

中图分类号:

单位代号: 10280

密 级:

学 号: 20721452

上海大学



专业学位硕士学位论文

题 目	基于 Zernike 矩的视频鲁棒 水印技术研究
--------	-----------------------------

作 者 陈诗怡

学科专业 电子信息

导 师 吴汉舟

完成日期 2023 年 5 月

姓 名：陈诗怡

学 号：20721452

论文题目：基于 Zernike 矩的视频鲁棒水印技术研究

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：陈诗怡

学 号：20721452

论文题目：基于 Zernike 矩的视频鲁棒水印技术研究

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：_____

上海大学工程硕士学位论文

基于 Zernike 矩的视频鲁棒
水印技术研究

姓 名：陈诗怡

导 师：吴汉舟

学科专业：电子信息

上海大学通信与信息工程学院

2023 年 5 月

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

Robust Video Watermarking Based on Zernike Moments

M.A. Candidate: Shiyi Chen

Supervisor: Hanzhou Wu

Major: Electronic Communication

School of Communication and Information Engineering

Shanghai University

May, 2023

摘 要

近年来，随着互联网和多媒体技术的飞速发展，数字信息以图像、视频、文字等形式被广泛传播。为保证多媒体数据的知识产权，避免引起版权纠纷，数字水印技术得到了广泛应用。该技术可以在不影响数据感知效果的情况下，嵌入版权信息以保护媒体的知识产权。然而，视频作为一种常见的多媒体载体，除了具有图像的特性，还包含时间轴信息，且现有的视频鲁棒水印方法中，针对抵抗高强度旋转攻击的成果仍存在一定局限性。在此背景下，本文研究基于 Zernike 矩的视频鲁棒水印算法，取得主要成果如下：

(1) 针对抵抗高强度旋转攻击的问题，提出了一种基于归一化 Zernike 矩的视频鲁棒水印算法。该算法利用 Zernike 矩的旋转不变性以嵌入水印信息，且由于 Zernike 矩的抗噪声能力，算法也能很好地抵抗噪声攻击。此外，该算法还充分考虑了人类视觉系统的特点，采用色度通道嵌入水印，进一步减少水印对原始视频视觉上的影响。为保证水印的不可感知性，在水印嵌入过程中，采用了自适应矩选择策略，在实现良好的鲁棒性的前提下，提高了水印的不可感知性。实验结果表明，相比于现有算法，该算法平均峰值信噪比值提高了约 7dB，具有更高的不可感知性，且在缩放、噪声以及高强度旋转攻击下，也有较好的表现。

(2) 针对 Zernike 矩计算耗时高的问题，提出了一种基于 Zernike 矩的快速视频鲁棒水印算法。该算法针对 Zernike 矩的高计算量进行了优化，降低了计算消耗，使其适用于高清视频。算法首先将视频帧进行分组，按照一定规则，选择一定数量的帧对来嵌入水印。然后，应用自适应分块法，并对每个块应用奇异值分解，最后获得一个方形特征矩阵。接着，通过计算该特征矩阵的 Zernike 矩，将水印信息嵌入到对应的 Zernike 矩中，以实现更好的鲁棒性和不可感知性。最后，根据视频编码框架的特性，应用补偿机制，将帧对间的帧差覆盖为水印，以抵抗压缩和转码攻击。实验结果表明，该算法比现有的水印算法更具优势。

关键词： 视频鲁棒水印，Zernike 矩，视频，鲁棒性，版权保护

ABSTRACT

In recent years, with the rapid development of Internet and multimedia technology, digital information has been widely disseminated in the form of images, videos, and texts. In order to ensure the intellectual property rights of multimedia data and avoid causing copyright disputes, digital watermarking technology has been widely used. This technology can embed copyright information to protect the intellectual property rights of media without affecting the perceptual effect of data. However, video, as a common multimedia carrier, contains timeline information in addition to the characteristics of images, and there are still limitations in the results of existing video robust watermarking methods for resisting high-intensity rotation attacks. In this context, this dissertation studies the video robust watermarking algorithm based on Zernike moments, and the main results are obtained as follows:

(1) A video robust watermarking algorithm based on normalized Zernike moments is proposed for the problem of resisting high-intensity rotation attacks. The algorithm uses the rotational invariance of Zernike moments to embed watermark information, and it also resists noise attacks well due to the anti-noise capability of Zernike moments. To ensure the imperceptibility of the watermark, an adaptive moment selection strategy is used in the watermark embedding process, which improves the invisibility of the watermark while achieving good robustness. In addition, the algorithm takes into full consideration the characteristics of the human visual system and uses the chromaticity channel to embed the watermark, which further reduces the visual impact of the watermark on the original video. The experimental results show that the algorithm improves the average peak signal-to-noise ratio value by about 7 dB compared with existing algorithms, has higher imperceptibility, and performs better under scaling, noise, and high-intensity rotation attacks.

(2) A fast video robust watermarking algorithm based on Zernike moments is

proposed to address the problem of high computational time consumption of Zernike moments. The algorithm is optimized for the high computation of Zernike moments to reduce the computation consumption and make it suitable for HD video. The algorithm first groups video frames and selects a certain number of frame pairs to embed the watermark according to certain rules. Then, the adaptive chunking method is applied, the singular value decomposition is applied to each block, and a square feature matrix is obtained. Then, the watermark information is embedded into the corresponding Zernike moments by calculating the Zernike moments of this feature matrix to achieve better robustness and imperceptibility. Finally, the compensation mechanism is applied to cover the frame difference between frame pairs as a watermark according to the characteristics of video coding framework to resist compression and transcoding attacks. The experimental results show that the method is more advantageous than existing watermarking methods.

Keywords: Robust Video Watermarking, Zernike moment, Video, Robustness, Copyright Protection.

目 录

摘 要.....	I
ABSTRACT.....	II
第一章 绪论.....	1
1.1 课题来源.....	1
1.2 课题研究的背景和意义.....	1
1.3 数字水印.....	3
1.3.1 数字水印的基本概念.....	3
1.3.2 数字水印技术的评价指标.....	5
1.3.3 视频水印技术分类.....	6
1.3.4 视频水印技术研究概况.....	8
1.4 论文的研究内容与结构安排.....	10
1.4.1 主要研究内容.....	10
1.4.2 论文结构安排.....	11
1.5 本章小结.....	11
第二章 相关理论与技术基础.....	12
2.1 Zernike 矩.....	12
2.1.1 Zernike 矩技术概要.....	12
2.1.2 基于 Zernike 矩的鲁棒水印方法.....	15
2.2 H.264/AVC 压缩编码标准.....	16
2.2.1 编码框架.....	17
2.2.2 预测模式.....	18
2.2.3 变换和量化.....	20
2.3 常见攻击类型和抵抗方法.....	21
2.3.1 常见视频攻击类型.....	21
2.3.2 抵抗几何攻击的视频鲁棒水印方法.....	23

2.4	本章小结	25
第三章	基于归一化 Zernike 矩的视频鲁棒水印算法	26
3.1	引言	26
3.2	算法设计	27
3.2.1	总体架构	27
3.2.2	预处理	27
3.2.3	水印的嵌入过程	29
3.2.4	水印的提取过程	31
3.3	实验结果与分析	32
3.3.1	实验设置及评价标准	32
3.3.2	不可感知性分析	34
3.3.3	鲁棒性分析	35
3.4	本章小结	37
第四章	基于 Zernike 矩的快速视频鲁棒水印算法	38
4.1	引言	38
4.2	算法设计	39
4.2.1	总体架构	39
4.2.2	预处理	40
4.2.3	水印的嵌入过程	41
4.2.4	水印的提取过程	42
4.3	实验结果与分析	45
4.3.1	实验设置及评价标准	45
4.3.2	不可感知性分析	47
4.3.3	鲁棒性分析	47
4.3.4	对比分析	50
4.4	本章小结	54
第五章	总结与展望	55

5.1 总结	55
5.2 展望	56
参考文献.....	57
作者在攻读硕士学位期间公开发表的论文.....	65
作者在攻读硕士学位期间所作的项目.....	66
致 谢.....	67

第一章 绪论

1.1 课题来源

本课题来源于国家自然科学基金青年项目“社交网络多用户协同的行为隐写”(项目编号: 61902235)。

1.2 课题研究的背景和意义

随着新一轮科技革命和移动互联网技术的快速发展,文字、图像、音频和视频等数字媒体资源的获取已更加便利,多媒体数据也逐渐成为全球信息产业的重要组成部分。其中,由于视频压缩标准的不断演进,视频信息作为多媒体数据中容量最大的载体,已进入便于传播的时代^[1]。目前,“抖音”、“快手”等短视频的创作平台日渐兴起,更加快了大众对视频信息和产品的认知与使用。未来视频产业将成为体量庞大的基础产业,“信息视频化”的时代即将到来。

视频的全面普及丰富了人们的生产和生活,但由于移动互联网技术的发展和多媒体数据的易复制性,视频常被非法获取、篡改、传播,从而引发版权纠纷等问题。侵权的盗版视频常在网络中肆意流通,严重影响了视频作品产销环境。因此,视频数据的版权保护已成为了重要且刻不容缓的工作^[2]。为保护视频作品的版权,有专家和学者提出了数字水印技术,在证明版权归属的同时,保证数据的完整性。与加密技术相比,数字水印技术中水印以不可感知形式隐藏到视频内容中,不会影响数字产品的使用。它利用视频自身的信息冗余,将版权信息嵌入视频载体。在发生版权纠纷时,可以通过提取视频中的版权信息证明版权归属。此外,在视频被非法传播时,为追踪盗版责任者,可在分发视频前对同一视频的不同复制产品嵌入不同的水印,在提取水印时通过对应水印标识以确定责任者。视频水印技术作为保护视频信息安全的一种有效方法,已应用于电影、影视剧版权保护等场景,对维护版权所有者的合法权益具有重要的意义^[3]。由于视频文件在网络传输过程中常受到有意和无意的攻击,故为避免水印丢失,实现更好的安全性和有效性,研究对抗这些攻击的视频鲁棒水印,具有一定实用价值。其中,鲁

鲁棒性是指水印在受到攻击后仍能被正确提取。同时，视频水印也需要拥有不可感知性，即水印嵌入后，像素失真不可感知，且不降低数据质量^[4]。

此外，当前针对数字媒体的攻击方式层出不穷，针对视频鲁棒水印的攻击方式也越来越多样化，例如旋转、缩放、剪切等几何攻击，对视频鲁棒水印的提取造成了很大的困难^[5]。这些攻击方式使得水印嵌入的位置和信息被改变，难以提取出水印。此外，视频传输过程中的网络攻击也是鲁棒水印技术面临的一个挑战。例如，攻击者可能通过流量劫持、中间人攻击等方式篡改视频文件，使得视频内容被破坏，水印提取算法无法正确提取水印^[6]。综上所述，对于上述的攻击，传统的鲁棒水印技术对部分攻击鲁棒性尚有提升空间，因此需要鲁棒性更强的数字水印技术来应对这些攻击，这也是进一步研究视频鲁棒水印技术重要意义所在。

目前，对抗多种攻击的视频鲁棒水印领域已取得了重大进展。首先，现有研究从不同嵌入域提出视频鲁棒水印相关算法。其中一些工作采用了频域方法，如基于小波变换的水印算法^[7-8]和基于离散余弦变换的水印算法^[9-10]。也有一些基于空间域方法的研究工作，如基于可变分块的水印算法^[11]，基于边缘检测的水印算法^[12]等。这些方法都有各自的特点，可以根据实际应用场景的需要选择。其次，也有一些研究工作侧重于水印的感知，如基于色度域嵌入水印的算法^[13-14]，以最小化嵌入水印对视频质量带来的影响，提高水印的不可感知性。这些工作主要以嵌入不同的水印位置、使用不同的水印强度等作为研究点。此外，目前还有研究更关注于实现更强的鲁棒性，以应对更复杂的攻击。内容主要涉及完善水印算法，使其对各种攻击有更强的鲁棒性，如几何攻击^[15-16]、视频压缩攻击^[17-18]等。

由于 Zernike 矩是一种基于多项式的正交函数系，对旋转、平移和缩放具有不变性，这意味着即使媒体文件发生了旋转或平移等变换，其 Zernike 矩仍然保持不变^[19]。此外，它还具有良好的抗噪声能力，这些都使得 Zernike 矩在数字媒体处理中得到了广泛应用。所以在鲁棒水印技术中，也可以利用 Zernike 矩的这些优势，实现更好的水印鲁棒性和不可感知性，且目前基于 Zernike 矩的鲁棒图像水印已有较为成熟的研究成果^[19-23]，而基于 Zernike 矩的视频鲁棒水印仍有进一步研究的空间，这也是研究基于 Zernike 矩的视频鲁棒水印技术重要意义所在。

综上所述，视频鲁棒水印技术的研究已经取得了不少重要进展，但对于高强

度的几何攻击，仍需进一步研究，以保证算法的鲁棒性和不可感知性。此外，由于 Zernike 矩为对旋转、平移和缩放具有较强的不变性，且在视频鲁棒水印领域尚未广泛应用，因此，研究基于 Zernike 矩的视频鲁棒水印技术具有一定价值。

1.3 数字水印

1.3.1 数字水印的基本概念

在当今数字时代，数字媒体内容已经无处不在，如图像、音频、视频和文本。然而，这些数字媒体可以很容易地被复制和篡改，从而导致了許多版权纠纷。为了解决这个问题，需要可靠和有效的技术来识别和保护数字媒体的版权，同时不影响原始数字媒体的质量和信息的完整性。其中，数字水印作为数字版权保护的一项重要技术，也因此得到广泛的研究和应用^[24]。

数字水印技术和隐写术都来源于信息隐藏技术。信息隐藏^[25]是一种在媒体中嵌入额外信息而不造成数字媒介明显视觉失真的技术，其通用模型如图 1-1 所示。其中载体信号多为多媒体信号（图像、音频、视频等），随后通过嵌入算法，生成的加密后的媒介，再通过一个不安全的信道发送到指定的接收器。然后该接收器可以使用预共享密钥和提取算法从加密媒介中提取嵌入的额外信息。这些信息通常是用于保密通信的加密信息、可供识别内容所有者、内容使用者以及其他的认证信息。这些信息可以被授权用户提取和控制，以实现秘密通信、内容认证和内容保护等功能。信息隐藏技术根据其目的可以分为两个领域：用于秘密通信的隐写术和用于内容认证或版权保护的数字水印。虽然这两种方法都属于信息隐藏技术，但它们在定义、关键性能指标（如安全性和可靠性）等方面都有较大差异。

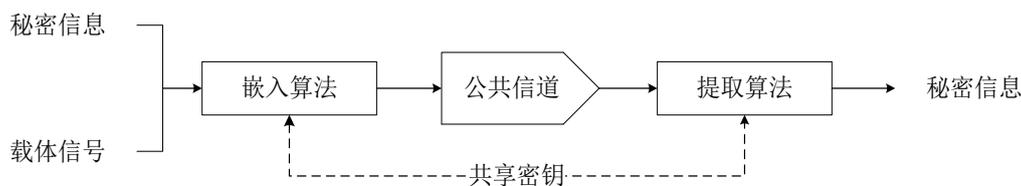


图 1-1 信息隐藏基本框架

隐写术^[26]，即信息的隐蔽传输，是一种隐蔽的通信技术，在这种技术中，秘密信息被嵌入到原始数据媒介中，而不明显改变媒介的内容和视觉效果。同时，隐写术也是一种秘密通信技术，在这种技术中，秘密信息被嵌入到原始媒介中，以便通过含有秘密信息的媒介来完成秘密信息的传输与通信。古希腊历史学家希罗多德^[27]曾在其著作中描述该技术：发送者在信使的头上写下秘密信息，等信使的头发长到足以覆盖时再传递信息，然后接收者剃除信使的头发后即可成功接收秘密信息。该技术早在 17 世纪就被用来发送秘密信息，并在上世纪的世界大战中仍作为一种秘密通信手段用于军事领域^[28]。

数字水印与隐写术不同，隐写术强调隐藏信息，而水印则侧重于通过信息的嵌入，来保护数字媒体^[29]。由于人们可以随时随地在网上下载和传播数字媒体，因此保护和管理数字媒体的需求越来越多，例如，通过监测非法传播的图像和视频来控制版权，从而确保商业和盈利，监测数字媒体的非法伪造等等。水印是应对这些挑战的一个有效工具。该技术可用于反盗版、追踪非法传播的媒体、文件认证、数字取证和隐私保护等领域。数字水印技术最重要的方面是嵌入和提取的信息的可靠性和准确性，同时确保嵌入和提取对原始数据的影响尽可能小。

数字水印可分为鲁棒水印和脆弱水印。鲁棒水印主要应用于版权保护、数字取证等场景，要求嵌入的水印能对抗各种常见的攻击：如压缩、旋转等，即在经历攻击后仍能保持嵌入信息的可提取性和完整性^[30]。而脆弱水印主要用于保护载体的完整性^[31]。相对于鲁棒水印，脆弱水印的优势在于能够更精准地检测信号的篡改，以保护数据的完整性和可靠性。由于脆弱水印对信号的变化很敏感，适用于检测发票和证明文件等不允许有细微改动的载体的内容变动情况。

鲁棒数字水印技术的实现必须考虑很多因素，包括嵌入规则、嵌入位置、嵌入强度、提取算法等^[32]。其中，嵌入原则是指数字水印在数字媒体中的嵌入算法规则，它直接影响到数字水印的可提取性和稳健性。嵌入位置指的是数字媒体中嵌入数字水印的部分，不同的嵌入位置也会影响数字水印的鲁棒性。嵌入强度是指嵌入数字媒体中的数字水印的强度。较高的嵌入强度可以提高数字水印的稳定性，但可能影响载体的质量。提取算法指的是从数字媒体中提取数字水印的算法。一个好的提取算法可以提高数字水印的提取能力和稳定性。

1.3.2 数字水印技术的评价指标

当评价数字水印技术时，通常会考虑以下几个方面的指标：鲁棒性、不可感知性、容量、安全性、计算效率等。这些指标都需要在实际应用中达到平衡，以达到最优的数字水印技术性能效果。以下对其进行详细介绍：

(1) 鲁棒性

鲁棒性表示数字水印算法对各种攻击的抵抗能力，即数字水印算法在受到各种攻击后，仍能正确地提取出数字水印信息的能力^[3]。首先，数字水印技术的鲁棒性通常表现在可以抵御各种攻击，例如噪声、压缩、滤波、亮度/对比度调整、几何攻击、裁剪攻击等。其次，水印算法的鲁棒性需要根据不同载体考虑。其中，音频水印的鲁棒性指标通常考虑抗噪能力、压缩和格式转换等。图像水印的鲁棒性指标通常包括抗攻击能力、图像处理和压缩等方面。而视频水印的鲁棒性指标则需要考虑针对视频的攻击，例如帧率、分辨率和编解码格式的修改。

(2) 不可感知性

不可感知性表示数字水印对原始媒体的影响程度，评价不可感知性的指标通常包括主观质量评估、客观质量评估等。主观质量评估是通过人的感知来衡量媒体质量，需要邀请一些受试者参与评估，以得出结论。客观质量评估是通过计算机算法对视频进行分析来衡量视频质量，这种方法的优点是可以自动化、高效、可重复性强。常见的客观质量评估方法包括 PSNR（峰值信噪比）^[33]、SSIM（结构相似性指数）^[34]、VQM（视频质量度量）^[35]等。此外，数字水印技术的不可感知性需要考虑嵌入的数字水印信息的大小和嵌入的位置。一般来说，数字水印信息越大，嵌入的位置就越难以保证不可感知性。其他如压缩、格式转换和传输等因素，也会对数字水印的不可感知性产生一定影响。

(3) 容量

容量表示数字水印在媒体中嵌入的信息量，通常受到透明度和鲁棒性的限制，所以在设计算法时需要在信息量与透明度和鲁棒性之间做出权衡。

(4) 安全性

安全性表示数字水印技术的安全性能，其中包括信息的保密性和完整性，数字水印技术需要保证水印信息不被未经授权的人所访问、复制和篡改，从而保证

信息的安全性。因此，安全性也是数字水印技术的重要指标。

(5) 计算效率

表示数字水印嵌入和提取的速度，即水印的处理时间需要在一个可接受的范围内。例如，在数字广播和音频实时流的应用中，涉及到实时数据处理，这时数字水印技术的计算效率变得尤为重要，数字水印必须能够在实时处理的情况下被嵌入和提取出来，以满足对于版权保护和内容安全的要求。因此，数字水印技术的计算效率可以被认为是衡量其应用价值的一个重要指标。除此之外，计算效率还与数字水印算法的复杂度和系统硬件的性能等因素密切相关。

1.3.3 视频水印技术分类

1993 年，Tirkel 等人^[36]首次提出了数字水印的概念。此后，数字水印技术引起了众多国内外学者的关注，并出现了许多关于数字水印的学术研究。近年来，基于视频的数字水印方法已取得了丰富的研究成果，并从灰度视频扩展至彩色视频。同时，该技术还可以根据不同的分类方式进行分类，如图 1-2 所示。常见的分类方式包括嵌入位置、提取方式和使用目的等。其中，基于嵌入位置的分类方式最为广泛，主要可将视频水印算法分为基于空域、压缩域和变换域等不同类型。

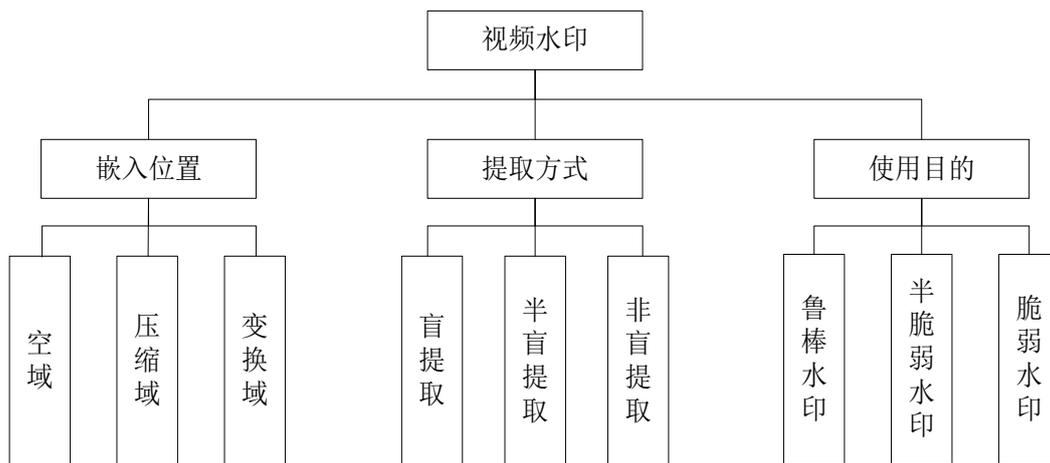


图 1-2 视频水印技术分类

基于空域的视频水印技术是指将水印信息嵌入至视频帧的某些特定区域中，例如边缘、纹理等。相对于其他嵌入域，空域水印技术对视频的影响较小，但是

由于嵌入的区域相对较少，因此水印的识别精度相对较低。

基于压缩域的视频水印是指将水印信息嵌入视频压缩编码的码流中，例如嵌入视频的 DCT 系数或运动矢量^[16]。由于视频在传输过程中，多使用经压缩的码流进行传输，所以直接将水印嵌入码流，具有较高的效率。但这类方法在应用时容易受到视频压缩算法的影响，即可能仅适用于一种压缩标准，具有一定局限性。

基于变换域的视频水印是指将水印信息嵌入到视频的某些变换域中，例如傅里叶变换域、小波变换域等。相对于其他嵌入域，变换域水印技术能够在一定程度上提高水印的鲁棒性和识别准确性。但是，这种方法需要对视频数据进行复杂的变换和处理，因此会增加水印的嵌入和提取计算复杂度。

若基于提取方式分类，现有方法具体可分为非盲提取、盲提取和半盲提取^[3]。其中，非盲提取是指通过比较接收方的原始视频信息和水印视频来提取水印的一种方法。但由于实际情况中，原始视频不易获得，且需要大量的存储空间，所以这类方法在具有较高的复杂性和使用成本。为解决这个问题，可以使用盲提取的方式，这类方法可以在不使用原始信息的情况下提取水印，更具普适性。半盲提取是指提取时不需要大量的原始视频数据，但需要原始水印信息或其他额外的信息。相对于盲提取和非盲提取而言，半盲提取是一个折中的解决方案。

若基于使用目的分类，现有方法可分为视频鲁棒水印、视频脆弱水印以及视频半脆弱水印^[37]。其中，视频鲁棒水印指视频即使受到蓄意或无意的攻击，水印也能正确地嵌入水印的视频中分离出，适用于保护知识产权、内容认证和版权管理等领域。脆弱和半脆弱水印经常被用作认证水印，即使用嵌入在载体视频中的水印被用来验证视频内容的真实性。视频脆弱水印的鲁棒性是所有级别中最弱的，这意味着嵌入的信息很容易被删除。半脆弱的视频水印是上述两种方式的折中方案，它既能抵抗常见的攻击，也能检测到恶意的数据改动。

此外，还有一些新兴的数字视频水印技术，比如基于深度学习的水印技术。这类技术利用了深度学习的特性，使用了大量的数据作为训练样本，在一定程度上提高了水印的鲁棒性和不可感知性^[30]。但由于深度神经网络的训练过程需要大量的数据支持和较高的设备性能需求，因此这也带来了较高的使用成本。此外，深度学习模型的黑盒特性也会使读者难以理解其决策，从而限制了算法的可解释

性。因此，在应用基于深度学习的水印技术时，需要仔细权衡上述限制。

1.3.4 视频水印技术研究概况

在现有的视频水印技术的研究成果中，最常见的分类方式是根据嵌入域进行分类，主要分为基于空域、压缩域和变换域的算法。在近几年的研究中，水印的鲁棒性和不可感知性也得到了较大的提升，且由于其相比深度学习方法而言，更为简单、快速、易于实现，从而得到了更为广泛的应用。

对于空间域视频水印技术，水印是通过修改视频帧的像素值来嵌入的。这类方法首先是由 Van^[37]等人提出的，该方法基于 LSB (least significant bit) 进行了数据嵌入，将水印序列嵌入到重要性最小的像素中。但在常见攻击下，重要性最小的像素最容易丢失，所以这种方案的鲁棒性并不理想。为增强该类方法的鲁棒性，Oh^[38]等将水印信息与辅助参考信息同时嵌入到与视频帧的相同空间位置中，使两个信息保持空间同步关系，达到抗几何攻击和同步攻击的目的。由于水印直接嵌入到帧中而无需任何变换，因此空间域方法在设计 and 执行时较为简单且计算效率相比其他方法更高。然而，这类方法的信息隐藏能力相对较低，对常见攻击（如有损压缩、裁剪）的鲁棒性有限，因为攻击过程中会丢失一些水印信息^[11]。

在压缩域，水印信息可以从两个位置进行嵌入，一种是在视频编码过程中嵌入水印信息；另一种是在符合视频编码标准的压缩码流中进行水印嵌入，实时性高，但水印的嵌入容量低^[39]。H.264/AVC 是目前最常用的视频压缩标准，它结合了各种提高压缩效率的新功能，具有更好的视频质量和更低的比特率。Tew 和 Wong^[40]对 H.264/AVC 压缩视频中的数字视频水印做出了概述，文中确定了水印嵌入发生的阶段，包括预测过程、变换、量化和熵编码，然后回顾了每个阶段的相关水印技术。Lee 等^[41]采用量化索引调制的方法修改视频部分解码后得到的低频 DCT 系数以嵌入水印信息，该方法具有一定鲁棒性但是无法抵抗几何攻击。

在变换域，常常通过改变变换域的系数来嵌入水印。这些方法包括奇异值分解 (SVD)、离散傅立叶变换 (DFT)、离散余弦变换 (DCT) 和基于小波的变换（例如离散小波变换 (DWT) 和双树复小波变换 (DT-CWT)）以及这些域的组合。Lai 等^[42]利用奇异值良好的稳定性，即对视频帧添加较小扰动时，这些值不

会发生显著变化，提出了一种基于 DWT 和 SVD 的水印方法。但是奇异值分解在计算上较耗时，且对 CPU 占用较大。基于 DCT 域的水印由于视频编码器利用 DCT 和量化的 DCT 系数来消除冗余，因此基于 DCT 域的水印方法对压缩是稳健的。目前大多数的工作是使用 DCT 的低子带或中子带来做水印^[43]。低子带更稳定，但于不可感知性尚有提高空间。相反，高子带能提供更好的视觉效果，但对量化不够稳健^[44]。对于双树复小波变换，Huan 等^[45]提出的方法是自动选择稳定的 DT CWT 系数来嵌入水印。实验表明，该方法可以抵御几何攻击。但视频的每一帧都要单独处理，且当旋转角度偏大时不具备鲁棒性。

然而，尽管近年来视频水印技术已取得了重大进展，但现有方法在应对几何攻击，如缩放、旋转以及转码攻击方面仍需要进一步的研究。几何攻击作为一种常见的攻击方式，通过破坏水印信息的同步性，导致水印提取准确率降低。由于不同视频网站采用的编码方式各异，因此在上传视频后，平台会对带水印的视频进行重压缩或转码。重压缩和转码都会使原视频丢失信息，导致水印无法提取。

为了能够抵抗几何攻击，Singh^[46]基于尺度不变特征变换(Scale-invariant Feature Transform)提出了抵抗旋转的视频水印方案。该方法在提取前，通过匹配 SIFT 特征点，计算旋转攻击的角度恢复视频帧。对旋转、裁剪等具有鲁棒性，但是对缩放不具有鲁棒性。王丽云^[47]利用整体 DCT 低频系数能够抵抗缩放攻击的特性，将水印信息嵌入在整体 DCT 系数的低频系数中，实现了抵抗缩放、裁剪和平移攻击的效果，但是整体 DCT 系数对旋转攻击不具有鲁棒性。

对于重压缩和转码攻击，Liu 等^[48]采用 BCH 综合码（BCH 码）技术对水印数据进行预处理，使提取的错误水印得以纠正。在重压缩过程中，该算法可以在 QP 不变的情况下抵抗重压缩攻击，但由于 BCH 码的纠错能力有限，当 QP 发生较大变化时，就不具备稳健性。Nair 等^[49]基于能量宏块（MB）嵌入水印，对视频转码具有一定稳健性，但抵御几何攻击的能力欠佳。

综上所述，近年来，视频鲁棒水印技术的研究已经取得较大进展，其主要分类方式可以根据水印嵌入位置分类，可分为基于空域、变换域和压缩域三大类。但现有算法中对于高强度几何攻击、转码攻击，以及对抗多种攻击融合的鲁棒性仍亟待进一步研究，以增加算法的应用范围。

1.4 论文的研究内容与结构安排

1.4.1 主要研究内容

本论文介绍了视频鲁棒水印的概念、研究进展和常用方法，并分析了现有方法的优缺点以及可以改进的方向。且从调研结果可知，针对高强度的几何攻击的视频水印算法仍亟待进一步研究。因此，本文基于 Zernike 矩研究视频鲁棒水印技术，并取得相应成果。首先，提出了一种基于归一化 Zernike 矩的视频鲁棒水印算法。其次，针对高清视频，提出了一种基于 Zernike 矩的快速视频鲁棒水印算法。两种算法的具体内容如下：

(1) 基于归一化 Zernike 矩的视频鲁棒水印算法

为解决高强度旋转攻击的问题，该算法利用 Zernike 矩的旋转不变性以嵌入水印信息，且由于 Zernike 矩的抗噪声能力，该算法也能很好地抵抗噪声攻击。为保证水印的不可感知性，在水印嵌入过程中，算法采用了自适应矩选择策略，选择合适的矩用于嵌入水印，最大程度均衡了鲁棒性和不可感知性。此外，该算法还充分考虑了人类视觉系统的特点，采用色度通道嵌入水印，以进一步减少水印对原始视频视觉上的影响。此外，为减小虚警率，本算法在提取水印时使用了多数表决法，以增强水印的鲁棒性。实验结果表明，相比于现有算法，该算法具有更高的不可感知性。同时，针对噪声攻击和高强度的几何攻击也有较好的表现。

(2) 基于 Zernike 矩的快速视频鲁棒水印算法

该方案针对 Zernike 矩的高计算量进行了优化，可用于高清视频并可以抵抗各种攻击。其中不仅包括高强度的几何攻击，还包含噪声攻击、压缩和转码攻击等。该算法首先将视频帧进行分组，按照一定规则，选择一定数量的帧对来嵌入水印。然后，应用自适应分块法，将每个帧分成不相交的块，并对每个块应用奇异值分解，最后获得一个方形特征矩阵。接着，计算该特征矩阵的 Zernike 矩。与计算原视频帧的 Zernike 矩相比，自适应分块法最大程度增加了计算效率。随后，将秘密信息嵌入到使用自适应矩选择策略选定的 Zernike 矩中，以实现更好的鲁棒性和不可感知性。最后，根据视频编码框架的特性，应用补偿机制，将帧对间的帧差覆盖为水印，以抵抗压缩和转码攻击。实验结果表明，该算法在鲁棒

性、计算效率、水印不可感知性等方面，较现有的水印算法更具优势。

1.4.2 论文结构安排

本论文是以作者攻读硕士学位期间承担课题的工作为基础，旨在研究基于 Zernike 矩的视频鲁棒水印技术，在充分调研了国内外已有的鲁棒水印算法的基础上，结合了视频的特性和不同攻击的特点，基于 Zernike 矩展开研究，提出了两种视频鲁棒水印算法。本文的各章内容安排如下：

第一章为绪论，首先介绍了课题的来源。其次，介绍了本文的研究背景及意义。然后，详细介绍了数字水印的基本概念、评价指标，以及视频水印的分类方法和研究概况。最后，介绍了论文的主要研究成果和各章的内容安排。

第二章是本论文的理论基础部分，主要介绍 Zernike 矩的基本概念，H.264/AVC 视频编码标准相关框架，以及常见的视频水印攻击类型和抵抗方法。

第三章介绍了一种基于归一化 Zernike 矩的视频鲁棒水印算法，其核心思想是使用归一化 Zernike 矩实现水印的嵌入和检测。与现有算法相比，该算法利用 Zernike 矩的旋转不变性和噪声稳健性，更好地对抗高强度的几何和噪声攻击。

第四章介绍了一种基于 Zernike 矩的快速数字视频水印算法，适用于高清视频，并对多种攻击具有较强鲁棒性。与现有算法相比，该算法在能够抵抗高强度几何攻击的同时，对压缩和转码攻击也有较高的鲁棒性，且具有较高的计算效率。

最后，第五章总结了本文的工作，并对工作中可以改进的地方进行了展望。

1.5 本章小结

本章为论文的绪论，分别介绍了课题的来源、论文的研究背景及意义和数字水印的基本概念。其中，着重介绍了数字水印的相关理论、发展状况，以及评价指标，随后阐述了视频水印的技术分类和研究概况，在列出国内外现有的研究成果的基础上，简要分析其优缺点，引出本文的主要研究方向。最后，介绍了论文的主要研究内容及成果，并阐述了论文的结构安排。

第二章 相关理论与技术基础

本章将介绍与论文相关的理论与技术基础。主要内容包括 Zernike 矩, H.264/AVC 编码标准, 以及常见的视频攻击方式及解决方法。

2.1 Zernike 矩

本文的技术路线主要围绕 Zernike 矩展开。本节主要介绍 Zernike 矩相关理论, 包括 Zernike 矩的定义、计算流程及其性质。此外, 介绍了基于 Zernike 矩的鲁棒水印研究现状, 并分析现有方法的优缺点及未来可以研究的方向。

2.1.1 Zernike 矩技术概要

在数学和统计学中, 矩是对变量分布和形态特点的一组度量, 不同的矩使用不同的基函数用于描述函数^[50]。由于 Zernike 矩的基函数为正交基, 矩信息具有冗余度低、描述能力强且各阶矩相互独立的特点, 因此这些矩与图像的独立特征相对应。并且在图像含有噪声的情况下, Zernike 矩与几何矩相比, 具有更强的鲁棒性^[51], 适用于鲁棒水印的研究。该节主要介绍 Zernike 矩的技术概要, 包括其计算过程和相关定义, 以及 Zernike 矩的性质, 具体内容如下:

(1) Zernike 矩的计算过程及相关定义

在图像与视频处理领域, 常使用 Zernike 多项式作为基函数, 通过计算 Zernike 矩来描述和分析图像或物体的特征。Zernike 多项式是定义在极坐标系下的复数多项式, 是一类用于描述圆形区域内的函数的正交基函数。每个 Zernike 多项式由其阶数 n 和重复度 m 决定, 其数学表达式为:

$$V_{n,m}(x,y) = R_{n,m}(\rho)e^{jm\theta} \quad (2-1)$$

其中, n 为自然数, m 为整数, 且 $0 \leq |m| \leq n$, $n - |m|$ 为偶数, $|m|$ 表示重复度 m 的绝对值; (x,y) 表示直角坐标系中的一点, ρ, θ 和 x, y 的关系可以表示为:

$\rho = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(y/x)$; $R_{n,m}(\rho)$ 为正交的径向多项式, 其表达式为:

$$R_{n,m}(\rho) = \sum_{s=0}^{n-|m|/2} \frac{(-1)^s (n-s)! \rho^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \quad (2-2)$$

当计算完毕 Zernike 多项式后，可以得出连续图像 $f(x,y)$ 的以 n 为阶数， m 为重复度的 Zernike 矩值，其计算公式为：

$$A_{n,m} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x,y) V_{n,m}^*(\rho,\theta) dx dy \quad (2-3)$$

其中， $V_{n,m}$ 为以 n 为阶数， m 为重复度的 Zernike 多项式，*表示共轭。

对于离散图像，可以使用求和的方式求得相应矩值，具体计算公式为：

$$A_{n,m} = \frac{n+1}{\pi} \sum_{\rho \leq 1} \sum f(\rho,\theta) V_{n,m}^*(\rho,\theta) \quad (2-4)$$

(2) Zernike 矩性质

Zernike 矩具有一些重要的性质，使其能够广泛应用于图像与视频处理。其性质主要包括逆变换、计算效率以及旋转不变性，具体介绍如下：

a) 逆变换

根据 Zernike 矩的定义，由于 Zernike 多项式具有正交性，且表示图像圆形区域内的信息，因此可以由逆变换重构图像内切圆中的数据。对于离散图像而言，其重构表达式为：

$$\hat{f}(x,y) = \sum_{n=0}^N \sum_{\{m:n \geq |m|, n-|m| \bmod 2=0\}} A_{n,m} V_{n,m}(\rho,\theta) \quad (2-5)$$

其中， $\hat{f}(x,y)$ 表示重构后的图像。值得注意的是，由于 Zernike 矩的阶数无法达到无穷大，所以根据公式(2-5)求出的结果只能是近似重构值，且当阶数越大时，重构的图像就越接近原始图像，如图 2-1 所示。

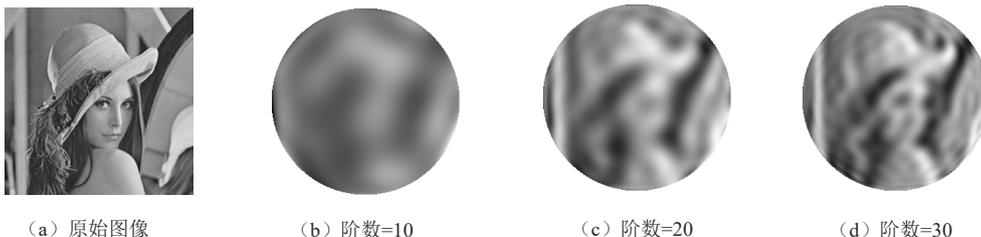


图 2-1 使用不同阶数重构的 Lena 图

在图 2-1 中，使用分辨率为 512×512 的灰度 Lena 图作为原始图像，分别以最大阶数 10、20、30 计算图像的 Zernike 矩并进行重构。由图 2-1 可知，由 Zernike 矩重构的图像均包含在图像内切圆内，当最大阶数越高时，重构图像越清晰。

b) 计算效率

Zernike 矩的计算效率由所求矩的最大阶数以及图像的分辨率大小所决定。随着阶数的增加，图像分辨率的增加，Zernike 矩在计算时就需要更多的时间^[52]。该结论在图 2-2 中得到验证，该图以不同分辨率的 Lena 图像为例，说明了在不同阶数和图像分辨率下，计算 Zernike 矩所需的时间。

实验中，为了获得不同尺寸图像时的计算时长，将 512×512 的 Lena 图通过双三次插值法^[53]分别转化为分辨率为 32×32 ， 64×64 ， 128×128 ， 256×256 的图像，并使用不同的阶数计算 Zernike 矩。由图 2-2 可知，图像分辨率越大，且阶数越大时，需要消耗的执行时间就越长。

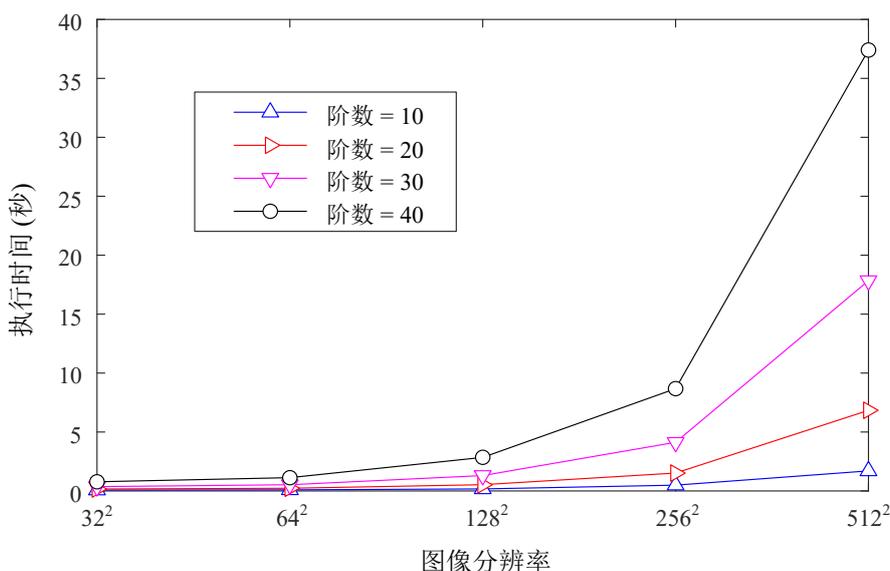


图 2-2 阶数和图像分辨率与 Zernike 矩计算执行时间的关系

c) 旋转不变性

由公式 (2-4) 可求得图像的 Zernike 矩 $A_{n,m}$ ，根据定义，可将其简化为极坐标形式表示，其表达式为：

$$A_{n,m} = |A_{n,m}| e^{jm\theta} \quad (2-6)$$

此时，假设将原始图像顺时针以 α 作为角度，进行旋转，对应的表达式为：

$$A'_{n,m} = A_{n,m} e^{-jm\alpha} = |A_{n,m}| e^{jm(\theta-\alpha)} \quad (2-7)$$

由公式 (2-7) 可以得出：当图像旋转一定角度后，影响的是 Zernike 矩的相位值，对应的幅度保持不变。所以 Zernike 矩的幅度可以作为图像的旋转不变特征，应用于数字水印领域，以抵御旋转攻击。

此外，将图像归一化后求得的 Zernike 矩具有尺度不变性，可以用来抵御尺度变换攻击；同时由于 Zernike 多项式具有正交性，使图像各阶 Zernike 矩之间彼此独立，因此对图像形状有较强的表达能力，可以达到较强的抗噪声效果^[19]。

2.1.2 基于 Zernike 矩的鲁棒水印方法

由于 Zernike 矩具有较强的旋转不变性和噪声鲁棒性，且通过归一化变换可以抵抗常见几何攻击，因此，国内外专家对 Zernike 矩在鲁棒水印中的应用进行了大量的研究。目前，Zernike 矩已广泛应用于图像鲁棒水印领域，具体描述与分析如下：

一般而言，现有的 Zernike 矩的鲁棒水印方法可以根据 Zernike 矩的使用方式分为两类。一种是将水印嵌入 Zernike 矩中^[19-23]，另一种是将 Zernike 矩用于在水印嵌入、提取之前的预处理部分^[55-57]。对于前者，主要的应用难点在于 Zernike 矩的计算效率，例如，许多现有的方法是计算整个图像的 Zernike 矩来得出水印的嵌入载体。但由图 2-2 可知，图像的分辨率越大，Zernike 矩的阶数越高，需要的计算时间越大，且当图像分辨率为 512×512 ，最大阶数为 30 时，计算一张图片的 Zernike 矩已需要近 20 秒的时间，所以此类方法将带来非常高的计算成本。对于后者，这类方法通常是使用 Zernike 矩来计算图像受到旋转攻击前和攻击后带来的角度差，以在后续步骤中进行补偿。由于不需要计算所有符合条件的 Zernike 矩，所以计算成本较第一类较低，但这类方法通常是非盲水印，需要原始的媒体数据作为参考才能提取水印，在实际应用上具有一定局限性。

此外，这些方法还可以根据算法是否将 Zernike 矩与其他变换结合，分为两类。第一类^[19-23]只使用 Zernike 矩，并没有使用其他变换。这类算法虽然对常见几何攻击有较高的鲁棒性，但是对其他攻击，如压缩攻击、信号处理攻击仍较为

敏感。第二类方法是将 Zernike 矩与其他变换结合^[55-57], 如 DWT (Discrete Wavelet Transform) 或 DCT (Discrete Cosine Transform)。相比于第一类方法, 第二类方法具有更高的鲁棒性, 但也提高了计算成本。

综上所述, 基于 Zernike 矩的鲁棒图像水印已有许多研究成果, 可以根据不同的需求选择合适的方法。然而, 由于视频是由许多静止的图像组成的, 而大多数采用 Zernike 矩的图像水印方法的计算量都较大, 因此现有的图像水印方法不能直接扩展到视频。此外, 由于视频总是通过特定的压缩标准进行有损压缩, 简单地将 Zernike 矩应用于视频不能确保较高的鲁棒性, 所以需要进一步研究基于 Zernike 矩的视频鲁棒水印。

目前关于 Zernike 矩的视频鲁棒水印已有一些研究成果, 但总体数量较少, 且主要用于分辨率较低的视频。He 等人^[58]提出了一种基于伪 Zernike 矩归一化的视频水印方案, 可以在抵抗旋转和缩放攻击方面取得良好的性能。然而, 由于视频的分辨率通常长宽不一致, 而该方法主要将水印信息嵌入到视频帧的中心圆区域, 仅适用于长宽一致的视频, 缺乏普适性。为解决上述问题, Xu 等人^[59]在视频中间裁剪了一个方形区域用于水印嵌入, 但是该方法当视频被缩放时, 不具备鲁棒性。Yuan 等人^[60]提出了一种基于特征提取和局部 Zernike 矩的新型视频水印方案, 在抵御信号处理和几何攻击方面具有良好的性能。但它将秘密数据嵌入所选视频帧的亮度通道中, 与色度通道相比, 影响了水印的不可感知性。此外, 由于 Zernike 矩的计算量很大, 且上述方法没有关注于如何增加计算效率, 所以算法执行时间较长, 且计算消耗较高, 并不适用于高清视频。因此, 尽管该领域目前已有了较大进展, 但仍需要进一步研究, 以保证计算效率和算法的鲁棒性。

2.2 H.264/AVC 压缩编码标准

视频由于其容量较大的特性, 在传播时通常需要进行压缩编码, 以节省存储、传输成本。因此, 在视频水印的设计中, 需要熟悉视频编码的原理及其带来的影响。在众多的视频压缩编码标准中, H.264/AVC 因其低码率、高质量、高网络适应性, 应用最为广泛^[39]。本节主要介绍该编码标准的主要框架, 并对其中包含的预测模式, 变换和量化步骤进行了详细介绍。

2.2.1 编码框架

H.264/AVC 标准分为视频编码层 VCL (Video Coding Layer) 以及网络提取层 NAL (Network Abstraction Layer), 前者主要负责将视频内容编码; 后者主要负责根据网络传输要求, 打包和传送编码后的视频数据。本节主要讨论 H.264/AVC 的视频编码层, 其编码框架如图 2-3 所示。该框架可以主要分为四个关键过程, 即预测、变换、量化和熵编码^[61]。以下对这四个步骤进行简要介绍。

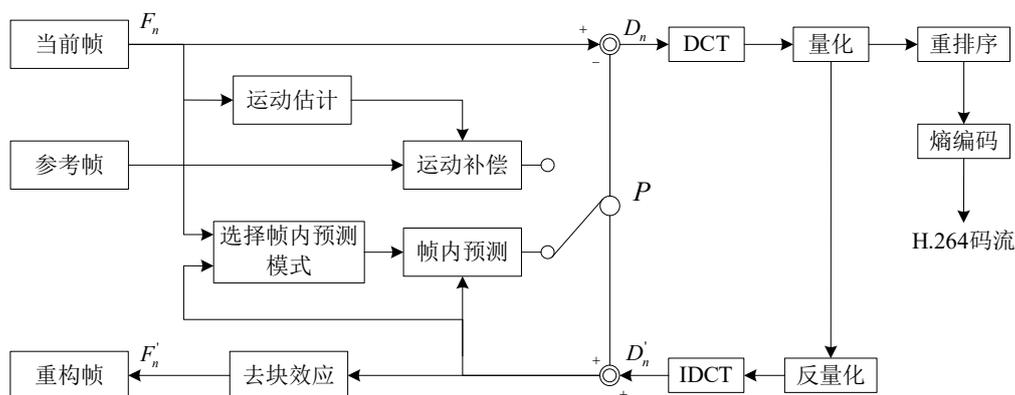


图 2-3 H.264/AVC 编码框架

(1) 预测

在 H.264/AVC 编码器中, 主要以预测的方式减少冗余信息。该方法主要利用像素间、帧间的相似性, 计算视频帧各采样点的预测值, 并与实际值相减得到差值, 使得最终 H.264 码流传输的是采样点的预测值和实际值之间的差值, 而不是像素点原本的采样值, 以达到压缩视频大小的目的。其中, 主要的预测方式可以分为帧内 (Intra) 和帧间 (Inter) 预测。帧内预测是指使用目标帧的帧内信息进行预测, 该方法利用帧内相邻像素相近的特性, 进行预测并获得残差; 帧间预测主要使用相邻帧的信息进行预测, 主要利用相邻帧之间, 相同位置像素值相近的特性进行预测, 以获取并存储残差, 用于后续步骤。

(2) 变换

获取残差后, 需要对相应的残差矩阵进行变换, 以进一步压缩数据大小。变换的目的是将图像信息由空域转化到频域。由于图像中直流分量和低频分量通常

占比较高，而高频分量占比较小，因此变换可以将图像信号的能量大致集中在低频区域中，以便后续压缩。此外，由于在 MPEG 标准中使用的 8×8 DCT 变换经常会导致失配问题，因此，H.264/AVC 视频编码标准中使用 4×4 整数 DCT 变换将残差矩阵变换至频域，以进一步保证压缩后的视频质量。

(3) 量化

量化是 H.264/AVC 编码过程中唯一的有损过程，它根据人类视觉系统划分量化表，以压缩视频的高频成分。由于高频成分主要表示视频的细节信息，因此可以在不明显降低视觉质量的情况下，提高视频压缩效果。

(4) 熵编码

熵编码是一种无损压缩编码，可以降低信源数据间的相关性，从而实现压缩。在 H.264/AVC 视频编码标准中，支持变长编码和算术编码两种熵编码方式。这两种编码方式均利用了信源的统计特性，将量化后的预测残差的 DCT 系数编码到 H.264 码流中，从而实现码率压缩。

2.2.2 预测模式

预测是视频编码的核心技术之一。在 H.264/AVC 视频编码标准中，预测前，首先会将视频帧分为若干个片，再以片中的宏块为基本单位进行预测。其中，预测模式主要可以分为两类，分别是帧内预测模式和帧间预测模式^[61]。

(1) 帧内预测模式

帧内预测是指用当前宏块周围的像素值来预测当前宏块。如图 2-4 所示，在大小为 4×4 分块中，小写字母代表待预测像素值，大写字母代表其周围的像素值。在预测完当前块后，将当前块和预测块作差，并保存差值用以压缩视频。

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

图 2-4 4×4 子块的帧内待预测像素及其相邻像素

通常情况下，帧内相邻像素之间存在相关性，这也是导致帧中存在大量的冗余信息的原因。帧内预测可以利用这种关联性来以压缩视频。预测后，如果当前块和预测块之间差值越小，压缩效率越高。在 H.264/AVC 视频编码标准中，使用不同的模式来预测当前块的像素值，具体预测模式由图 2-5 所示。

以大小为 4×4 的亮度分块为例，在这种情况下，主要存在 9 种预测模式^[61]，分别用数字 0 到 8 对其进行标识，依次代表 8 种 AC 预测模式和一种 DC 预测模式。其中模式 0 和 1 为 AC 预测模式，分别表示将对应的相邻像素值覆盖至预测块、模式 2 为 DC 预测模式，表示将相邻的像素，即图中的 A~D 和 J~L 代表的像素值相加取平均，作为预测值。模式 3 到 8 是 AC 预测模式，需要根据图中所示的不同方式使用参考像素 A~M，使用加权平均以得到待预测的像素。

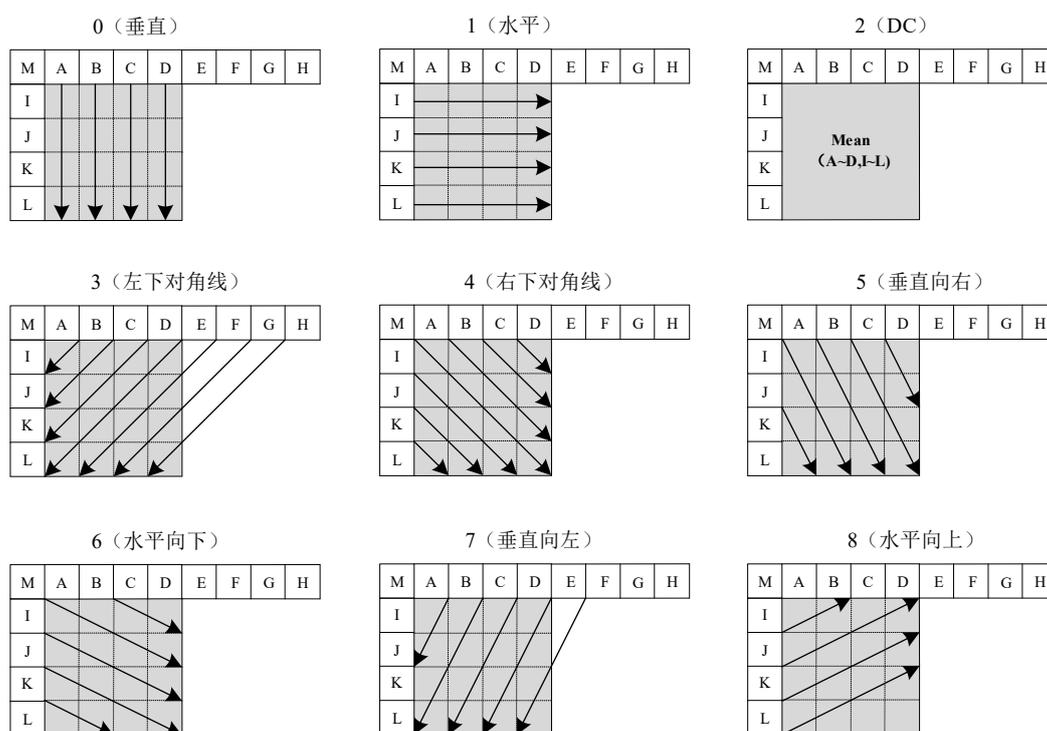


图 2-5 4×4 子块的 9 种帧内预测模式示意图

(2) 帧间预测模式

与帧内预测不同，帧间预测主要用于减少空间冗余，由于视频的相邻帧间具有一定的相似性，所以在帧间预测中，通过从参考帧中寻找最优的参考块作为预测块，然后计算两者的偏移和差值以减少冗余。在进行预测前，需要将视频帧内

的宏块进行划分。一般有 8 种划分模式可供选择，如图 2-6 所示。对于纹理越多的图像，宏块划分越细，而对于纹理平坦的图像，可以选择较大的划分尺寸。

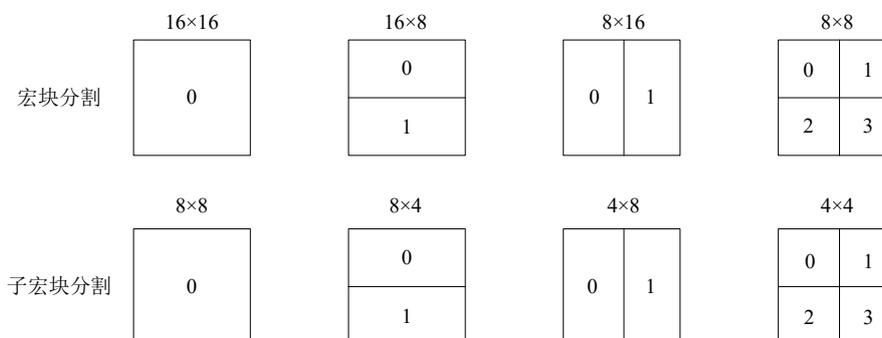


图 2-6 帧间预测宏块划分模式

帧间预测主要包括运动估计 (Motion Estimation, ME) 和运动补偿 (Motion Compensation, MC) 两个步骤。其中运动估计也称为运动搜索，是指在相邻帧中找出与当前像素块最相似的块作为匹配块，用于进行运动补偿，即使用搜索到的预测块与当前编码块的对应像素相减，获得残差块，并结合匹配块与当前块的相对位移，即运动矢量 (Motion Vector, MV) 来进一步描述残差，如图 2-7 所示。

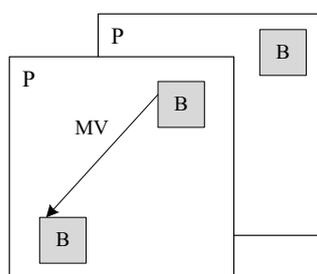


图 2-7 H.264/AVC 编码器的运动矢量示意图

2.2.3 变换和量化

为了进一步提高压缩效率，在 H.264/AVC 编码器中，需要对残差块进行变换和量化。其过程如下：首先，将预测残差矩阵经过 4×4 整数 DCT 变换，再判断该宏块是否是色度块或是 16×16 的亮度块。若是 16×16 的亮度块，则需要提取其直流系数组成矩阵，对其进行 4×4 的 Hadamard 变换并量化；若是色度块，则进行 2×2 的 Hadamard 变换并量化。若以上两种情况均不是，在完成 4×4 的 DCT

整数变换后，直接对其进行交流系数量化。

2.3 常见攻击类型和抵抗方法

鲁棒性是一个评价鲁棒数字水印方法好坏的重要指标，目前设计的鲁棒数字水印方法应对常见的攻击具有鲁棒性。本节主要介绍常见攻击类型和抵抗方法。

2.3.1 常见视频攻击类型

对于视频鲁棒水印技术，常见的攻击可以分为以下几类：

(1) 几何攻击

几何攻击是指对视频帧进行缩放、旋转、平移、镜像、裁剪等攻击^[62]。旋转攻击指对视频帧进行旋转变换，包括顺时针和逆时针旋转；平移攻击指对其进行平移变换，包括水平和垂直平移；镜像攻击指对视频帧进行镜像变换，包括水平和垂直镜像；缩放攻击指对帧进行缩放变换，包括放大和缩小。

在缩放攻击中，攻击者会对载体图像进行缩放操作，从而使水印信息发生形变，例如振幅和相位的变化^[62]。缩放攻击通常用一个缩放因子 s 来控制。假设载体图像为 $I(x, y)$ ，其宽为 w ，高为 h ，缩放后的图像 $I_1(x, y)$ 为：

$$I_1(x, y) = I(s_x x, s_y y) \quad (2-8)$$

其中， s_x 和 s_y 分别是 x 和 y 方向上的缩放因子。当 $s_x \neq s_y$ 时，缩放操作会使载体图像在 x 和 y 方向上发生不同的缩放，从而导致水印信息在不同方向上发生不同程度的形变，进而影响水印的提取。

旋转攻击则是通过对载体图像进行旋转来改变水印信息的相位和振幅^[62]。在实现方式上，通常是将载体图像旋转一个固定的角度。设载体图像为 $I(x, y)$ ，旋转后的图像 $I_2(x, y)$ 为：

$$I_2(x, y) = I(R(x, y)) \quad (2-9)$$

其中， $R(x, y)$ 是一个二维旋转矩阵，控制图像旋转的角度，通常表示为：

$$R(x, y) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2-10)$$

旋转操作会使旋转后的图像可能会导致数字水印失真、位移或者完全丢失。载体图像在不同位置上的像素值发生改变, 从而影响水印信息的相位和振幅, 从而降低水印提取的准确性。

(2) 噪声攻击

噪声攻击是一种常见的数字水印攻击方式, 它通常会在水印数据中引入一些随机噪声或干扰, 导致质量下降, 同时会破坏嵌入的水印信息, 从而破坏水印的可读性或提取性能。噪声攻击主要分为加性噪声和乘性噪声两种。

加性噪声攻击是指将随机噪声加到水印信号上^[63], 使其失真或无法读取。这种攻击方式通常通过在水印信号中添加高斯噪声或椒盐噪声来实现。高斯噪声是指按照正态分布随机分布的噪声, 它具有平滑的性质; 而椒盐噪声是指将一些像素点变成黑色或白色, 从而模拟帧图像中的噪声点。加性噪声攻击会降低水印的信噪比, 从而影响水印的可读性和提取性能。

乘性噪声是指噪声与原始信号的乘积而不是加和的一种噪声形式^[64]。在乘性噪声攻击下, 嵌入的数字水印会随着载体信号的变化而变化, 使得水印信号的振幅和相位都发生改变, 从而影响到水印的鲁棒性。由于噪声与原始信号乘积的影响, 水印提取时可能会出现失真或无法提取的情况, 从而降低水印提取的准确性。

(3) 压缩攻击

在视频鲁棒水印领域, 压缩攻击是指将带有数字水印的视频进行压缩, 使其文件大小更小, 从而达到节省存储空间或网络带宽的目的。这种压缩可能会改变视频的一些特征, 如帧率、分辨率、码率等, 从而使得数字水印信号失真或无法提取, 降低了数字水印的鲁棒性^[40]。

通常, 在数字水印嵌入到视频中后, 由于视频容量较大, 需要进行压缩才能便于传输。其中, 压缩可以通过采用不同的编解码标准来实现, 如 MPEG 和 H.264 等。这些标准在压缩视频时通常会采用预测、变换和量化等步骤, 从而减少视频的冗余信息并降低文件大小^[61]。然而, 这些步骤会对数字水印信号造成破坏, 并导致水印信号无法提取或失真。目前, 网络中传播的多媒体数据都会经过压缩处

理, 视频可能还会按照网站要求的编码方式重新进行编码压缩。由于压缩带来的量化噪声以及水印强度的衰减可能会使视频中水印提取失败。

(4) 码率变化攻击

码率变化攻击通常指修改视频的码率。码率, 即比特率 (bits per second, bps) 是指视频文件中单位时间内的平均数据量, 其单位通常是 Kbps 或 Mbps。一般来说, 在相同分辨率下, 视频文件的比特率越高, 压缩率越低, 视频帧的质量就越好。反之则越粗糙。因此, 当嵌入水印的视频的比特率降低时, 嵌入视频序列中的水印信息也会随之改变, 从而导致水印无法提取。

(5) 其他视频处理攻击

其他的视频处理攻击主要包括色彩增强、锐化、超分辨率、模糊等, 这些攻击均可能影响水印信息, 导致水印图像内部信息发生变化从而破坏水印信号的同步性, 使得水印的提取效果受到影响, 最终导致提取结果错误。其中, 对于模糊攻击而言, 基于特征点检测的水印算法^[46,65]会受到较大的影响, 因为在受到模糊攻击后, 视频帧的细节部分在模糊中丢失, 相应的检测算法中提取特征点的匹配度将降低, 从而导致检测到错误的特征点, 使得水印提取错误。

2.3.2 抵抗几何攻击的视频鲁棒水印方法

几何攻击作为一种常见的攻击方式, 是视频水印领域的一大难点。其本质是通过扰乱水印与载体之间的同步性, 从而导致水印提取准确率降低。目前抵抗几何攻击的视频鲁棒水印技术主要通过以下方式实现:

(1) 穷举搜索法

该方法主要是对受到几何攻击后的帧尝试所有可能的攻击类型, 比如定义可能的失真及其参数范围后, 测试所有提出的组合, 以求出水印最可能产生的预期失真, 以获取最高的水印提取准确率^[66-68]。但是, 由于失真种类繁多, 且不同参数会带来不同的影响, 所以需要大量的组合用于搜索, 这也是导致此类方法计算量大的原因。即随着搜索空间的增加, 计算次数也会增加, 带来较大的计算量, 从而使得计算效率低, 且计算消耗较大。另一方面, 多次使用同一水印检测器也将出现较高的误报率。可见, 虽然穷举搜索法的思路较为简单, 但实际应用中需

要较大的使用成本，因此，此类方法应用较少。

(2) 几何矫正

这个方法的基本思想是，在提取视频水印之前，先计算出每一帧图像所受到的几何攻击的参数和类型，然后对接收到的视频进行逆变换，再提取水印信息。其中，主要采用的方法是同步模板法，即在水印嵌入时，除了嵌入水印之外，还嵌入一个用于同步校正的模板。在提取水印时，首先根据同步模板计算宿主图像受到的几何攻击参数和类型，然后对水印图像进行相应的逆变换，实现水印的同步后再提取水印信息。例如，Niu 等人^[69]使用同步模板法携带多个时间轴上分布的像素点位置信息的模板进行水印嵌入，即使遭受几何攻击也可以正确检测到模板。通过模板可以找到新的位置信息，从而保证了水印检测在空间上的同步性。然而，该算法选择水印嵌入点的方法过于复杂，而且不能抵抗各种空间同步攻击，这些攻击可能导致水印嵌入点在帧内产生相对偏移。

(3) 几何不变量

这类方法的思想是将水印信息嵌入在具有旋转、缩放和平移不变性(Rotation, scaling and translation, RST)的特征上，以提高水印的鲁棒性。例如，基于变换域的方法^[42-45,70]和基于 Zernike 矩的方法^[58-60]等。其中，Huan 等人^[45]提出了一种基于双树复小波的视频水印方法，该方法对每一帧进行多级双树复小波分解，并将水印嵌入到分解后的稳健子带中。该方法利用双树复小波分解的无损重构、平移不变和方向选择等特点，在面对几何攻击时表现良好。但双树复小波并不具备旋转和缩放等几何不变性，因此只能对一定强度的几何攻击做出抵抗。Liu 等人^[70]提出了一种基于一维傅里叶变换(1D-DFT)和 Radon 变换的视频水印算法。他们将视频序列划分为许多图像组，然后在时间轴上对每个划分好的组做 1D-DFT 变换，这种变换不仅能够完整保存视频的空域信息，还能得到时域信息。接着在该时频域中应用 Radon 变换，最后选择最高时频来嵌入水印。1D-DFT 变换使得水印能够抵抗视频压缩处理，而基于 Radon 变换的水印嵌入则使得水印能够抵抗几何攻击。此外，Zernike 矩的视频水印方法在上一节中已有详细介绍，这里不再赘述。总之，这类方法的优点在于无需确定和恢复几何攻击，但如何选择合适的特征、提取和应用这些特征仍需进一步研究。

(4) 特征点

该类方法提取载体中具有几何不变特性的特征点，由于此类特征点具有稳定性，所以在遭受攻击后，仍能准确地提取水印信息。常见的方法为基于 SIFT 特征点的方法，如文献[65]，该方法使用 *contourlet* 变换将视频的每一帧转换成不同方向的低频带。通过改变低频带系数直方图的分布来嵌入水印信息。为了提高鲁棒性，水印提取时使用尺度不变特征变换 (SIFT) 进行几何形变校正，然后将低频带系数直方图与原始视频进行比较。该算法具有很好的不可感知性，并且能够抵抗常见的几何攻击。然而，其缺点是水印提取需要使用原始视频。因此，这类方法需要使用原始载体来提取水印信息，不能进行盲提取。此外，如果载体收到了模糊攻击等对特征点影响较大的攻击，特征点在检测时就容易出现误检或漏检，从而导致水印提取的错误。

2.4 本章小结

本章主要围绕本文提出的研究方案，提供相关理论基础。内容主要包括 Zernike 矩、H.264/AVC 压缩编码标准的基本知识，以及视频鲁棒水印领域中常见的攻击类型和抵抗方法。关于 Zernike 矩，本章详细介绍了其定义、性质和计算方法，并阐述了它在数字水印领域的应用情况。对于 H.264/AVC 压缩编码标准，本章介绍了其基本框架以及主要步骤。最后，介绍了视频鲁棒水印领域中常见的几种攻击类型，包括几何攻击、噪声攻击和压缩攻击，并详细讨论了它们的特点和影响，以及抵抗几何攻击的常用方法。

第三章 基于归一化 Zernike 矩的视频鲁棒水印算法

本章研究针对高强度旋转攻击的视频鲁棒水印技术,提出了一种基于归一化 Zernike 矩的视频鲁棒水印算法。首先介绍了该算法的整体架构,并详细阐述了算法的预处理及水印嵌入和提取步骤,最后通过实验验证算法的有效性。

3.1 引言

根据第二章对数字水印中 Zernike 矩的理论基础和现有算法的介绍,可以了解到, Zernike 矩具有旋转不变性,能够有效抵抗高强度旋转攻击,可以应用于数字水印。然而,目前大部分使用 Zernike 矩的算法都基于图像,而在视频领域中的应用仍较为缺乏,且现有的算法均在不同方面存在一定的局限性,如只适用于长宽一致的视频、只具有旋转不变性、对其他常见攻击的鲁棒性较差等。因此,这促使我们提出一种新的视频水印方案,该方案能够在保持高强度旋转鲁棒性的同时,对其他常见攻击也具有鲁棒性,且适用于通用分辨率的视频序列。

在本章中,提出了一种基于归一化 Zernike 矩的视频鲁棒水印算法。为了达到抵抗高强度旋转攻击的目的,使用 Zernike 矩作为载体,作为稳定不变的特征值,用于嵌入和提取水印。为解决现有使用 Zernike 矩的算法只适用于长宽一致视频,以及对缩放攻击不具备鲁棒性的问题,使用归一化方法,将视频序列进行预处理。对于水印的嵌入和提取,使用改进后的 QIM (Quantization Index Modulation) 方法,即 DM-QIM (Dither Modulation-QIM)。在 QIM 方法中,发送端主要将水印信息通过使用量化器将载体的系数调制到不同的索引区间,然后接收方通过这些索引区间,以识别水印信息^[71]。为更好地权衡鲁棒性和水印的不可感知性,应用改进后的 QIM 方法,即 DM-QIM,用于水印嵌入和提取。在该算法中,使用抖动矢量将载体调制到不同的群组中,以提高水印的鲁棒性,同时减少原视频失真。

此外,根据人类视觉系统特性,亮度的失真比色度的失真更明显^[72],所以为了尽量减小水印对视觉上的影响,该算法将水印信息嵌入到原始视频序列的色度 U 通道中,以提高水印的不可见性。实验结果表明,该算法保持了良好的视觉质

量,且具有较高的鲁棒性,与现有算法相比,该算法除了具有较强的旋转鲁棒性,还可以抵抗其他常见攻击,且适用于通用分辨率的视频序列,在鲁棒性和失真之间也取得了更好的权衡。

3.2 算法设计

3.2.1 总体架构

本章所提算法的总体架构如图 3-1 所示,包括视频序列预处理、水印信息的嵌入和水印信息的提取三个主要步骤。

为了提高算法对高强度旋转攻击的鲁棒性,本算法采用了 Zernike 矩作为水印嵌入的载体。为了保证嵌入水印后的视频质量,方案考虑将水印信息嵌入视频的色度分量中。另外,由于 Zernike 矩仅能表示对应图像内切圆区域的特征,因此在嵌入水印前,采用自适应归一化方法将对应帧转化为方形矩阵,以便进行后续特征提取。在水印信息的嵌入阶段,使用了 DM-QIM 方法,将水印信息嵌入对应帧的 Zernike 矩的模值内,并进行一系列逆变换,得到包含水印信息的 YUV 视频序列。在提取水印信息时,重复了上述预处理工作,并使用 DM-QIM 方法中的提取步骤,即可成功提取出水印信息。

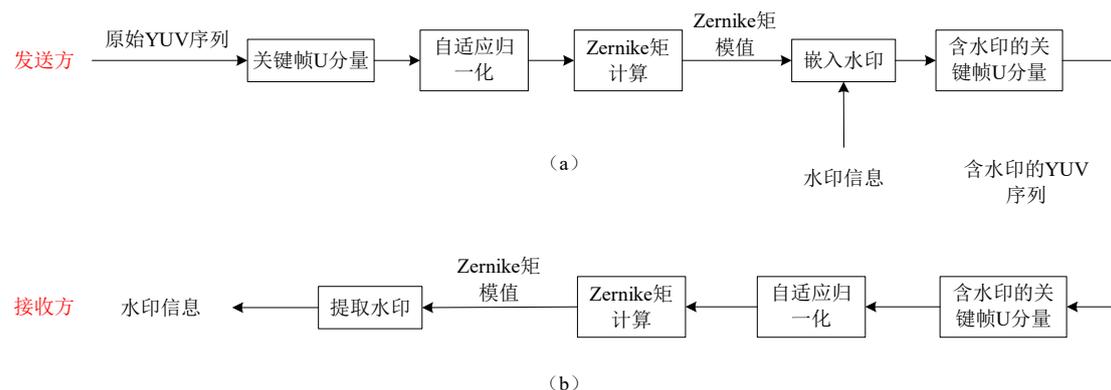


图 3-1 基于归一化 Zernike 矩的视频鲁棒水印算法总体架构 (a) 水印嵌入 (b) 水印提取

3.2.2 预处理

如图 3-1 总体架构所示,算法的预处理部分主要包括选择关键帧、提取色度

分量、自适应归一化以及 Zernike 矩计算部分。其中选择关键帧包含两个步骤，即首先将视频帧分组，使每个组内包含相同数量的帧，其次，在每个组内选择一个关键帧用于水印的嵌入。提取色度分量表示提取 YUV 视频序列中的色度分量 U，用于后续水印嵌入。其中自适应归一化的步骤如下所示：

步骤一：假设 $f(x, y)$ 表示待处理的视频关键帧。首先计算该帧的中心点 (\bar{x}, \bar{y}) ，其计算公式为：

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (3-1)$$

其中， m_{10}, m_{01} 和 m_{00} 表示 $f(x, y)$ 的相应矩，其计算公式为：

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad (3-2)$$

随后，将 $f(x, y)$ 转换为 $f_1(x, y) = f(x - \bar{x}, y - \bar{y})$ ，使对应关键帧居中。

步骤二：在 x 方向对 $f_1(x, y)$ 应用剪切变换至 $f_2(x, y)$ ，其计算公式为：

$$f_2(x, y) = \mathbf{A} f_1(x, y) \quad (3-3)$$

其中， \mathbf{A} 为 $\mathbf{A}_x = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}$ ； β 根据文献[73]，可由 $\mu_{30}^{(1)} + 3\beta\mu_{21}^{(1)} + 3\beta^2\mu_{12}^{(1)} + \beta^3\mu_{03}^{(1)} = 0$

求得，用以确保 $\mu_{30}^{(2)}$ 为零； $\mu_{pq}^{(i)}$ 表示对应帧的中心距，可表示为：

$$\mu_{pq}^{(i)} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f_i(x, y) \quad (3-4)$$

其中， M 和 N 分别表示对应帧图像的长和宽。

步骤三：在 y 方向对 $f_2(x, y)$ 应用剪切变换至 $f_3(x, y)$ ，其计算公式为：

$$f_3(x, y) = \mathbf{A} f_2(x, y) \quad (3-5)$$

其中， \mathbf{A} 为 $\mathbf{A}_y = \begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix}$ ； γ 根据文献[73]，可由 $\gamma\mu_{20}^{(2)} + \mu_{11}^{(2)} = 0$ 求得，使 $\mu_{11}^{(3)}$ 为零。

步骤四：将 $f_3(x, y)$ 在 x 和 y 两个方向上进行缩放，使最终分辨率大小为

$K \times K$ ，用以计算对应的 Zernike 矩。 K 的值是根据视频帧原始尺寸自适应设置的。例如，当输入视频序列的 U 通道的分辨率为 176×144 时，通常可以将 K 设置为 256，以在保留信息和精度之间达到平衡。 K 的值可以根据不同的需求进行相应调整。

根据文献[73]可得，图像和其仿射变换在经过上述归一化步骤后将获得相同的图像。因此，在应用上述归一化步骤后，再结合第二章对 Zernike 矩性质的介绍，对应帧在经过归一化后，其 Zernike 矩的振幅对旋转和缩放攻击均保持不变。

Zernike 矩的计算过程已在第二章中有详细介绍，在本章就不加以赘述。

3.2.3 水印的嵌入过程

在计算完毕 Zernike 矩后，首先需要删除一部分不适用于嵌入水印的矩，以保证算法的适用性和优越性最大化，再进行水印的嵌入步骤。具体描述如下：

(1) 筛选适用于嵌入水印的矩

假设所求得的 Zernike 矩为 $A_{n,m}$ ，其对应的删除规则有三，具体如下所示：

规则一：根据文献[70]中描述，当 Zernike 矩的重复度 m 为 4 的倍数时，即 $m = 4j$ ， j 为整数时，所求得的 Zernike 矩值将偏离正交性，即无法准确计算。因此，为了保证水印的鲁棒性，需要在嵌入步骤前删除这部分矩。

规则二：根据文献[71]中描述， $|A_{00}|$ 和 $|A_{11}|$ ，即 A_{00} 和 A_{11} 的幅值，与对应图像无关。因此，由于 A_{00} 和 A_{11} 中不包含图像信息，这些矩值不能合适地描述关键帧的特征，故不适用于水印嵌入，需要在嵌入前删除。

规则三：由公式 (2-4) 可得： $|A_{n,m}| = |A_{n,-m}|$ ，因此，为减少嵌入水印后，水印信息对原始帧带来的修改，在嵌入时将不考虑后者，以达到更高的不可感知性。

上述三条删除规则提供了有效的特征选择机制，帮助筛选适用于嵌入水印的 Zernike 矩。通过这些规则，待处理的视频关键帧的特征可以更可靠地被描述，从而在算法的鲁棒性和不可感知性之间寻求平衡，以最大化算法的使用效果。这些删除规则在水印嵌入和提取过程中起到了关键作用，确保所选取的 Zernike 矩

能够提供准确而稳定的特征表示，从而增强了水印算法的性能和可靠性。

(2) 嵌入步骤

选择完毕适用于嵌入的矩后，需要将水印信息嵌入关键帧中对应的矩。嵌入水印共有四个关键步骤，具体描述如下：

步骤一：使用 DM-QIM 方法，将水印数据嵌入对应关键帧中，所有被选中的矩的振幅，且每个目标帧内包含 1 比特的水印数据，其具体计算表达式为：

$$|A_{n,m}^w| = \begin{cases} Q(|A_{n,m}|, \Delta) & , w=0; \\ Q(|A_{n,m}| + \Delta/2, \Delta) - \Delta/2 & , w=1. \end{cases} \quad (3-6)$$

其中， $|A_{n,m}^w|$ 的上标 w 表示该振幅已嵌入水印信息。 $Q(x, \Delta)$ 表示量化器，其数学表达式定义如下：

$$Q(x, \Delta) = \Delta \times \text{round}\left(\frac{x}{\Delta}\right) \quad (3-7)$$

其中，函数 $\text{round}(\cdot)$ 用于计算与输入值最接近的整数； w 代表水印位，取值为 0 或 1； Δ 表示量化步长，需要根据 $|A_{n,m}|$ 的值来设置。

步骤二：构建水印信号 \mathbf{R} ，其表达式如下：

$$\mathbf{R}(x, y) = \sum_{n,m} \left(\frac{|A_{n,m}^w|}{|A_{n,m}|} - 1 \right) \times A_{n,m} V_{n,m}(\rho, \theta) \quad (3-8)$$

其中， (x, y) 表示水印信号中对应的像素坐标， $V_{n,m}(\rho, \theta)$ 表示 Zernike 多项式。

步骤三：使用系数 α ，将水印信号叠加至目标帧中，其数学表达式为：

$$\mathbf{I}^w = \mathbf{I} + \alpha \mathbf{I}_w \quad (3-9)$$

其中，系数 α 用于控制水印的嵌入强度，以增强水印的不可感知性，其表达式为：

$$\alpha = \frac{\Theta \mathbf{I}}{\Theta \mathbf{I}_r} \quad (3-10)$$

其中，函数 $\Theta(\cdot)$ 用于计算输入的方形矩阵的内切圆区域中，所有像素点数据的均值； \mathbf{I}_r 表示对原始帧的重构，其表达式为：

$$\mathbf{I}_r(x, y) = \sum_n \sum_m^{N_{\max}} A_{n,m} V_{n,m}(\rho, \theta) \quad (3-11)$$

其中， N_{\max} 表示计算 Zernike 矩时使用的最大阶数。

步骤四：对嵌入水印的特征矩阵 \mathbf{I}^w 根据上述步骤进行一系列逆变换，得到嵌入水印后的关键帧。对每个关键帧重复上述步骤后可得到嵌入水印后的视频序列。

3.2.4 水印的提取过程

如图 3-1 (b) 所示，在执行水印提取步骤前需要进行预处理，预处理步骤已详细描述于 3.2.2 节。在完成预处理后，进行水印提取。提取水印共有四个关键步骤，具体如下所示：

步骤一：同嵌入过程所述，在计算完毕对应矩阵的 Zernike 矩后，首先需要删除不适用于嵌入水印的矩，具体删除标准同嵌入过程一致。

步骤二：对每个关键帧中，经过筛选的 Zernike 矩取模值。将模值作为载体以提取水印。假设 $|B_{n,m}|$ 表示嵌入水印的 Zernike 矩的模值， b 代表提取的水印位，水印提取表达式为：

$$b = \operatorname{argmin}_{w_b \in \{0,1\}} \left| |B_{n,m}| - |B_{n,m}^{w_b}| \right| \quad (3-12)$$

其中，使用函数 $\operatorname{argmin}(\cdot)$ 可以得到使输入值最小时使用的自变量； w_b 为水印的可能取值，取值范围为 0 或 1； $|B_{n,m}^{w_b}|$ 用于模拟嵌入水印，其计算表达式为：

$$|B_{n,m}^{w_b}| = \begin{cases} Q(|B_{n,m}|, \Delta) & , w_b = 0; \\ Q(|B_{n,m}| + \Delta / 2, \Delta) - \Delta / 2, & w_b = 1. \end{cases} \quad (3-13)$$

其中，上标 w_b 表示在 $|B_{n,m}^{w_b}|$ 中嵌入的水印位； Δ 代表量化步长，需要与嵌入步骤时使用相同的参数值。

步骤三：在提取完毕该目标帧内所有选定矩中嵌入的水印后，为了减少提取结果存在突变的可能，将所有矩中选择频率最高的一个作为每个目标帧的最终提取水印位，以减少误码率，提高提取结果准确性。

步骤四：对所有关键帧重复上述步骤，并将提取结果按顺序组合，即可得到嵌入的水印信息。

3.3 实验结果与分析

3.3.1 实验设置及评价标准

本文中的所有实验都是在 Matlab R2016a 的环境中实现的，计算机系统为 64 位 Windows 10，内存 8GB，且使用 2.3GHz 英特尔酷睿 i5 处理器。为了客观地评估本章算法，使用六个以 CIF 为格式（352×288）的标准视频序列作为实验数据集，即 Akiyo, Foreman, Hall, Mother and Daughter, Paris 和 Silent。每个测试视频包含 300 帧，可以通过网址 <http://trace.eas.edu/yuv/index.html> 获取。同时，Huan 等人[45]所提出的视频鲁棒水印方案也在实验中被复现，用以进行对比分析。

实验中，对于本章算法，将测试视频在 U 通道的帧图像归一化为分辨率为 256×256 的矩阵，并设置 GOP（Group of Picture）长度为 6；再使用密钥，用于伪随机生成水印。水印长度设置为 50，以保证每个 GOP 可携带一个水印位。实验中，将水印嵌于一组 GOP 的第一帧。为了公平比较，将文献[45]中的 GOP 也相应设定为 6，嵌入强度 T 设置为 400，即文中推荐值。

为获取量化步长 Δ 的最佳设定值，使用上述六个标准测试视频序列设计实验，以评估使用不同量化步长时的鲁棒性。实验中，仅改变量化步长，测试不同量化步长下水印的 NCC（Normalized Cross Correlation）值，（NCC 常用于评价算法的鲁棒性，其定义将在后文详细描述）。相关实验结果如图 3-2 所示，从中可以得出，NCC 值随着量化步长 Δ 的变化而变化，且当量化步长 Δ 设置为 30000 和 40000 时，水印的 NCC 达到最大值。因此，30000 和 40000 可以作为最佳量化步长应用于实验，以保证较高的鲁棒性。在后续讨论中，若无特殊说明，量化步长 Δ 将设置为 30000。

关于不可感知性，实验中使用 PSNR 和 SSIM 作为评价标准。PSNR 全称为“Peak Signal-to-Noise Ratio”，即峰值信噪比，是一种衡量数字图像或视频质量的指标。它主要是通过对比原始图像或视频和被处理后的图像或视频之间的差异来计算的，其表达式为：

$$PSNR = 10 \cdot \log_{10} \frac{max^2}{E}, E = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n |I(i, j) - K(i, j)|^2 \quad (3-13)$$

其中，若以图像为例， max 表示图像中像素点的最大可能取值，通常为 255； E 表示图像 \mathbf{I} 与图像 \mathbf{K} 的均方误差； m 和 n 分别表示图像 \mathbf{I} 与 \mathbf{K} 的长和宽。

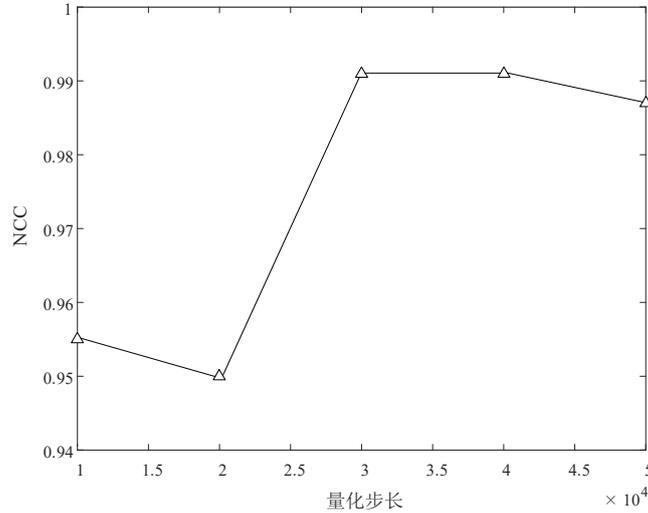


图 3-2 使用不同量化步长时提取水印的 NCC 值

PSNR 的单位是分贝 (dB)，其数值越高，表示图像或视频的质量越好，失真程度越低。常见的 PSNR 阈值为 30dB，一般认为超过 30dB 的 PSNR 值就可以被视为无明显失真的图像或视频^[33]。

SSIM 是结构相似性衡量指标 (Structural Similarity Index)，它用于评估两帧之间的相似度^[34]。其计算公式为：

$$SSIM(\mathbf{X}, \mathbf{Y}) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3-14)$$

其中， \mathbf{X} 和 \mathbf{Y} 可表示视频的两帧， μ_x 和 μ_y 分别表示 \mathbf{X} 和 \mathbf{Y} 的均值， σ_x 和 σ_y 表示方差， σ_{xy} 表示 \mathbf{X} 和 \mathbf{Y} 的协方差， c_1 和 c_2 是两个常数，用于避免分母为 0。SSIM 指标的取值范围为 [0,1]，其中，‘1’表示两帧完全相似，‘0’表示两帧完全不同。

SSIM 指标的计算基于人类视觉系统在观看图像或视频时会提取其中的结构化信息的假设。其中，将结构化信息定义为独立于亮度、对比度的反映场景中物体结构的属性，并将失真建模为亮度、对比度和结构三个不同因素的组合^[34]。因此，SSIM 在一定程度上可以更好地反映人类视觉系统的感受，用于评估图像或视频的失真程度。

关于鲁棒性，实验使用 NCC（Normalized Cross Correlation）作为评价标准。NCC，即归一化互相关系数，是一种用于评估两个信号相似度的度量方法。在图像和视频水印中，NCC 可以用于判断算法鲁棒性，其计算公式为：

$$NCC(\bar{x}, \bar{y}) = \frac{\sigma_{\bar{x}\bar{y}}}{\sqrt{\sigma_{\bar{x}} \cdot \sigma_{\bar{y}}}} \quad (3-15)$$

其中， \bar{x} 和 \bar{y} 表示原始水印信号和经攻击后提取的水印信号，通常为一个向量。 $\sigma_{\bar{x}}$ 和 $\sigma_{\bar{y}}$ 表示对应水印信号的方差， $\sigma_{\bar{x}\bar{y}}$ 表示 \bar{x} 和 \bar{y} 的协方差。NCC 的取值范围为[-1,1]，其中‘1’表示水印完全匹配，‘-1’表示水印完全不匹配。

在图像或视频处理中，当一个水印被添加到原始信号后，可以通过计算该信号经过攻击后提取的水印以及原始水印信号的 NCC 值，来评估算法的鲁棒性。如果经攻击后提取的 NCC 值较大，那么可以认为算法对该攻击具有较高的鲁棒性，即攻击给水印带来的影响很小；反之，如果 NCC 值很低，说明经攻击后，提取的水印较原始水印信号的差异较大，即算法对该攻击的鲁棒性较弱。

3.3.2 不可感知性分析

水印的不可感知性是指载体在嵌入水印后，水印对于人类视觉系统的可感知程度。如果水印在嵌入过程中对视频质量产生明显的影响，它将不符合实际应用的要求。因此，在数字水印技术中，保持水印的不可感知性具有一定重要意义。

为评估本章算法的不可感知性，实验设计计算测试视频序列在使用本章算法和文献[45]的算法嵌入水印后的 PSNR 值（dB）和 SSIM 值。对于给定的视频序列，其 PSNR 值和 SSIM 值由未嵌入水印的原始视频帧和嵌入水印后的视频帧共同计算得出。相关实验结果由表 3-1 所示。

由表 3-1 可得，根据本章算法嵌入水印的视频序列的 PSNR 值约为 37dB，而文献[45]中的 PSNR 值为 30dB。因此，本章算法较文献[45]中的算法的 PSNR 值平均高约 7 dB。同理也可以得出，本章算法的 SSIM 值较文献[45]中算法平均高 0.01。因此，由上述实验分析可得，在 PSNR 和 SSIM 方面，本章算法与文献[45]中的算法相比，具有更高的不可感知性，拥有更好的视觉效果。

表 3-1 含水印视频序列的 PSNR 值 (dB) 和 SSIM 值

视频序列	文献[45]		本章算法	
	PSNR	SSIM	PSNR	SSIM
Akiyo	32.472	0.982	36.233	0.993
Foreman	30.123	0.975	38.997	0.991
Mother and Daughter	29.455	0.983	37.235	0.992
Hall	28.272	0.978	37.126	0.992
Paris	30.158	0.986	36.583	0.991
Silent	29.409	0.980	36.993	0.990

3.3.3 鲁棒性分析

鲁棒性指算法或系统对于各种干扰、攻击或噪声的抵抗能力和稳定性。在图像、视频处理，以及信息安全领域中具有重要的意义。因此，在实际应用中，相关算法需要能够在面对各种干扰和攻击的情况下保持其性能和可靠性。本节主要讨论本章算法就缩放、旋转以及噪声攻击的鲁棒性，具体分析如下：

(1) 缩放攻击

为了验证算法在面对缩放攻击时的鲁棒性，实验设计了对使用本章算法嵌入水印的视频进行不同缩放系数的攻击测试，其中，使用的缩放系数为：150%、200%、300%以及 400%，并列出无攻击时的 NCC 值，用于分析算法在缩放攻击下的鲁棒性表现。此外，也将本章算法和文献[45]中算法进行对比，进一步评估算法在鲁棒性方面相对于现有算法的优势。

实验结果如表 3-2 所示，从中可以得出，无攻击时，本章算法的 NCC 值较文献[45]中算法的 NCC 值高 2%。在缩放攻击后，本章算法中的 NCC 值相对保持了与攻击前相同的数值。而文献[45]中，当缩放系数从 200%增加到 300%，以及 300%增加到 400%时，NCC 值出现显著下降的趋势。当系数为 400%时，文献[45]的 NCC 值比本章算法低 40%。综上所述，本章算法在上述缩放系数下，与文献[45]中算法相比，具有较高的鲁棒性，能够较好地抵抗缩放攻击。

表 3-2 在缩放攻击下的视频水印 NCC 值

攻击类型	文献[45]	本章算法
无攻击	0.980	0.993
缩放至 150%	0.856	0.987
缩放至 200%	0.927	0.993
缩放至 300%	0.694	0.993
缩放至 400%	0.610	0.987

(2) 旋转攻击

为了验证算法在面对旋转攻击时的鲁棒性,实验设计了使用不同旋转角度的攻击进行测试,并将本章算法和文献[45]中算法进行对比,其结果如表 3-3 所示。由实验结果可得,当旋转角度增加时,本章算法的 NCC 值可以保持不变,其平均值较无攻击时的 NCC 值略低。而文献[45]中的 NCC 值随着角度的增加而明显下降,特别是在 90° 到 120° 之间时,对应的 NCC 值比本章算法的 NCC 值低 60% 到 80%。综上所述,本章算法对旋转攻击的鲁棒性方面明显优于文献[45]中的算法,对于旋转攻击具有较高的鲁棒性。

表 3-3 在旋转攻击下的视频水印 NCC 值

攻击类型	文献[45]	本章算法
旋转 30°	0.604	0.942
旋转 60°	0.475	0.936
旋转 90°	0.405	0.975
旋转 120°	0.154	0.929

(3) 噪声攻击

为了验证算法在面对噪声攻击时的鲁棒性,实验设计了添加不同类型噪声进行测试。具体而言,实验设计对嵌入水印的视频添加了强度为 0.15 的椒盐噪声、方差为 0.04 的斑点噪声、均值为 0、方差为 0.02 的高斯噪声以及泊松噪声。并将本章算法和文献[45]中算法进行对比,其实验结果如表 3-4 所示。观察实验结果可知,在经过噪声攻击后,本章算法的 NCC 值较未攻击时略有下降,且基本稳定在 0.98,而由文献[45]算法求得的 NCC 值较本章方法低 14%,且对椒盐噪

声的鲁棒性最低。综上所述，本章算法在面对噪声攻击时表现出较好的鲁棒性。尽管在经历噪声攻击后，算法的 NCC 值会有所下降，但整体性能稳定，并且相对于文献[45]具有更高的鲁棒性。

表 3-4 在噪声攻击下的视频水印 NCC 值

攻击类型	文献[45]	本章算法
椒盐噪声	0.806	0.976
斑点噪声	0.874	0.980
高斯噪声	0.847	0.985
泊松噪声	0.928	0.988

3.4 本章小结

本章介绍了基于归一化 Zernike 矩的视频鲁棒水印算法。该算法采用了归一化 Zernike 矩来提取视频的特征，进而实现对水印的嵌入和提取。本章详细介绍了该算法的实现步骤，包括预处理、水印嵌入和提取的具体过程，以及归一化 Zernike 矩的计算方法等。实验表明，基于归一化 Zernike 矩的视频鲁棒水印算法具有较强的鲁棒性和抗攻击性，可以在视频经过高强度旋转攻击、缩放攻击、噪声攻击后仍然能够准确地提取出水印信息。此外，该算法能够在嵌入水印的同时保证视频的视觉质量，不会对视频的清晰度和色彩产生明显的影响，且与现有水印算法对比后可得，该算法在鲁棒性和水印的不可感知性方面都具有一定优势。

第四章 基于 Zernike 矩的快速视频鲁棒水印算法

本章主要研究适用于高清视频的基于 Zernike 矩的视频鲁棒水印技术，并提出了一种快速的视频鲁棒水印算法。本章首先介绍了算法的总体架构，并详细阐述了预处理、水印嵌入以及提取步骤。最后，通过实验验证了算法的有效性。

4.1 引言

根据第三章提出的，基于归一化 Zernike 矩的视频鲁棒水印算法可得，Zernike 矩适用于视频水印，并可以达到较高的抗旋转性和抗缩放效果。但是，由第二章对 Zernike 矩的理论介绍可知，Zernike 矩的计算消耗较大，且当分辨率为 256×256 时，若所求 Zernike 矩的最大阶数为 30，就已需要近 5 秒的时间来计算一帧中所有对应的 Zernike 矩值。因此，这类算法并不适用于高清视频，即分辨率较高的视频。所以这促使我们研究新的视频鲁棒水印方案，以在保持水印鲁棒性和不可感知性的同时，拥有较高的计算效率，适用于高清视频。

为了实现上述目标，本章提出了一种基于 Zernike 矩的快速视频鲁棒水印算法。为了适应高清视频，该算法采用了自适应分块法，将视频帧分为特定大小的块。对于每个块，进行 SVD (Singular Value Decomposition)，即奇异值分解，并提取最大奇异值。随后提取所有分块中最大的奇异值，并将其按顺序依次合并，以得到一个方形的特征矩阵，用于水印的嵌入和提取。由于该特征矩阵较原先视频帧相比，分辨率大幅降低，因此可以显著提高 Zernike 矩的计算效率。

此外，为了更好地平衡算法的鲁棒性和不可感知性，本算法在第三章算法的基础上，进一步筛选了适合的嵌入载体。且受到第二章视频编码技术框架的启发：视频编码主要传递冗余信息。因此在本算法中，在每一组 GOP 中选择适当的帧对，将水印信息作为冗余信息嵌入帧对之间，以增强水印对压缩攻击和转码攻击的鲁棒性。此外，在水印提取过程中，引入了两种可选的方法，即自适应尺寸调整和角度补偿，以进一步提高算法的适用范围。实验结果表明，该算法在面对常见的几何、噪声和压缩攻击时均有较高的鲁棒性。与最新的研究成果相比，该算法不仅具有较高的鲁棒性，还能保持良好的视觉效果和较高的计算效率。

4.2 算法设计

4.2.1 总体架构

本章所提算法的总体架构如图 4-1 所示。该研究方案主要包括三个步骤：视频序列预处理、水印信息的嵌入和水印信息的提取。为了增强算法对压缩和转码攻击的鲁棒性，在每一组 GOP 中选择合适的帧对作为关键帧，并将水印信息作为冗余信息嵌入帧对之间。这样可以增强水印对压缩攻击和转码攻击的鲁棒性。同时，为了适用于高清视频，提出了自适应分块法，并结合奇异值分解，在每个分块中选择最大的奇异值以降低分辨率。由于 Zernike 矩的数学特性，仅能表示对应矩阵内切圆区域的特征，而大多数视频的宽和高的大小并不一致，因此设计了自适应分块方法，根据视频分辨率的大小自适应选择块的大小，以得到方形的特征矩阵，便于计算 Zernike 矩，用于水印的嵌入和提取。

在水印信息的嵌入阶段，采用了与第三章一致的 DM-QIM 方法，将水印信息嵌入对应帧的 Zernike 矩的模值内，并进行一系列逆变换，得到包含水印信息的 YUV 视频序列。在水印信息的提取阶段，为了增强算法的应用范围，提出了两种可选的预处理方法：自适应尺寸调整和角度补偿，用于应对分辨率的特殊变化和高强度的旋转攻击。随后，重复嵌入水印时的预处理工作，并使用 DM-QIM 中的提取步骤来提取水印信息。

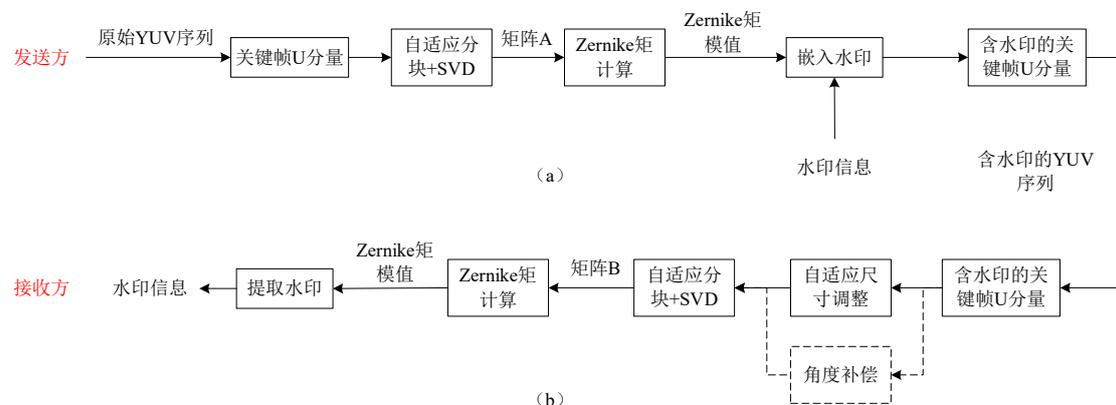


图 4-1 基于 Zernike 矩的快速视频鲁棒水印算法总体架构 (a) 水印嵌入 (b) 水印提取

4.2.2 预处理

如图 4-1 总体架构所示，算法的预处理部分主要包括选择关键帧、提取色度分量、自适应分块、奇异值分解以及 Zernike 矩计算部分。其详细步骤如下：

(1) 选择关键帧

选择关键帧包含两个步骤，即首先提取 YUV 视频序列的色度分量 U，再对视频帧进行分组，每个分组内包含相同数量的帧。其次，对于每个分组，计算每两个相邻帧之间的差异，以选定用于嵌入水印的帧，其数学表达式为：

$$D(i, i+1) = \sum_{x,y} |f(x, y, i) - f(x, y, i+1)|^2 \quad (4-1)$$

其中， $f(x, y, i)$ 和 $f(x, y, i+1)$ 代表一组分组中第 i 帧和第 $i+1$ 帧中，坐标为 (x, y) 的像素值。

在计算完毕组内相邻帧的差值后，选择差值最小的一组相邻帧，以构成一组帧对，用于后续水印嵌入步骤。

由上述步骤收集完毕所有的帧对后，在每一组帧对中，使用前一帧覆盖后一帧，并使用被前一帧覆盖的后一帧作为关键帧，用于嵌入水印。水印可以作为两帧之间的差值进行传递。两帧关系具体可以表述如下：令 I_i 和 I_{i+1} 为公式 (4-1) 中的第 i 帧和第 $i+1$ 帧。假设 (I_i, I_{i+1}) 是待嵌入水印的帧对，水印信号为 W_i 。嵌入水印后，帧对 (I_i, I_{i+1}) 将转换为 (I'_i, I'_{i+1}) ，其中 $I'_i = I_i$ ， I'_{i+1} 的值由 (I_i, W_i) 决定。

(2) 自适应分块

假设 (I_i, I_{i+1}) 代表需要嵌入水印的帧对，在将前一帧覆盖至后一帧后，可以得到 (I_i, I_{i+1}^*) ，其中 $I_{i+1}^* = I_{i+1}$ 。其次，对 I_{i+1}^* 进行自适应分块，将其分为互不相交的块，且在水平和垂直方向上，分块数量相同。假设 h 和 w 为 I_{i+1}^* 的高度和宽度。为了均衡水平方向和垂直方向上的块数、计算 h 和 w 的最大公约数 (GCD) 作为两个方向的块数。因此，一个自适应分块的尺寸可以表示为：

$$h_b = \frac{h}{\gcd(h, w)}, w_b = \frac{w}{\gcd(h, w)} \quad (4-2)$$

其中, h_b 表示分块的高度, w_b 为分块的宽度, $\text{gcd}(h, w)$ 代表 h 和 w 的最大公约数。

(3) 奇异值分解

对关键帧经过自适应分块后的每个分块执行奇异值分解, 其数学表达式为:

$$\mathbf{F} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (4-3)$$

其中, \mathbf{F} 代表特定分块, 大小为 $h_b \times w_b$; \mathbf{U} 为大小为 $h_b \times h_b$ 的正交矩阵; $\mathbf{\Sigma}$ 为大小为 $h_b \times w_b$ 的对角矩阵, 包含奇异值; \mathbf{V} 为大小为 $w_b \times w_b$ 的正交矩阵, 上标 T 表示转置操作。

由于奇异值中的最大值代表对应矩阵的大部分信息, 所以相对其他奇异值而言, 鲁棒性更强, 所以在该步骤中, 提取每个分块中奇异值的最大值作为该分块的特征值。收集完毕关键帧内所有的特征值后, 将其按顺序排列, 可以得到特征矩阵 \mathbf{A} , 其分辨率大小为 $\text{gcd}(h, w) \times \text{gcd}(h, w)$, 可用于后续水印嵌入。

4.2.3 水印的嵌入过程

在计算完毕可以 Zernike 矩后, 同第三章算法的嵌入过程一致, 总共分为两步, 一是需要删除一部分不适用于嵌入水印的矩, 以提高算法的适用性; 二是使用嵌入算法嵌入水印。具体如下:

(1) 筛选适用于嵌入水印的矩

本章的筛选步骤在第三章算法提及的筛选步骤的基础上, 引入了比例系数 α 以进一步筛选合适的矩, 其补充步骤如下:

步骤一: 假设按第三章所述的步骤, 删除不适用的矩后, 共有 L 个矩, 且由 $A_{n,m}$ 表示。将上述矩按下标 (n, m) 进行升序排序。其中, n 拥有较高的优先级。

步骤二: 由于低阶 Zernike 矩在对抗几何、噪声攻击时具有更好的表现^[19,66], 所以在该步骤中, 使用系数 α , 其取值范围为 $[0, 1]$, 来选择排序后的前 $\alpha \times L$ 个 Zernike 矩, 用于水印嵌入, 以最大限度地提高水印的鲁棒性和不可感知性。

(2) 嵌入步骤

筛选完毕适合嵌入水印的矩后, 执行水印的嵌入过程, 嵌入过程与第三章算

法类似，需要用到相似的公式。本章将对其简单描述，具体步骤如下：

步骤一：根据公式 (3-6)，将水印数据嵌入所有这些被选中的矩的振幅中，且每个目标矩阵内包含 1 比特的水印数据。

步骤二：使用公式 (3-8) 构建水印信号 \mathbf{R} 。

步骤三：使用系数 β ，将水印信号 \mathbf{R} 叠加至特征矩阵 \mathbf{A} 中，具体数学表达式与公式 (3-9) 相似，需要将其中 \mathbf{I} 转换为 \mathbf{A} ， α 转换为 β ，其余上标表示含义不变。 β 与 \mathbf{A}_r 可同理公式 (3-10) 以及公式 (3-11) 求得，分别表示强度系数和对原始特征矩阵的重构。

步骤四：对嵌入水印的特征矩阵 \mathbf{A}'' 根据上述步骤进行一系列逆变换，得到嵌入水印后的关键帧，并最终得到嵌入水印后的视频序列。

4.2.4 水印的提取过程

如图 4-1 (b) 中的水印提取过程所示，与嵌入过程不同的是，在预处理过程中，增加了自适应尺寸调整和角度补偿这两个可选的功能。

(1) 自适应尺寸调整

由于在数据嵌入前，需要将每个关键帧划分为小块。然而，如果其高度和宽度是互为素数，即最大公约数为 1，此时，就不能通过计算它们的最大公约数来获得适当的块大小。例如，如果原始视频的分辨率为 1080p (即 1920×1080)，在提取 U 通道分量后，对应分辨率将变为 540p (即 960×540)，其长和宽的最大公约数为 60。但是，如果该视频被缩放至宽度和高度互为素数时，如 1921×1081，就无法得到合适的分块尺寸。

因此，为解决这一难点，实施一个调整大小的策略，如表 4-1 所示。该表格以 1080p 视频为例，展示了不同缩放系数下，其 U 通道分量经过自适应分块后，每个分块的尺寸。当原始视频是 1080p，则如果缩放后对应视频序列的高度和宽度互为素数，需要将缩放后的视频尺寸调整至与该表格中宽度最接近的分辨率。为了尽可能多地适应实际场景，该功能缺省开启，如图 4-1 中以实线显示。同时，用户也可以根据特定需求，设计其他调整大小的策略。

表 4-1 1080p 视频的 U 通道分量在缩放后的自适应分块尺寸

缩放系数	分辨率		最大 公约数	分块尺寸	
	宽	高		宽	高
1	960	50	60	16	9
0.95	912	513	57	16	9
0.90	864	486	54	16	9
0.85	816	459	51	16	9
0.80	768	432	48	16	9
0.75	720	405	45	16	9
0.70	672	378	42	16	9
0.65	624	351	39	16	9
0.60	576	324	36	16	9
0.55	528	297	33	16	9
0.50	480	270	30	16	9

(2) 角度补偿

当视频经历高强度旋转攻击时，相应的像素的信息会发生很大变化，若此时再根据原始尺寸进一步裁剪图像，许多信息将会丢失。在这种情况下，就很难通过原来的方法提取水印。为解决这一问题，提出一种方法来补偿较大角度的旋转攻击，其计算简图总结于图 4-2，具体步骤描述如下：

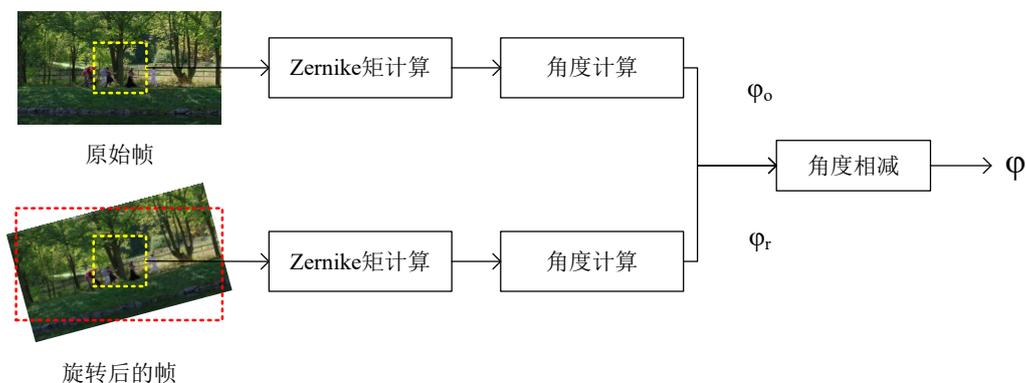


图 4-2 旋转角度计算简图

步骤一：将视频的中心点作为原点，并从原点提取一个边长为 50 的正方形。为了估计旋转角度，用公式 (2-4) 计算原始视频中被裁剪图像的 Zernike 矩。由于 Zernike 矩是由复数表示，其旋转角度可以通过复数的性质求出，其表达式为：

$$\varphi = \varphi_r - \varphi_o = \text{angle}(A_{n,m}^r) - \text{angle}(A_{n,m}^o) \quad (4-10)$$

其中，函数 $\text{angle}(\bullet)$ 用于计算输入矩的角度； r 和 o 分别用于上标和下标，用以代表旋转后的视频和原始视频。

步骤二：为了准确估计旋转角度并保持较低的计算成本，使用参数集 $S = \{(n,m) | (3,1), (5,1), (7,1), (4,2), (6,2), (8,2)\}$ 以计算 Zernike 矩，并将其取平均，得到最终的旋转角度估计值，其数学表达式为：

$$\varphi = \frac{1}{|S|} \sum_{(i,j) \in S} [\text{angle}(A_{i,j}^r) - \text{angle}(A_{i,j}^o)] \quad (4-11)$$

其中， S 代表参数集 S 中包含的元素数量。

选择集合 S 作为参数集的原因如下：其一，为了得到最准确的估计结果，需要选择相对稳定的 Zernike 矩，而根据 Zernike 矩的特性， n 的值越小，对应的 Zernike 矩会越稳定^[19]，所以需要选择较小的阶数用于参数集，以增加结果的准确性。其二，如 3.3.2 节中讨论，当 m 为 4 的倍数时，其中的 Zernike 矩无法准确计算，而当 $m = n = 0$ 或 1 时，相应的 Zernike 矩与图像无关，所以它们不适合作为参数使用，将不作考虑。其三，由于视频由一组帧序列构成，所以对于每一帧而言，需要较小的计算时间才能便于方法的应用。而根据 Zernike 矩的特性，当阶数 n 较小时，计算时间较低，所以需要选择阶数较小的 Zernike 矩。其四，为使角度估计的结果尽可能的准确，减小误差的同时，保证计算效率，有必要选择适当数量的参数对，以在准确率和计算效率中取得均衡。

综上所述，由于一和三结论一致，所以选择适当的参数对需要满足三个条件：首先， n 的值需要较低；其次， m 不是 4 的倍数，而且 m 和 n 都不是 0 或 1；第三，应该使用适当数量的参数对。所以在该方法中，选择集合 S 作为参数集。

步骤三：在获得补偿角度 φ 后，为了最小化误差，补充使用 $\varphi+1$ 和 $\varphi-1$ 这两个角度提取水印，并选择三者中，获得最高 NCC 值的角度作为补偿角度。

(3) 水印提取

同第三章类似，当完成预处理工作后，提取水印共需三个主要步骤。假设预处理后获得特征矩阵 \mathbf{B} 用于提取水印，提取步骤具体如下所示。

步骤一：同嵌入过程所述，在计算完毕对应矩阵的 Zernike 矩后，首先需要删除不适用于嵌入水印的矩，具体删除标准同嵌入过程一致。

步骤二：对筛选后的 Zernike 矩取模值，作为载体以提取水印。假设 $|B_{n,m}|$ 表示嵌入水印的 Zernike 矩的模值， b 代表提取的水印位，可由公式 (3-12) 求得。 $|B_{n,m}^{w_b}|$ 可由公式 (3-13) 计算得出，其中， Δ 代表量化步长，需要与嵌入步骤使用相同的参数值。

步骤三：在提取完毕该目标帧内所有选定矩中嵌入的水印后，为了减少提取结果存在突变的可能，将所有矩中选择频率最高的一个作为每个目标帧的最终提取水印位，以减少误码率，提高提取结果准确性。

步骤四：为了获得完整的水印信息，以上述步骤对所有嵌入水印的关键帧进行重复操作，并将提取的水印结果按顺序组合，从而得到完整的嵌入水印信息。

4.3 实验结果与分析

4.3.1 实验设置及评价标准

本文中的所有实验都是在 Matlab R2016a 的环境中实现的，计算机系统为 64 位 Windows 10，内存 8GB，且使用 2.3GHz 英特尔酷睿 i5 处理器。为尽可能客观地评估该算法，使用九个分辨率为 1920×1080 的高清 (HD) 标准视频序列作为实验数据集，即 Ducks Take Off, In To Tree, Life, Old Town Cross, Park Joy, Pedestrian Area, Rush Hour, Station 以及 Sunflower。每个测试视频包含 300 帧，可以通过网址 <http://trace.eas.edu/yuv/index.html> 获取。

本章算法共有三个参数需要设置，分别是量化步长 Δ ，Zernike 矩的最大阶数 N ，以及比例系数 α 。为了找到近乎最优的参数，对带有水印的视频应用不同的攻击方法进行测试，并计算提取水印与原始水印之间的归一化互相关系数 (NCC)。测试中使用了三种常见的攻击方式，分别是缩放、旋转和椒盐噪声攻击，

且就第一段所述的九个标准序列中，随机选择了四个视频序列用于模拟测试，序列名称分别是 In To Tree、Rush Hour、Old Town Cross 和 Sunflower。

为确定最优量化步长 Δ ，设置 $N = 30$ ， $\alpha = 1/3$ ，并枚举水印嵌入的参数 Δ 从 10000 到 50000，步长为 5000。对于每个带水印的视频，分别进行系数为 0.4 的缩放攻击、旋转角度为 5 度的旋转攻击和强度为 0.2 的椒盐噪声攻击。实验结果如图 4-3 (a)，(b)，(c) 所示，从中可以得出，当 $\Delta=30000$ 时，四个测试视频的相应 NCC 值能均衡地达到较高的数值。

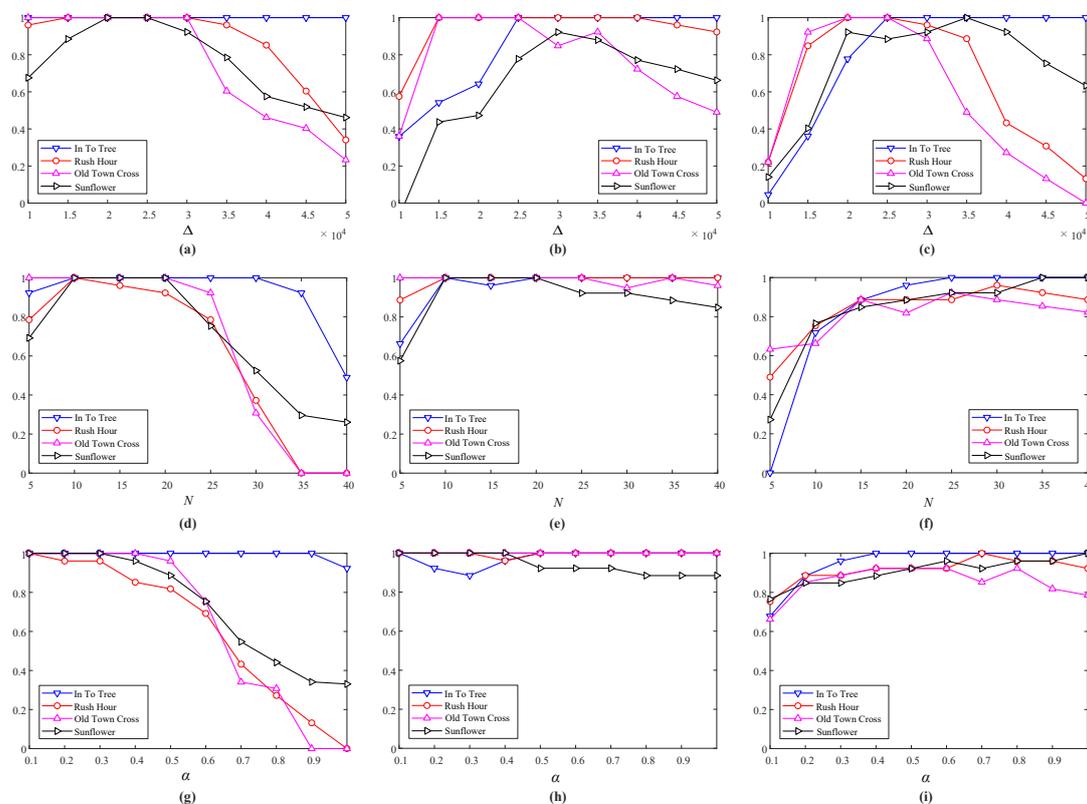


图 4-3 在不同攻击下变化不同参数后的 NCC 值

同样地，为确定最优阶数 N ，由于上一步中可以确定， $\Delta=30000$ 时能平衡测试视频的鲁棒性，故将 Δ 设置为 30000， α 设置为 1/3，阶数 N 从 5 枚举到 40，步长为 5，进行水印嵌入。对应结果如图 4-3 (d)，(e)，(f) 所示，从中可以得出，当 $N = 20$ 时能得到四个测试序列的最优解，故取 20 作为最优阶数。

最后，设置最佳比例系数 α ，取 $\Delta=30000$ ， $N = 20$ ，枚举 α 从 0.1 到 1.0，步长为 0.1 进行水印嵌入，实验结果如图 4-3 (g)，(h)，(i) 所示，从中可以得

出, $\alpha = 1/2$ 时可实现较高鲁棒性。综上所述, 为了达到较高的鲁棒性, 在水印嵌入和提取中, 推荐使用 $\Delta=30000$, $N = 20$ 和 $\alpha = 1/2$ 作为参数设置。在后续实验中, 若未有特殊说明, 这些参数将缺省使用。

关于不可感知性, 实验中使用 PSNR、SSIM 和 VQM 三种客观评价标准。其中, PSNR 和 SSIM 在第三章中已经详细描述。VQM^[35], 即视频质量度量 (Video Quality Metric), 是一种用于评估视频质量的算法。它可以测量视频的失真和质量损失程度, 并将其转化为一个数值, 以便进行比较和分析。VQM 算法的设计目的在于模拟人类对视频质量的主观感受, 同时还考虑了视频的特征、场景等因素, 使其更为准确。计算 VQM 时需要输入原始视频和嵌入水印的视频, 其取值范围为[0,1], 数值越低代表失真越小, 即水印的不可感知性越强。关于鲁棒性, 实验使用 NCC 作为评价标准, 已第三章详细描述。

4.3.2 不可感知性分析

由于在实际应用场景中, 该算法主要可用于版权保护, 所以要求水印嵌入引起的视觉失真不应被察觉, 否则, 嵌入水印的视频的商业价值将会相应受损。因此, 水印需要有较高的不可感知性, 以得到较高的应用价值。

为测试水印的不可感知性, 实验计算相关视频序列的 PSNR、SSIM、VQM 值, 对其不可感知性进行评估, 实验结果如表 4-2 所示。其中, 嵌入水印视频的平均 PSNR 约为 47.5dB, 平均 SSIM 高于 0.99, 平均 VQM 约为 0.02。由于 PSNR 和 SSIM 越高, VQM 越低, 代表水印的不可感知性越强。所以在上述评价标准下, 可以得出结论: 该算法具有优秀的不可感知性。

4.3.3 鲁棒性分析

本节主要研究算法对旋转、缩放、噪声、压缩及转码攻击的鲁棒性。由本章算法整体架构可得, 在水印提取过程中, 角度补偿功能并非缺省开启, 需要使用者设置是否启用, 故本节将对鲁棒性的分析实验分为两部分, 分别为“不使用角度补偿”和“使用角度补偿”两种情况。在使用角度补偿时, 主要研究其对不同角度旋转攻击的鲁棒性。实验结果如下所示:

表 4-2 嵌入水印后不同视频序列的平均 PSNR、SSIM 和 VQM

视频序列	PSNR	SSIM	VQM
Ducks Take Off	47.7343	0.99945	0.0194
In To Tree	47.4308	0.99908	0.0206
Life	47.2791	0.99889	0.0208
Pedestrian Area	47.6556	0.99914	0.0220
Old Town Cross	47.7225	0.99927	0.0203
Park Joy	47.4437	0.99926	0.0196
Station	47.6440	0.99920	0.0204
Rush Hour	47.5341	0.99913	0.0214
Sunflower	47.4863	0.99904	0.0199
均值	47.5478	0.99920	0.0208

(1) 不使用角度补偿

不使用角度补偿时，主要研究算法对缩放攻击、噪声攻击、旋转攻击、旋转和裁剪攻击、压缩和转码攻击时的鲁棒性。相关实验结果如表 4-3 所示，具体分析如下：

a) 缩放攻击

为测试算法对缩放攻击的鲁棒性，实验应用三个缩放系数，即 90%，60%，30%进行测试。在模拟缩放攻击时，仅考虑使用缩放系数的情况，这是由于在默认情况下，该算法应用自适应尺寸调整方法，因此经自适应调整大小后的输入视频大小与原始视频大小有固定的比例关系，所以，仅使用不同缩放系数模拟缩放攻击，即可包含大部分缩放场景。

实验结果如表 4-3 所示，当将测试视频的分辨率缩放至视频原尺寸的 90% 或 60%时，NCC 值为 1.000，当将其缩小至 30%时，其 NCC 值略微下降至 0.942。这表明本算法对缩小攻击具有很高的鲁棒性。

b) 噪声攻击

为测试算法对噪声攻击的鲁棒性，实验中应用三种不同类型的噪声，即椒盐噪声、斑点噪声和高斯噪声进行测试。对噪声的参数设置如下：对于椒盐噪声，

强度设置为 0.1。对于斑点噪声，使用 0.04 作为方差，对于高斯噪声，均值和方差分别设置为 0 和 0.02。由表 4-3 可得，在上述噪声攻击下，所得的 NCC 值均为 1，证明了该算法对噪声攻击具有较高的鲁棒性。

c) 旋转攻击

为测试算法对旋转攻击的鲁棒性，实验分别设计将嵌入水印的视频旋转 90 度、180 度和 270 度，并将镜像翻转视为一种特殊情况，考虑在旋转攻击内。实验结果如表 4-3 所示，从中可以得出，在镜像翻转和上述不同角度旋转攻击之后，所有的 NCC 值都为 1，这表明所提出的算法具有很好的抵抗旋转攻击的能力。

表 4-3 不使用角度补偿时的平均 NCC 值

攻击类型	NCC	攻击类型	NCC
缩放至 90%	1.000	旋转 270°	1.000
缩放至 60%	1.000	旋转 3°并裁剪	1.000
缩放至 30%	0.942	旋转 5°并裁剪	0.943
椒盐噪声	1.000	以 QP=28 压缩	1.000
斑点噪声	1.000	转码至 QP=15	1.000
高斯噪声	1.000	转码至 QP=20	1.000
镜像翻转	1.000	转码至 QP=25	0.987
旋转 90°	1.000	转码至 QP=30	1.000
旋转 180°	1.000	转码至 QP=32	0.946

d) 旋转和裁剪攻击

为测试算法对于旋转和裁剪攻击的鲁棒性，实验中使用两个角度，即 3 度和 5 度进行测试。其中，旋转和裁剪攻击定义为：将带水印的视频旋转一定角度并进行裁剪以保持视频的尺寸。实验结果如表 4-3 所示，从中可得，当使用 3 度进行旋转和裁剪攻击时，NCC 值为 1，在 5 度的情况下略微下降到 0.943，但仍然可以表明算法具有较高的鲁棒性。

e) 压缩和转码攻击

为测试算法对于压缩和转码攻击的鲁棒性，实验设计使用 H.264/AVC 编码器对嵌入水印的视频进行压缩。其中，将 Profile 配置为 Baseline，量化参数 QP

(Quantization Parameter) 设置为 28, 并使用不同的 QP 进行同质转码, 即保留最初的编码方式: 若最初使用 H.264 编码器进行编码, 则转码后仍输出 H.264 码流。相关实验结果如表 4-3 所示, 从中可得, 在经历压缩和转码攻击后, 大多数 NCC 值为 1, 其余 NCC 值也均高于 0.94。因此, 该算法在抵抗压缩和转码攻击方面的有效性也得到了验证。

(2) 使用角度补偿

由于自适应分块法的限制, 当旋转角度不是 90 的整数倍的情况时, 无法获得令人满意的提取结果。因此, 为了涵盖所有可能的情况, 该算法在提取步骤中应用了角度补偿。相关实验结果如表 4-4 所示, 从中可以得出, 在应用角度补偿后, 嵌入水印的视频在旋转不同角度后的 NCC 值。此外, 由表 4-4 可得, 经过不同角度的旋转攻击后, 对应的 NCC 值通常高于 0.96, 这表明在角度补偿后, 该算法对不同类型的旋转攻击都有较高的鲁棒性。

表 4-4 使用角度补偿后的平均 NCC 值

攻击类型	NCC	攻击类型	NCC
旋转 10°	1.000	旋转 190°	0.987
旋转 15°	0.991	旋转 195°	0.987
旋转 30°	1.000	旋转 210°	0.978
旋转 45°	0.996	旋转 225°	0.965
旋转 75°	0.960	旋转 255°	0.974

4.3.4 对比分析

本节将从不可感知性、鲁棒性计算复杂度和三个方面对本章提出的算法与相关工作, 即文献[13]、[45]、[58]、[74]、[75]进行对比。其中文献[13]、[45]均基于 DT CWT 域的视频鲁棒水印算法, 文献[58]和[75]为基于 Zernike 矩的视频水印算法, 文献[74]为基于 DWT 的数字水印算法。

为了公平比较, 实验中将所有算法的 GOP 设置为 6, 并使用各自论文中推荐的参数进行仿真。由于文献[13]中推出了多种算法, 在后续实验中选择其中的算法二进行对比实验。此外, 在本节中, 不对本章算法应用角度补偿。

(1) 不可感知性

为测试不可感知性, 实验使用 PSNR、SSIM 以及 VQM 作为评价标准, 相关实验结果如表 4-5 所示, 从中可以观察到, 虽然该算法的平均 PSNR 略低于文献 [74] 中提出的算法, 但明显高于其他对比算法。此外, 该算法在平均 SSIM 方面表现出最高的值, 证明其具有较高的不可感知性能。同时, 该算法在 VQM 方面也呈现最低值, 这意味着在视觉质量方面具有出色的表现。因此, 本章算法和其他对比算法相比, 具有较高的不可感知性。

表 4-5 不同算法下含水印视频的平均 PSNR、SSIM 和 VQM

评价标准	文献[13]	文献[45]	文献[74]	文献[75]	本章算法
平均 PSNR	39.546	41.599	49.573	35.363	47.548
平均 SSIM	0.987	0.990	0.994	0.982	0.999
平均 VQM	0.048	0.037	0.030	0.060	0.021

(2) 鲁棒性

该小节主要讨论本章算法的鲁棒性较其他算法的对比, 分析各自对于缩放攻击、旋转攻击、压缩转码攻击, 以及噪声攻击的鲁棒性。具体实验分析结果如下:

a) 缩放攻击

为测试本章算法较现有算法相比, 对缩放攻击的鲁棒性, 实验设计对使用不同算法嵌入水印的视频应用缩放系数, 将视频按比例缩放, 以模拟缩放攻击。实验主要测试了应用三个缩放系数时的鲁棒性, 即 80%, 60% 和 40%, 分别表示将原始视频缩放至相应的比例的分辨率大小。实验结果如表 4-6 所示, 从中可知, 对于缩放攻击, 本章算法的平均 NCC 值显著高于其他算法, 且当缩放系数越小时, 效果越明显。这表明本章提出的算法在抵抗缩放攻击方面较具有较大的优势。

表 4-6 不同算法在应用缩放攻击后的平均 NCC 值

攻击类型	文献[13]	文献[45]	文献[74]	文献[75]	本章算法
无攻击	0.986	0.921	0.970	0.964	1.000
缩放至 80%	0.943	0.911	0.968	0.953	1.000
缩放至 60%	0.936	0.955	0.972	0.928	1.000
缩放至 40%	0.896	0.982	0.970	0.847	0.991

b) 旋转攻击

为测试对本章算法较现有算法相比,对旋转攻击的鲁棒性,实验设计对每个使用不同算法嵌入水印的视频进行旋转、镜像翻转以及旋转和裁剪攻击。对比结果如表 4-7 所示,从中可以看出,在前四种情况下,本章算法的 NCC 值较高,即具有较高的鲁棒性,而在最后一种情况时,NCC 略有下降,但仍高于其他算法的结果。对于文献[74]和[75]中的算法,其 NCC 均值比本章算法的对应均值低约 4%。此外,在文献[13]和[45]的算法中,当受到 90 度旋转攻击时,水印无法提取,即检测失败。同时,在其他四种旋转攻击的情况下,本章算法的 NCC 值显著高于文献[13]和[45]的算法的 NCC 值。综上所述,与上述四种算法相比,本章算法在旋转攻击方面的具有较好的表现,且能同时对旋转和裁剪攻击具有鲁棒性。

表 4-7 不同算法在应用旋转攻击后的平均 NCC 值

攻击类型	文献[13]	文献[45]	文献[74]	文献[75]	本章算法
旋转 90°	-	-	0.957	0.928	1.000
旋转 180°	0.817	0.833	0.943	0.967	1.000
镜像翻转	0.843	0.867	0.957	0.933	1.000
旋转 3°并裁剪	0.828	0.854	0.972	0.964	1.000
旋转 5°并裁剪	0.746	0.809	0.938	0.936	0.943
均值	0.809	0.841	0.953	0.946	0.986

c) 压缩和转码攻击

为测试对本章算法较现有算法相比,对于压缩攻击的鲁棒性,实验使用 H.264/AVC 编码器模拟压缩攻击,且将 Profile 配置为 Baseline,QP 设置为 28,帧率设置为在每秒 25 帧。对于转码攻击的鲁棒性,实验设计将原始水印视频的 H.264/AVC 视频流(QP = 28,每秒 25 帧)转换为具有不同 QP 和分辨率的 H.264/AVC 视频流。其中,720p 表示分辨率为 1280×720 的视频,540p 表示分辨率为 960×540 的视频。具体实验结果如表 4-8 所示,从中可得,虽然在文献[45]所述的算法中,转码至 540p 时所得的 NCC 值略高于本章算法所得的 NCC 值,但本章算法在其他情况下比所有算法都有更好的表现,即在抵抗压缩和转码攻击时具有较高的优越性。

表 4-8 应用压缩和转码攻击后不同算法的 NCC 值

攻击类型	文献[13]	文献[45]	文献[74]	文献[75]	本章算法
压缩攻击	0.912	0.927	0.967	0.946	1.000
转码至 720p	0.856	0.887	0.972	0.928	0.996
转码至 540p	0.828	0.986	0.967	0.894	0.983
转码至 QP=25	0.906	0.927	0.967	0.912	0.987
转码至 QP=30	0.894	0.925	0.967	0.882	1.000

d) 噪声攻击

为测试算法较现有算法相比,对噪声攻击的鲁棒性,实验分别设计在使用不同算法嵌入水印的视频添加强度为 0.15 的椒盐噪声、方差为 0.04 的斑点噪声、均值为 0,方差为 0.02 的高斯噪声和泊松噪声,并从这些视频中提取 NCC 值。实验结果如表 4-9 所示,从中可以看出,本章算法在上述所有情况下都优于其他对比算法,且平均 NCC 值显著高于文献[13]和[45]的算法。

表 4-9 应用噪声攻击后的不同算法的 NCC 值

攻击类型	文献[13]	文献[45]	文献[74]	文献[75]	本章算法
椒盐噪声	0.547	0.629	0.957	0.914	0.979
斑点噪声	0.754	0.837	0.977	0.948	1.000
高斯噪声	0.729	0.795	0.972	0.927	1.000
泊松噪声	0.853	0.901	0.977	0.948	1.000

(3) 计算效率

为进一步比较了所提出的算法与其他相关算法的计算效率,分别列出了不同算法每个 1080p 帧的平均执行时间(以秒为单位),其中包括水印嵌入和水印提取的时间,如表 4-10 所示。与文献[74]和[75]相比,所提出的算法需要明显较低的执行时间。虽然在嵌入时间方面,本章算法与文献[45]所提算法消耗时间接近,但在提取水印时,本章算法具有更低的执行时间。此外,文献[13]中的算法和本章算法的计算效率处于相同水平。此外,也测试了文献[58]中算法的计算效率,由于该算法使用 Zernike 矩,但并未设计提高效率的步骤,所以具有非常高的计算复杂度而不能应用于高分辨率视频,所以并未将结果展示于表 4-10。综上所述,

与现有算法相比，所提出的算法在计算效率上具有一定的优势。

表 4-10 不同算法的执行时间对比(秒)

算法	嵌入步骤	提取步骤
文献[13]	0.984	0.476
文献[45]	0.860	2.376
文献[74]	11.052	2.778
文献[75]	6.724	4.284
本章算法	0.899	0.296

4.4 本章小结

本章主要介绍了一种基于 Zernike 矩的快速视频鲁棒水印算法，适用于高清视频。本章首先介绍了算法架构，包括预处理、以及水印的嵌入和提取模块，并详细阐述了每个模块的具体实现方法。其中自适应分块法是使该算法能够应用于高清视频的主要原因。其次，设计实验，评估水印的不可感知性、鲁棒性以及计算效率，并与相关算法进行了比较分析。

实验结果表明，使用该算法嵌入水印后，视频水印具有较好的不可感知性，且对于多种攻击都具有较高的鲁棒性，并且与现有算法相比，具有一定优势，且有更好的性能表现。综上所述，该算法能够在不影响视频质量的前提下有效地抵抗多种攻击，具有更高的不可感知性和鲁棒性，在计算效率上也具有一定优势，具有一定应用价值，可以在数字版权保护和信息安全等方面得到一定应用。

第五章 总结与展望

5.1 总结

作为信息安全研究领域中的一大分支,数字水印在近些年来的研究成果层出不穷。且由于近几年视频信号的广泛传播和使用,视频水印技术也得到了越来越广泛的关注和研究。其中,针对高强度旋转攻击,仍是当前视频水印研究的难点问题。*Zernike* 矩作为一种广泛应用于数字图像处理的工具,具有较强的抗旋转特性,但由于其较低的计算效率,并未广泛应用于视频水印。因此本文致力于研究基于 *Zernike* 矩的视频鲁棒水印,取得的主要研究成果如下:

(1) 本文在第三章提出了一种基于归一化 *Zernike* 矩的视频鲁棒水印算法,用于解决抵抗高强度旋转攻击的问题。该算法利用了 *Zernike* 矩的旋转不变性来嵌入水印信息,并且由于其抗噪声能力,也能很好地抵抗噪声攻击。为了保证水印的不可感知性,该算法采用了自适应矩选择策略,并考虑了人类视觉系统的特点,采用色度通道嵌入水印,以进一步减少对原始视频视觉上的影响。实验结果表明,相比于现有算法,该算法平均峰值信噪比提高了约 7dB,具有更高的不可感知性,且在缩放、噪声以及高强度旋转攻击下也有较好的表现。由于该算法对高强度旋转攻击有较强的鲁棒性,在版权保护等领域有一定的应用价值。

(2) 本文在第四章提出了一种基于 *Zernike* 矩的快速视频鲁棒水印算法,用于解决 *Zernike* 矩的计算量过大,不适用于高清视频的问题。该算法中,应用自适应分块法,将每个帧分成不相交的块,并对每个块应用奇异值分解,最后获得一个方形特征矩阵,以大幅降低特征矩阵的分辨率,从而减少计算量。为增强抵抗压缩和转码攻击的能力,该算法根据视频编码框架的特性,应用补偿机制,将帧对间的帧差覆盖为水印。与传统的 *Zernike* 矩方法相比,该算法能够在保证水印鲁棒性的同时,大大减少计算时间。此外,经过大量对比实验后可以得出,该算法能够在不影响视频质量的前提下有效地抵抗多种攻击,包括旋转、缩放、噪声、压缩和转码等攻击,具有更高的不可感知性和鲁棒性。综上所述,该算法提供了一种快速且鲁棒的视频水印方案,可以较好地保护视频的版权信息。

5.2 展望

本文以基于 Zernike 矩的视频鲁棒水印算法为研究目标，在已有的水印理论的基础上，提出了两种算法，分别详细介绍与第三和第四章。实验结果表明了所提出的算法的有效性，但是这些算法依然有可以改进的地方。因此，对本文的研究提出以下几点展望：

(1) 对于第三章提出的基于归一化 Zernike 矩的视频鲁棒水印算法。在优点方面，具有较强的鲁棒性和抗攻击性，可以在视频经过一定程度的攻击后仍然能够提取出水印信息。另外，该算法能够在嵌入水印的同时保证视频的视觉质量，不会对视频的清晰度和色彩产生明显的影响。在不足之处方面，该算法的计算量较大，需要耗费一定的时间和计算资源，同时水印嵌入的容量也存在一定的限制。

(2) 对于第四章提出的基于 Zernike 矩的快速视频鲁棒水印算法。在优点方面，首先，该算法利用 Zernike 矩具有的旋转和尺度不变性，能够有效抵御旋转和缩放攻击。另外，算法使用了自适应分块算法，使得嵌入和提取过程简单高效，能够在较短的时间内完成水印的嵌入和提取。此外，实验结果显示，该算法对于噪声和压缩攻击也具有较好的鲁棒性。最后，与现有算法相比，该算法在鲁棒性和提取质量等方面表现优越。然而，该算法仅适用于高分辨率的视频，不适用于低分辨率视频，且同样有嵌入容量受限的限制。

综上所述，基于归一化 Zernike 矩的视频鲁棒水印算法是一种有效的水印技术，具有良好的鲁棒性和抗攻击性。未来，可以进一步优化该算法的计算效率，提高水印嵌入的容量，以适应更加广泛的应用场景。同时，关于基于 Zernike 矩的快速视频鲁棒水印算法，可以研究如何将其扩展应用至低分辨率视频，且增加水印嵌入的容量。此外，这两种算法尚未考虑时域攻击，帧同步、帧删除等攻击的鲁棒性，将在未来进一步的研究。

参考文献

- [1] QASIM A F, MEZIANE F, ASPIN R. Digital watermarking: Applicability for developing trust in medical imaging workflows state of the art review [J]. *Computer Science Review*, 2018, 27:45-60.
- [2] KOCH J. D, SMITH M D, TELANG R. *Camcording and Film Piracy in Asia-Pacific Economic Cooperation Economies* [R]. USA: International Intellectual Property Institute, 2011.
- [3] ASIKUZZAMAN M, PICKERING M. An Overview of Digital Video Watermarking [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(9):2131-2153.
- [4] ANAND A, SINGH A K. Watermarking techniques for medical data authentication: A survey [J]. *Multimedia Tools and Applications*, 2021, 80(20):30165-30197.
- [5] 罗茂. 基于仿射矩阵校正的抗几何攻击水印算法 [D]. 云南大学, 2019.
- [6] XIAO Y, CHENG W, XIAO Z. A survey on robust video watermarking algorithms for copyright protection [J]. *Applied Sciences*, 2018, 8(10):1-26.
- [7] LIU Q, YANG S, LIU J, et al. A discrete wavelet transform and singular value decomposition-based digital video watermark method [J]. *Applied Mathematical Modelling*, 2020, 85:273-293.
- [8] LIU W, LIU X, ARGYRIOU A, et al. 360-degree VR video watermarking based on spherical wavelet transform [J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2021, 17(1):38.
- [9] CEDILLO A, CEDILLO M, MIYATAKE M N, et al. A spatiotemporal saliency-modulated JND profile applied to video watermarking [J]. *Journal of Visual Communication and Image Representation*, 2018, 52:106-117.
- [10] LIU X, ZHAO R, LI F, et al. Novel robust zero-watermarking scheme for digital rights management of 3D Videos [J]. *Signal Processing: Image Communication*, 2017, 54:140-151.

- [11] KIMPAN S, LASAKUL A, CHITWONG S. Variable block size based adaptive watermarking in spatial domain [C]. Proceedings of the IEEE International Symposium on Communications and Information Technologies, 2004: 374-377.
- [12] 韩涛, 祝跃飞. 基于 Canny 边缘检测的自适应空域隐写术 [J]. 电子与信息学报, 2015, 37(5):1266-1270.
- [13] ASIKUZZAMAN M, ALAM M, LAMBERT A, et al. Imperceptible and robust blind video watermarking using chrominance embedding: a set of approaches in the DT-CWT domain [J]. IEEE Transactions on Information Forensics and Security, 2014, 9(9):1502-1517.
- [14] YANG L, WANG H, ZHANG Y, et al. A robust DCT-based video watermarking scheme against recompression and synchronization attacks [C]. Proceedings of the International Workshop on Digital-forensics and Watermarking, 2022: 149-162.
- [15] WANG Y L, PEARMAIN A. Blind MPEG-2 video watermarking robust against geometric attacks: a set of approaches in DCT domain [J]. IEEE Transactions on Image Processing, 2006, 15(6):1536-1543.
- [16] MAREEN H, DE J, VAN G, et al. A scalable architecture for uncompressed-domain watermarked videos [J]. IEEE Transactions on Information Forensics and Security, 2019, 14(6):1432-1444.
- [17] MADINE F, AKHAEI A, ZARMEHI N. A multiplicative video watermarking robust to H.264/AVC compression standard [J]. Signal Processing: Image Communication, 2018, 68:229-240.
- [18] MANSOURI A, MAHMOUDI A. Toward a secure video watermarking in compressed domain [J]. Journal of Information Security and Applications, 2019, 48:102370.
- [19] KIM H, LEE H. Invariant image watermark using Zernike moments [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(8):766-775.

- [20] LUTOVAC B, DAKOVIC M, STANKOVIC S, et al. An algorithm for robust image watermarking based on the DCT and Zernike moments [J]. *Multimedia Tools and Applications*, 2017, 76(22):23333-23352.
- [21] HU R, XIANG S. Cover-lossless robust image watermarking against geometric deformations [J]. *IEEE Transactions on Image Processing*, 2021, 30:318-331.
- [22] SINGHAL N, LEE Y, KIM C, et al. Robust image watermarking using local Zernike moments [J]. *Journal of Visual Communication and Image Representation*, 2009, 20(6):408-419.
- [23] WANG H, CHEN Y, ZHAO T. Modified Zernike moments and its application in geometrically resilient image zero-watermarking [J]. *Circuits, Systems, and Signal Processing*, 2022, 41(21):6844-6861.
- [24] LI S, ZHANG X. Toward construction-based data hiding: from secrets to fingerprint images [J]. *IEEE Transactions on Image Processing*, 2019, 28(3):1482-1497.
- [25] XIANG Y, HUANG J, GONZALEZ F, et al. Latest advances and emerging applications of data hiding [J]. *IEEE Access*, 2016, 4:9740-9742.
- [26] TAO J, LI S, ZHANG X, et al. Towards robust image steganography [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(2):594-600.
- [27] HERODOTUS J. *The histories* [M]. London: Penguin Books, 1996.
- [28] LAREW K, KAHN D. The codebreakers: the story of secret writing [J]. *The American Historical Review*, 1968, 74(2):537-538.
- [29] THANKI R, BORISAGAR K. A technical review of digital image watermarking techniques [J]. *International Journal of Advanced Research on Computer Science Software Engineering*, 2013, 3(5):1290-1299.
- [30] WANG Y, ZHOU Y, QIAN Z, et al. Review of robust video watermarking [J]. *Journal of Image and Graphics*, 2022, 27(1):27-42.
- [31] HU Y, LU W, MA M, et al. A semi fragile watermarking algorithm based on

- compressed sensing applied for audio tampering detection and recovery [J]. *Multimedia Tools and Applications*, 2022, 81(13):17729-17746.
- [32] 秦莉文. 鲁棒数字水印性能优化方法研究 [D].北京交通大学,2021.
- [33] KHOSRAVI M, YAZDI M. A lossless data hiding scheme for medical images using a hybrid solution based on IBRW error histogram computation and quartered interpolation with greedy weights [J]. *Neural Computing and Applications*, 2018, 30(7):2017-2028.
- [34] WANG Z, BOVIK A, SHEIKH H, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4):600-612.
- [35] PINSON M.H., WOLF S. A new standardized method for objectively measuring video quality [J]. *IEEE Transactions on Broadcasting*, 2004, 50(3):312-322.
- [36] TIRKEL A Z, RANKIN G A, VAN S R M, et al. Electronic watermark [J]. *Digital Image Computing, Technology and Applications*, 1993, 1:666-673.
- [37] VAN G, TIRKEL Z, OSBORNE F. A digital watermark [C]. *Proceedings of the 1st International Conference on Image Processing*, 1994: 86-90.
- [38] OH T, LEE M, KIM K, et al. Robust high- definition video watermarking based on self- synchronizing signals against composite distortions [J]. *Optical Engineering*, 2010, 49(49):662.
- [39] WIEGAND T, SULLIVAN G, BJONTEGAARD G, et al. Overview of the H.264/AVC video coding standard [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2003, 13(7): 560-576.
- [40] TEW Y, WONG K. An overview of information hiding in H.264/AVC compressed video [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, 24(2):305-319.
- [41] LEE M, IM D, LEE H, et al. Real-time video watermarking system on the compressed domain for high-definition video contents: Practical issues [J]. *Digital Signal Processing*, 2012, 22(1):190-198.

- [42] LAI C, TSAI C. Digital image watermarking using discrete wavelet transform and singular value decomposition [J]. *IEEE Transactions on Instrument and Measurement*, 2010, 59(11):3060-3063.
- [43] WANG Y, PEARMAIN A. Blind MPEG-2 video watermarking robust against geometric attacks: A set of approaches in DCT domain [J]. *IEEE Transactions on Image Processing*, 2006, 15(6):1536-1543.
- [44] ASIKUZZAMAN M, ALAM M, LAMBERT A, et al. Robust DT CWT-based DIBR 3D video watermarking using chrominance embedding [J]. *IEEE Transactions on Multimedia*, 2016, 18(9):1733-1748.
- [45] HUAN W, LI S, QIAN Z, et al.: Exploring stable coefficients on joint sub-bands for robust video watermarking in DT CWT domain [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 32(4):1955-1965.
- [46] SINGH K M. A robust rotation resilient video watermarking scheme based on the SIFT [J]. *Multimedia Tools and Applications*, 2018, 77(13):16419-16444.
- [47] 王丽云. 抗几何攻击的视频水印算法研究 [D]. 华中科技大学, 2011.
- [48] LIU Y, ZHAO H, LIU S, et al.: A robust and improved visual quality data hiding method for HEVC [J]. *IEEE Access*, 2018, 6:53984-53997.
- [49] NAIR R, VARADHARAJAN V, JOGLEKAR S, et al. Robust Transcoding Resistant Watermarking for H.264 Standard [J]. *Multimedia Tools and Applications*, 2014, 73(2):763-778.
- [50] CHO T, ROLAND C. On image analysis by the methods of moments [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1988, 10(4):496-513.
- [51] GRUBER M, HSU K. Moment-based image normalization with high noise-tolerance [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(2): 136-139.
- [52] KOTOULAS L, ANDREADIS I. Real-time computation of Zernike moments [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005,

- 15(6):801-809.
- [53] HUANG Z, CAO L. Bicubic interpolation and extrapolation iteration method for high resolution digital holographic reconstruction [J]. *Optics and Lasers in Engineering*, 2020, 130:106090.
- [54] LI J, WANG J, MA W, et al. Geometric attack resistant face watermarking using Zernike moments [C]. *Proceedings of the IEEE International Conference on Communication Technology*, 2008:753-756.
- [55] YE X, DENG M, WANG Y, et al. A robust DWT-SVD blind watermarking algorithm based on Zernike moments [C]. *Proceedings of the Communications Security Conference*, 2014:1-6.
- [56] LI J, ZHU Y. A geometric robust image watermarking scheme based on DWT-SVD and Zernike moments [C]. *Proceedings of the International Conference on Computer Science and Information Technology*, 2010:367-371.
- [57] YANG C, LI J, BHATTI U, et al. Robust zero watermarking algorithm for medical images based on Zernike-DCT [J]. *Security and Communication Networks*, 2021:1-8.
- [58] HE W, SUN J, YANG Z, et al. Video watermarking scheme based on normalization of pseudo-Zernike moment [C]. *Proceedings of the International Conference on Measuring Technology and Mechatronics Automation*, 2010:1080-1082.
- [59] XU G, WANG R. A blind video watermarking algorithm resisting to rotation attack [C]. *Proceedings of the International Conference on Computer and Communications Security*, 2009:111-114.
- [60] YUAN X, PUN C. Feature based video watermarking resistant to geometric distortions [C]. *Proceedings of the IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, 2013:763-767.
- [61] 毕厚杰. 新一代视频压缩编码标准:H.264/AVC [M].北京:人民邮电出版社, 2005.

- [62] DONG P, BRANKOV J, GALATSANOS N, et al. Digital watermarking robust to geometric distortions [J]. *IEEE Transactions on Image Processing*, 2005, 14(12):2140-2150.
- [63] SOURY H, YILMAZ F, ALOUINI M. Average Bit Error Probability of Binary Coherent Signaling over Generalized Fading Channels Subject to Additive Generalized Gaussian Noise [J]. *IEEE Communications Letters*, 2012, 16(6):785-788.
- [64] CHEN D, CHENG L. Spatially Adapted Total Variation Model to Remove Multiplicative Noise [J]. *IEEE Transactions on Image Processing*, 2012, 21(4):1650-1662.
- [65] LI Z, CHEN S, CHENG X. Dual Video Watermarking Algorithm Based on SIFT and HVS in the Contourlet Domain [J]. *IEEE Access*, 2019, 7:84020-84032.
- [66] ALGHONIEMY M, TEWFIK A H. Progressive quantized projection approach to data hiding [J]. *IEEE Transactions on Image Processing*, 2006, 15(2):459-472.
- [67] 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.
- [68] WANG K, LAVOUÉ G, DENIS F, et al. Hierarchical watermarking of semiregular meshes based on wavelet transform [J]. *IEEE Transactions on Information Forensics and Security*, 2008, 3(4):620-634.
- [69] NIU X, SCHMUCKER M, BUSCH C. Video watermarking resistance to rotation, scaling, and translation [C]. *Proceedings of the Security and Watermarking of Multimedia Contents IV*, 2002, 4675:512-519.
- [70] LIU Y, ZHAO J. A new video watermarking algorithm based on 1D DFT and Radon transform [J]. *Signal Processing*, 2010, 90(2):626-639.
- [71] CHEN B, WORNELL G. Quantization index modulation: a class of provably good methods for digital watermarking and information embedding [J]. *IEEE Transactions on Information Theory*, 2001, 47(4):1423-1443.

- [72] PARRAGA C, BRELSTAFF G, TROSCIANKO T, et al. Color and luminance information in natural scenes [J]. *Journal of the Optical Society of America*, 1998, 15(3):563-569.
- [73] XIN Y, LIAO S, PAWLAK M. Geometrically robust image watermarking on a circular domain [J]. *Pattern Recognition Letters*, 2007, 40(1):3740-3752.
- [74] ERNAWAN F, KABIR M N. A blind watermarking technique using redundant wavelet transform for copyright protection [C]. *Proceedings of the 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, 2018: 221-226.
- [75] CHEN S, CHEN Y, CHEN Y, et al. Robust Video Watermarking Using Normalized Zernike Moments [C]. *Proceedings of the 8th International Conference on Artificial Intelligence and Security*, 2022: 323-336.
- [76] WU C F, HSIEH W S. Digital watermarking using zero tree of DCT [J]. *IEEE Transactions on Consumer Electronics*, 2000, 46(1):87-94.
- [77] ROY S D, LI X, SHOSHAN Y, et al. Hardware implementation of a digital watermarking system for video authentication [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 23(2):289-301.
- [78] SHARMA S S, THAPA S, TANAY C P. A robust color video watermarking technique using DWT, SVD and frame difference [C]. *Proceedings of the International Conference on Pattern Recognition and Machine Intelligence*, 2017:148-154.
- [79] RASTI P, SAMIEI S, AGOYI M, et al. Robust non-blind color video watermarking using QR decomposition and entropy analysis [J]. *Journal of Visual Communication and Image Representation*, 2016, 38:838-847.

作者在攻读硕士学位期间公开发表的论文

- [1] CHEN S, CHEN Y, CHEN Y, ZHOU L, WU H. Robust Video Watermarking Using Normalized Zernike Moments [C]. Proceedings of the 8th International Conference on Artificial Intelligence and Security, 2022:323-336.
- [2] CHEN S, MALIK A, ZHANG X, FENG G, WU H. A fast method for robust video watermarking based on Zernike moments [J]. IEEE Transactions on Circuits and Systems for Video Technology, Early Access, 2023. (已录用)
- [3] 陈诗怡, 吴汉舟. 一种基于泽尔尼克矩的快速视频鲁棒水印方法. (申请号: 202310601266.0, 专利, 2023)

作者在攻读硕士学位期间所作的项目

[1] 国家自然科学基金青年项目“社交网络多用户协同的行为隐写”(项目编号: 61902235).

致 谢

时光荏苒，我即将结束在上海大学的硕士生涯。回想这三年来的学习和生活，是我人生中最宝贵的时光之一。在这段时间里，我学习了很多知识，也收获了许多珍贵的人际关系和成长经历。在这里，我想向许多人表达我的感激和敬意。

首先，我要感谢我的导师。我的导师是一位严谨认真，专注科研的学者。每当我在研究工作中有困惑时，他都能给我一针见血地指出问题本质，且用心指导我，让我明白如何去开展研究、写作表达。在我印象中，老师一直都很忙，每次只有晚上的时间有空，但每次他都会和我们分享很多他的经验，给我们很多指导。比较遗憾的是，自己很多时候心有余而力不足，不能帮助老师做更多的事情。希望以后还能有机会帮到老师，也希望老师以后能更加顺利。

此外，我也要感谢课题组其他老师。通常，和课题组其他老师的交流是在每个答辩场合。虽然自己不是他们的学生，但是他们也总会给我提出很中肯的建议和指导。他们都是非常优秀且专注科研的老师，能在此环境中学习，得到指导，真的让我感觉非常幸运，充满感激。

其次，我要感谢我的家人。尽管我知道自己非常普通，在似乎在他们眼里我总是最努力最优秀的。他们总是会在我身边，支持我，给予我鼓励和信心。很多时候会有困难却不能和家人说，但只要想着他们一直在我身边，就会感觉很幸福。

此外，我要感谢那些给予我帮助和支持的课题组同学。虽然我经常不在学校，和他们的交流也少，但他们总是会无保留地和我分享信息，和我交流问题，也会和我分享生活中有趣的事，把我当成朋友。当然，我也要感谢我的室友。我们四个人四种性格，有着不同的爱好，会互相包容，给予关心。我仍然记得我们在宿舍里分享生活，分享快乐的时光，这也将是我一生的美好回忆。

最后，感谢评审老师们，感谢您们抽出宝贵的时间对我的论文进行审阅和指导，您的建议和意见对我的学术研究和人生规划都有很大的帮助。在未来的道路上，我会一如既往地严谨、努力，争取取得更多的成就。

再次感谢所有为我提供支持和帮助的人，虽然我可能只是你们人生的过客，但希望我也有给你们带来快乐、能给你们带来帮助的时候。