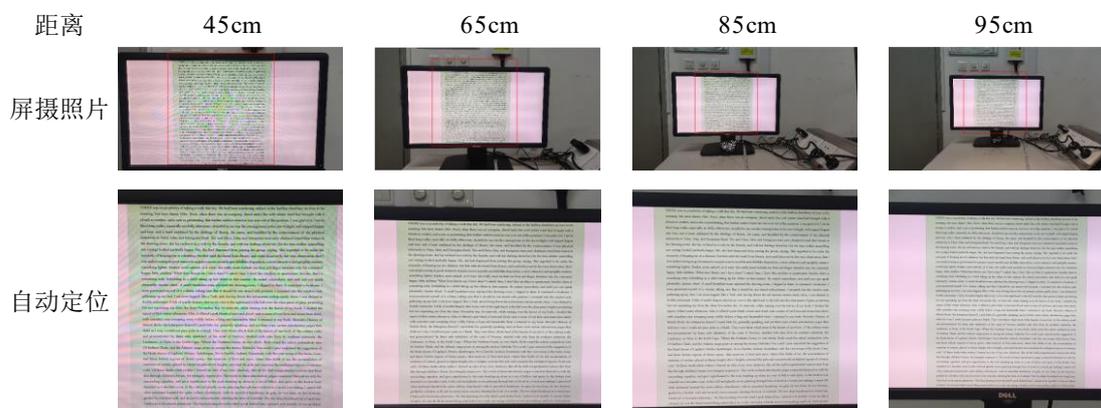


图 4.11 正面不同距离拍摄后的自动定位展示

表 4.5 正面不同距离拍摄的水印比特提取准确率 (%)

距离 (cm)		35	45	55	65	75	85	95
DocCN	绿色底纹 III	94.5	99.9	99.9	99.8	99.2	99.2	99.2
	黑色底纹 III	97.6	99.2	99.2	98.4	96.1	95.3	92.9
DocEN	绿色底纹 III	97.6	99.2	98.4	99.2	99.2	99.2	99.2
	黑色底纹 III	94.5	97.6	96.9	96.9	95.3	93.7	90.6

在水平角度偏转的拍摄条件下, 自动定位与自动矫正算法的性能对最终水印提取算法的效果都有影响。图 4.12 展示了在 40 厘米拍摄距离下, 不同水平角度偏转拍摄后, 本算法的自动定位与自动矫正结果。观察结果可以发现, 即便

在高达 65° 的水平角度偏转下，本章提出的算法依然能够成功完成水印文档的自动定位与自动矫正。表 4.6 展示了在正面 45 厘米拍摄距离下，不同水平角度偏转拍摄条件下，采用自动水印提取算法后的中英文文档水印比特提取准确率。根据表中数据，当中文水印文档左偏 65° 和 15° 时，水印比特提取准确率降至 90% 以下。这一现象可归因于在极端偏转角度下，自动矫正算法可能引入的偏转误差，该误差会导致频谱中的水印环也发生相应偏转，进而影响水印提取的准确率。因此，未来的研究应考虑引入具有抗旋转特性的 DFT 水印技术，以增强算法对一定旋转偏转的鲁棒性。而在其他水平角度偏转情况下，本章提出的自动水印提取算法有着水印比特提取准确率均高于 90% 的良好性能，表明该算法对于水平透视角度变化具有良好的鲁棒性。

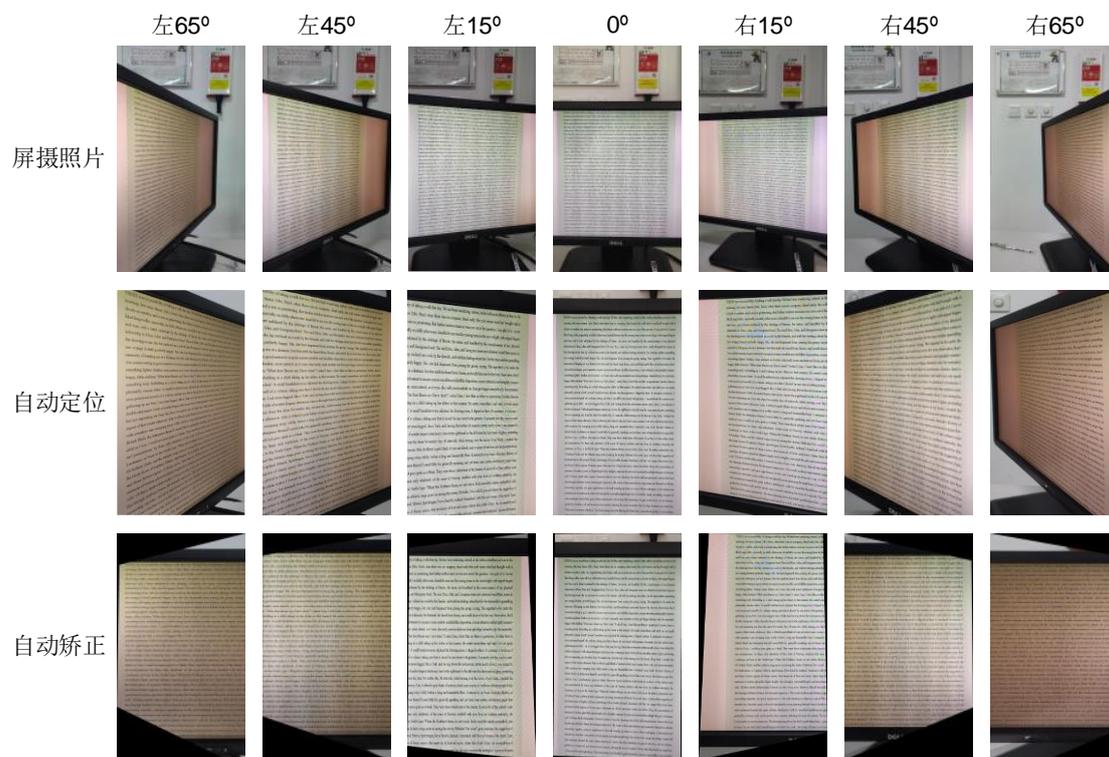


图 4.12 40cm 水平不同角度拍摄后的自动定位与矫正展示

表 4.6 40cm 处水平不同角度拍摄的水印比特提取准确率 (%)

水平角度 ($^\circ$)		左 65	左 60	左 45	左 30	左 15	0	右 15	右 30	右 45	右 60	右 65
Doc	绿色底纹 III	88.2	98.4	99.7	99.8	99.2	99.2	98.4	99.2	98.4	96.9	97.6
	CN 黑色底纹 III	88.2	98.4	92.9	97.6	88.2	94.5	97.6	93.7	99.2	98.4	92.9
Doc	绿色底纹 III	96.9	92.9	99.2	96.9	96.1	96.9	99.2	97.6	99.2	96.9	91.3
	EN 黑色底纹 III	98.4	97.6	96.9	98.4	98.4	98.4	99.2	95.3	96.9	99.2	99.2

4.4.5 算法通用性测试

为了评估本章提出的算法在不同设备上的屏摄鲁棒性，即验证算法的通用性，本研究开展了一系列的屏摄测试。测试采用了多款不同品牌的显示器和手机，这些设备在分辨率和摄像像素数上有所差异。实验条件与前文所述保持一致：利用 WORD 软件加载含水印底纹的文档，并使用程序默认设置进行展示；拍摄操作固定在屏幕正面 40 厘米处进行。表 4.7 展示了使用不同手机拍摄不同显示器屏幕后，水印比特提取的准确率。根据表中数据可知，本章提出的算法在不同设备上显示出了良好的普适性，这是因为当采用更高像素的相机和更高分辨率的显示设备，水印提取的准确率也会随之提升。

表 4.7 不同设备的水印比特提取准确率 (%)

手机型号 显示器型号	iPhone 14 (12 MP)	RedMi K40pro (16 MP)	Huawei Mate50 (12 MP)
MacbookPro (2560×1600p)	99.2	99.6	99.6
AOC 2730 (2560×1440p)	99.2	99.2	98.4
Huawei AD80HW (1920×1080p)	99.2	96.9	98.4
HP 400 G6 AiO (1920×1080p)	99.2	99.6	99.6

4.5 本章小结

本章提出了一种抗屏摄全盲自动文档水印算法，旨在解决现有抗屏摄水印研究中普遍存在的手动定位与校正难题。本章提出的自动提取系统能够直接处理屏摄照片，实现对照片中水印文档的自动定位、矫正和水印提取。通过设计含有定位环的新型 DFT 圆环水印模板，有效解决了水印区域的自动定位问题，并克服了矫正后尺寸变形的挑战，实现了全盲水印提取。此外，由于水印底纹由多个小尺寸水印模板拼接构成，因此即使仅拍摄到部分文档内容，也能完整且准确地提取水印信息。本章还根据不同的应用场景，设计了多种嵌入强度方案。在采用较低嵌入强度的方案时，拥有更佳视觉质量的同时，对常见的图像信号处理攻击也拥有一定的鲁棒性。在采用较强嵌入强度方案时，经大量实验证明，在维持良好视觉质量的同时，对常见的图像处理攻击和屏摄攻击均表现出了良好的鲁棒性，并且具备良好的设备通用性。

第五章 总结与展望

5.1 总结

随着数字化办公与智能手机的普及，屏幕拍摄已成为常见的信息传播方式，但同时也成为难以防范的隐私泄露方式。数字鲁棒水印技术作为版权保护和溯源认证的有效手段，也很难抵抗屏摄这种跨媒体信息传输。现有针对屏摄文档的鲁棒水印研究的鲁棒性较弱，无法有效抵抗屏摄过程中的组合失真，并且无法有效地平衡文档特性、水印隐蔽性和屏摄鲁棒性。此外，现有的屏摄文档水印方案的实用性也较差，在水印检出时需要手动干预，因此无法批量处理屏摄照片的版权认证，并且也无法对局部拍摄的文档实现水印提取。针对上述问题，本文首先深入分析了屏摄过程中可能引入的五种失真类型，主要包括相机失真、拍摄失真、屏幕显示失真、光照失真和摩尔纹失真。然后结合 DFT 频谱的特性、环形水印模板的鲁棒性、BCH 纠错编码和图像预处理算法等，设计了相应的鲁棒水印算法，并对算法生成的水印文档的视觉质量和抗屏摄鲁棒性进行了客观评估，也与现有的抗屏摄文档水印算法进行了客观比较。实验结果表明，本文所提出的基于 DFT 环形水印的抗屏摄文档底纹水印算法在视觉质量和鲁棒性上均有提升，提出的抗屏摄文档水印的自动提取方案则实现了对屏摄照片中水印信息的自动提取功能，具体研究内容如下：

1) 提出了基于 DFT 环形水印的文档底纹水印算法，在保持文档可编辑性的同时，还拥有良好的视觉质量和鲁棒性。通过结合 DFT 频谱的平移不变性、线性和叠加特性，将多个 DFT 水印模板拼接在一起构成完整的水印底纹，无需再定位水印区域，当在完整水印底纹上截取大于模板尺寸的方形水印底纹时，其频谱中的水印圆环的半径和强度是等比放大的。并且其良好的抗裁剪性，也可以有效抵抗文字失真。将水印信息嵌入在纯色底纹的亮度分量中，在保持良好阅读观感的同时也保留了文档的可编辑性。本算法选择在底纹亮度分量的中频区域嵌入水印环，并进行了多重冗余设计，最终结合水印提取前的预处理操作便可以准确提取经历屏摄失真的文档水印。大量的实验结果表明，本算法可

以有效抵抗屏摄过程中的组合失真，并且优于现有的水印方案。

2) 提出了抗屏摄文档水印的自动提取方案，在保留已有方案优点的同时解决了水印提取前的人工预处理问题，并且实现了水印的盲提取。本算法在嵌入端设计了一种新的 DFT 环形水印模板，在定位环内外分别嵌入两个水印圆环，增加水印容量的同时也改进了水印提取算法。将提取端分为自动定位、自动矫正和自动提取三个步骤。通过充分利用定位环的高能特征以及文档的文本行特征，实现对屏摄照片的水印区域定位，变形文档的矫正以及变形水印环的恢复和提取。得益于定位环的多重定位功能，本方案不仅可以实现水印的自动盲提取，并且可以在仅拍摄部分文档内容时仍能准确提取水印信息。经大量实验证明，在 1 米距离内以及 65° 水平偏转角度的拍摄下，本方案可以准确地从屏摄照片中直接提取水印信息，并且还可以有效抵抗一些常见的图形处理攻击：JPEG 压缩、高斯模糊、中值滤波、尺度缩放、平移和裁剪攻击。

5.2 展望

尽管实验结果表明了本文所提出方法有着较高的文档图像视觉质量和抗屏摄鲁棒性，但在水印的嵌入以及自动提取算法上仍可以继续优化：

1) 提升水印的嵌入容量，重新设计水印模板，嵌入包含不同水印信息的多重水印圆环，根据应用场景适当减少纠错能力，选择更合适的 BCH 纠错编码多项式；

2) 提升文档图像的视觉质量和抗屏摄鲁棒性，结合深度神经网络，对不同的屏摄失真进行仿真以组成失真层，以水印文档图像的视觉质量和水印比特提取准确率为损失函数进行水印的嵌入训练，训练得到最优化的嵌入半径和嵌入强度，进一步提升视觉质量和鲁棒性。

3) 进一步优化文档水印的自动提取算法，仍然拆分为自动定位、自动矫正和自动提取三个步骤。自动定位可以通过创建不同拍摄条件下的水印文档屏摄照片数据集，结合深度神经网络训练分辨有无水印的频域区别，进一步实现更为精准的水印区域定位；自动矫正则可以继续优化矫正算法，以实现在垂直大偏转角度下仍可顺利矫正。

参考文献

- [1] Habes A. Information hiding in BMP image implementation, analysis and evaluation[J]. Information Transmissions in Computer Networks, 2006, 6(1): 1-10.
- [2] Zeki A M, Manaf A A. A novel digital watermarking technique based on ISB (Intermediate Significant Bit)[J]. World Academy of Science, Engineering and Technology, 2009, 50: 989-996.
- [3] Wang X Y, Zhao H. A novel synchronization invariant audio watermarking scheme based on DWT and DCT[J]. IEEE Transactions on Signal Processing, 2006, 54(12): 4835-4840.
- [4] 王向阳, 杨红颖, 赵红. 一种可抵抗 MP3 压缩的音频水印算法[J]. 自动化学报, 2007, 33(03): 248-252.
- [5] Lu C S, Chen J R, Fan K C. Real-time frame-dependent video watermarking in VLC domain[J]. Signal Processing: Image Communication, 2005, 20(7): 624-642.
- [6] Seo J S, Yoo C D. Image watermarking based on invariant regions of scale-space representation[J]. IEEE Transactions on Signal Processing, 2006, 54(4): 1537-1549.
- [7] Jalil Z, Mirza A M, Jabeen H. Word length based zero-watermarking algorithm for tamper detection in text documents[C]//Proceedings of the 2nd International Conference on Computer Engineering and Technology, April 16-18, 2010, Chengdu, China. Piscataway: IEEE. 2010: V6-378.
- [8] Malik A, Sikka G, Verma HK. A high capacity text steganography scheme based on LZW compression and color coding[J]. Engineering Science and Technology, 2017, 20(1): 72-79.
- [9] Pereira S, Pun T. Robust template matching for affine resistant image watermarks[J]. IEEE Transactions on Image Processing, 2000, 9(6): 1123-1129.
- [10] Kang X, Huang J, Shi YQ, et al. A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(8): 776-786.
- [11] Pramila A, Keskinarkaus A, Seppänen T. Multiple domain watermarking for print-scan and JPEG resilient data hiding [C]//Proceedings of the 6th International Workshop on Digital Watermarking, December 3-5, 2007, Guangzhou, China.

- Berlin: Springer, 2008: 279-293.
- [12] Kang X, Huang J, Zeng W. Efficient general print-scanning resilient data hiding based on uniform log-polar mapping[J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(1): 1-12.
- [13] Gourrame K, Douzi H, Harba R, et al. Robust print-cam image watermarking in fourier domain[C]//*Proceedings of the 7th International Conference on Image and Signal Processing*, May 30-June 1, 2016, Trois-Rivières, QC, Canada. Cham: Springer, 2016: 356-365.
- [14] Liang S, Wang X. Robust image watermarking in the print-cam process[C]//*Proceedings of the 19th International Conference on Communication Technology*, October 16-19, 2019, Xi'an, China. Piscataway: IEEE, 2019: 1626-1630.
- [15] Tancik M, Mildenhall B, Ng R. Stegastamp: Invisible hyperlinks in physical photographs[C]//*Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, June 16-20, 2020, California, USA. Piscataway: IEEE, 2020: 2117-2126.
- [16] Jia J, Gao Z, Zhu D, et al. Learning invisible markers for hidden codes in offline-to-online photography[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-24, 2022, New Orleans, USA. Piscataway: IEEE, 2022: 2273-2282.
- [17] Fang H, Zhang W, Zhou H, et al. Screen-shooting resilient watermarking[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 14(6): 1403-1418.
- [18] Lowe D G. Object recognition from local scale-invariant features[C]//*Proceedings of the 7th IEEE International Conference on Computer Vision*, September 20-27, 1999, Kerkyra, Greece. Piscataway: IEEE, 1999, 2: 1150-1157.
- [19] Chen W, Ren N, Zhu C, et al. Screen-Cam robust image watermarking with feature-based synchronization[J]. *Applied Sciences*, 2020, 10(21): 7494.
- [20] Chen W, Zhu C, Ren N, et al. Screen-cam robust and blind watermarking for tile satellite images[J]. *IEEE Access*, 2020, 8: 125274-125294.
- [21] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF)[J]. *Computer Vision and Image Understanding*, 2008, 110(3): 346-359.
- [22] Li L, Bai R, Zhang S, et al. Screen-Shooting resilient watermarking scheme via

- learned invariant keypoints and QT[J]. *Sensors*, 2021, 21(19): 6554.
- [23] DeTone D, Malisiewicz T, Rabinovich A. Superpoint: Self-supervised interest point detection and description[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition workshops*, June 8–22, 2018, Salt Lake City, UT, USA. Piscataway: IEEE, 2018: 224-236.
- [24] Moxey C E, Sangwine S J, Ell T A. Color-grayscale image registration using hypercomplex phase correlation[C]//*Proceedings of the IEEE International Conference on Image Processing*, September 22–25, 2002, New York, USA. Piscataway: IEEE, 2002: 247-250.
- [25] Wang X, Wang C, Yang H, et al. A robust blind color image watermarking in quaternion Fourier transform domain[J]. *Journal of Systems and Software*, 2013, 86(2): 255-277.
- [26] Deng B, Li S, Qian Z. An SVD-based screen-shooting resilient watermarking scheme[J]. *Multimedia Tools and Applications*, 2022, 81(23): 32841-32855.
- [27] Dong L, Chen J, Peng C, et al. Watermark-preserving keypoint enhancement for screen-shooting resilient watermarking[C]//*Proceedings of the IEEE International Conference on Multimedia and Expo*, July 18-22, 2022, Taipei, Taiwan. Piscataway: IEEE, 2022: 1-6.
- [28] Fang H, Chen D, Huang Q, et al. Deep template-based watermarking[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 31(4): 1436-1451.
- [29] Fang H, Chen D, Wang F, et al. TERA: Screen-to-Camera Image Code with Transparency, Efficiency, Robustness and Adaptability[J]. *IEEE Transactions on Multimedia*, 2021, 24: 955-967.
- [30] Jia J, Gao Z, Chen K, et al. RIHOOP: Robust invisible hyperlinks in offline and online photographs[J]. *IEEE Transactions on Cybernetics*, 2020, 52(7): 7094-7106.
- [31] Zhu J, Kaplan R, Johnson J, et al. Hidden: Hiding data with deep networks [C]//*Proceedings of the European Conference on Computer Vision*, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018: 657-672.
- [32] Bai R, Li L, Zhang S, et al. SSDeN: Framework for screen-shooting resilient watermarking via deep networks in the frequency domain[J]. *Applied Sciences*, 2022, 12(19): 9780.

- [33] Yakushev A, Markin Y, Obydenkov D, et al. Docmarking: Real-Time Screen-Cam Robust Document Image Watermarking[C]//Proceedings of the Ivannikov Ispras Open Conference, December 01-02, 2022, Moscow, Russian Federation. Piscataway: IEEE, 2022: 142-150.
- [34] Ge S, Fei J, Xia Z, et al. A screen-shooting resilient document image watermarking scheme using deep neural network[J]. IET Image Processing, 2023, 17(2): 323-336.
- [35] Aman M, Khan A, Ahmad B, et al. A hybrid text steganography approach utilizing Unicode space characters and zero-width character[J]. International Journal on Information Technologies and Security, 2017, 9(1): 85-100.
- [36] Chotikakamthorn N. Document image data hiding technique using character spacing width sequence coding[C]//Proceedings of the International Conference on Image Processing, October 24-28, 1999, Kobe, Japan. Piscataway: IEEE, 1999, 2: 250-254.
- [37] Huang D, Yan H. Interword distance changes represented by sine waves for watermarking text images[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(12): 1237-1245.
- [38] Choo H G, Kim W Y. Data-hiding capacity improvement for text watermarking using space coding method[C]//Proceedings of the 2nd International Workshop on Digital Watermarking, October 20-22, 2003, Seoul, Korea. Berlin: Springer, 2004: 593-599.
- [39] Kim Y W, Moon K A, Oh I S. A text watermarking algorithm based on word classification and inter-word space statistics[C]//Proceedings of the 7th International Conference on Document Analysis and Recognition, August 06, 2003, Edinburgh, UK. Piscataway: IEEE, 2003:775-779.
- [40] Borges PVK, Mayer J. Document watermarking via character luminance modulation[C]//Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, May 14-19, 2006, Toulouse, France. Piscataway: IEEE, 2006, 2: 317-320.
- [41] Brassil J, Low S, Maxemchuk NF, et al. Electronic marking and identification techniques to discourage document copying[J]. IEEE Journal on Selected Areas in Communications, 1995, 13(8): 1495-1504.

- [42] Brassil J, Low S, Maxemchuk NF. Copyright protection for the electronic distribution of text documents[J]. Proceedings of the IEEE, 1999, 87(7): 1181-1196.
- [43] Atallah M J, Raskin V, Hempelmann C F, et al. Natural language watermarking and tamperproofing[C]//Proceedings of the International workshop on Information Hiding, December 18, 2002, Heidelberg, German. Berlin: Springer, 2002: 196-212.
- [44] Sun X, Luo G, Huang H. Component-based digital watermarking of Chinese texts[C]//Proceedings of the 3rd International Conference on Information Security, November, 2004, New York, USA. New York: ACM, 2004: 76-81.
- [45] Bennett K. Linguistic steganography: Survey, analysis, and robustness concerns for hiding information in text[J]. Center for Education and Research in Information Assurance and Security, 2004-13.
- [46] Murphy B, Vogel C. The syntax of concealment: reliable methods for plain text information hiding[C]//Proceedings of the Security, Steganography, and Watermarking of Multimedia Contents IX, January 29-February 01, 2007, San Jose, CA. Bellingham: SPIE, 2007: 351-362.
- [47] Meral H M, Sankur B, Özsoy A S, et al. Natural language watermarking via morphosyntactic alterations[J]. Computer Speech & Language, 2009, 23(1): 107-125.
- [48] Lou M, Liu J. Watermarking text document based on structure and semantic of chinese characters[C]//Proceedings of the 7th International Conference on System of Systems Engineering, July 16-19, 2012, Genoa, Italy. CA: IEEE, 2012: 866-869.
- [49] Chang C Y, Clark S. Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method[J]. Computational Linguistics, 2014, 40(2): 403-448.
- [50] Wu M, Liu B. Data hiding in binary image for authentication and annotation[J]. IEEE Transactions on Multimedia, 2004, 6(4): 528-538.
- [51] Suzaki M, Suto M. A watermark embedding and extracting method for printed documents[J]. Electronics and Communications in Japan Part III-Fundamental Electronic Science, 2005, 88(7): 43-51.
- [52] 张驰. 二值文本图像数字水印技术研究[D]. 重庆:重庆大学, 2007. DOI:10.7666/d.y1140282.

- [53] 赵星阳, 孙继银, 李琳琳. 基于字符阶梯边沿调整的文本水印算法[J]. 计算机应用, 2008, 28(12): 3175-3178,3182.
- [54] Tirandaz H, Davarzani R, Monemizadeh M, et al. Invisible and high capacity data hiding in binary text images based on use of edge pixels[C]//Proceedings of the International Conference on Signal Processing Systems, May 15-17, 2009, Singapore, Singapore. CA: IEEE, 2009:130-134.
- [55] Yazdani H, Doostari M, Yazdani V. A new method to persian text watermarking using curvaceous letters[J]. Journal of Basic and Applied Scientific Research, 2013, 3(4).
- [56] Gugemann D, Sommer D, Lenders V, et al. Screen watermarking for data theft investigation and attribution[C]//Proceedings of the 10th International Conference on Cyber Conflict, May 30-June 01, 2018, Tallinn, Estonia. NY: IEEE, 2018: 391-408.
- [57] Fang H, Zhang W, Ma Z, et al. A camera shooting resilient watermarking scheme for underpainting documents[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(11): 4075-4089.
- [58] Yang X, Zhang W, Fang H, et al. Language universal font watermarking with multiple cross-media robustness[J]. Signal Processing, 2023, 203: 108791.
- [59] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [60] 袁占亭, 张秋余, 陈宁. 数字水印的鲁棒性分析与研究[J]. 计算机工程与设计, 2005, 26(3): 614-616. DOI:10.3969/j.issn.1000-7024.2005.03.019.
- [61] 方涵. 屏摄鲁棒水印方法研究[D]. 安徽:中国科学技术大学, 2021.
- [62] 陈玮彤. 遥感影像抗屏摄鲁棒水印模型与算法[D]. 江苏:南京师范大学, 2021.
- [63] Kitamura I, Kanai S, Kishinami T. Copyright protection of vector map using digital watermarking method based on discrete Fourier transform[C]//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, July 09-13, Sydney, Australia. Piscataway: IEEE, 2001, 3: 1191-1193.
- [64] Solachidis V, Pitas I. Watermarking polygonal lines using Fourier descriptors[J]. IEEE Computer Graphics and Applications, 2004, 24(3): 44-51.
- [65] Solachidis V, Nikolaidis N, Pitas I. Fourier descriptors watermarking of vector

- graphics images[C]//Proceedings of the International Conference on Image Processing, September 10-13, 2000, Vancouver, Canada. Piscataway: IEEE, 2000, 3: 9-12.
- [66] Piva A, Bartolini F, Boccardi L, et al. Watermarking through color image bands decorrelation[C]//Proceedings of the IEEE International Conference on Multimedia and Expo. Latest Advances in the Fast Changing World of Multimedia, July 30-August 02, 2000, York, USA. Piscataway: IEEE, 2000, 3: 1283-1286.
- [67] Thitimajshima P, Thitimajshima Y, Rangsanseri Y. Hiding confidential signature into digital images via frequency domain[C]//Proceedings of the 10th International Conference on Electrical and Electronic Technology, August 19-22, 2001, Singapore. Piscataway: IEEE, 2001, 1: 246-249.
- [68] Licks V, Hordan R. On digital image watermarking robust to geometric transformations[C]//Proceedings of the International Conference on Image Processing, September 10-13, 2000, Vancouver, Canada. Piscataway: IEEE, 2000:690-693.
- [69] Solachidis V, Pitas L. Circularly symmetric watermark embedding in 2-D DFT domain[J]. IEEE Transactions on Image Processing, 2001, 10(11): 1741-1753.
- [70] Chotikakamthorn N, Pholsomboon S. Ring-shaped digital watermark for rotated and scaled images using random-phase sinusoidal function[C]//Proceedings of the 10th International Conference on Electrical and Electronic Technology, August 19-22, 2001, Singapore. Piscataway: IEEE, 2001, 1: 321-325.
- [71] Guo H, Georganas N D. Multi-resolution image watermarking scheme in the spectrum domain[C]//Proceedings of the IEEE Conference on Electrical and Computer Engineering, May 12-15, 2002, Winnipeg, Canada. Piscataway: IEEE, 2002, 2: 873-878.
- [72] Poljicak A, Mandic L, Agic D. Discrete Fourier transform-based watermarking method with an optimal implementation radius[J]. Journal of Electronic Imaging, 2011, 20(3): 033008.
- [73] Bose R C, Ray-Chaudhuri D K. On a class of error correcting binary group codes[J]. Information and Control, 1960, 3(1): 68-79.
- [74] Sun X C, Lu Z M, Wang Z, et al. A geometrically robust multi-bit video watermarking algorithm based on 2-D DFT[J]. Multimedia Tools and Applications,

- 2021, 80: 13491-13511.
- [75] Tensmeyer C, Davis B, Wigington C, et al. Pagenet: Page boundary extraction in historical handwritten documents[C]//Proceedings of the 4th International Workshop on Historical Document Imaging and Processing, November 10-11, 2017, Kyoto, Japan. NY: ACM, 2017: 59-64.
- [76] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, October 5-9, 2015, Munich, Germany. Cham: Springer, 2015: 234-241.
- [77] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation[C]//Proceedings of the IEEE International Conference on Computer Vision, December 11-18, 2015, Santiago, Chile. NY: IEEE, 2015: 1520-1528.
- [78] Liang J, DeMenthon D, Doermann D. Geometric rectification of camera-captured document images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(4): 591-605.
- [79] Liang J, DeMenthon D, Doermann D. Flattening curved documents in images[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 20-25, 2005, Diego, USA. Piscataway: IEEE, 2005, 2: 338-345.
- [80] Zhu Y, Dai R, Xiao B, et al. Perspective rectification of camera-based document images using local linear structure[C]//Proceedings of the 23rd Annual ACM Symposium on Applied Computing, March 16-20, 2008, Fortaleza, Brazil. NY: ACM, 2008: 451-452.
- [81] Lu S, Chen B M, Ko C C. Perspective rectification of document images using fuzzy set and morphological operations[J]. Image and Vision Computing, 2005, 23(5): 541-553.
- [82] Takezawa Y, Hasegawa M, Tabbone S. Camera-captured document image perspective distortion correction using vanishing point detection based on radon transform[C]//Proceedings of the 23rd International Conference on Pattern Recognition, December 04-08, 2016, Cancun, Mexico. CA: IEEE, 2016: 3968-3974.

作者在攻读硕士学位期间发表的论文

- [1] **Zhang Y**, Wu H, Huang C, et al. Screen-shooting resistant robust document watermarking in the Discrete Fourier Transform domain[J]. International Journal of Network Management. 2024;e2278. doi:10.1002/nem.2278
- [2] Yang Z, **Zhang Y**, Zhang X, et al. Robust and High-Fidelity Image Watermarking with Mutual Mapping and Anti-Spectral Aliasing[J]. Journal of Electronic Imaging, 2024, 33(2): 023006. (SCI 检索)

作者在攻读硕士学位期间所作的项目

CCF-腾讯犀牛鸟基金项目，项目编号：CCF-Tencent Open Fund - RAGR20220126。

致 谢

时间如白驹过隙，三年的研究生学习生涯悄然而逝。在这段难忘的时间里，有幸与认真负责的导师们相遇，有幸与实验室同门一起努力奋斗，有幸与爱我助我的女友结识。值此论文完成之际，我想由衷感谢这些在我三年求学路上给予我帮助的老师、朋友和家人。

首先，我要感谢我的导师——张新鹏教授。幸得张老师的悉心教导，让我可以顺利完成研究生学业。当我在科研中遇到难题时，张老师总是能提出建设性的指导意见，对我的研究方向也有预见性的看法和独到的见解。张老师还为我们提供了良好的科研环境和科研设备，让我们能够更加专注于科研工作，并更加高效地对课题进行各种探索和尝试。

其次，我想感谢同样为我的科研工作提供极大帮助的吴汉舟老师。吴老师给我的研究方向提供了很多针对性的建议，也对我的研究课题给予了耐心的指导和许多细节上的建议，并且在方案设计和论文撰写方面也一直提供细心的指导。我在硕士期间取得的学术成果也离不开吴老师的辛勤付出。

再次，我想感谢实验室的李晨、张力、金玺、叶志诚、张烜博、杨紫云、汪悦等同门，三年的同门之谊使我们成为了科研路上的战友，尽管大家的研究方向各不相同，但都会不遗余力地互相帮助彼此，共同创造了实验室和谐互助的科研氛围。我还想感谢刘勇师兄，刘师兄非常的热心和无私，在我小论文的撰写过程中提供了极大助力。

然后，我想感谢陪伴我三年的女友王从娟，感谢她在生活上给予我的关心和爱护，抚慰我在遇到科研难题时焦躁的心，照顾在疫情期间感染病毒的我，陪伴我共度三年的求学时光。

最后，感谢各位评审专家在百忙之中抽出宝贵的时间审阅我的论文。