

研究生选题报告

**题目：基于数字水印的图像认证算法研究**

|  |  |  |
| --- | --- | --- |
| **学号** | **：** | **M201570357** |
| **姓名** | **：** | **汪向红** |
| **专业** | **：** | **机械电子工程** |
| **指导教师** | **：** | **万立** |
| **院（系、所）** | **：** | **机械学院** |

**华中科技大学研究生院制**

一、课题来源、目的和意义

本课题来源于武汉天喻软件公司数据安全防护的一项技术研究，该技术的目标是向图像中嵌入数字水印，以达到对图像内容的真实性和完整性进行认证的目的。

课题的目的是研究出一项能满足实际应用需求的数字水印技术，该技术能够实现对图像的内容保护，不允许图像遭受丝毫的无意或有意的篡改，可对篡改区域定位，并可对篡改区域进行近似修复，以查看篡改前的内容，应用领域主要是含大量的图像的新闻，司法举证，电子商务等多个领域。

由于信息技术的飞速发展，越来越多的资料以数字的形式存在，然而数字媒体存在着易扩散和易修改的不足，数字媒体的内容是否安全、完整、可信问题日益凸显。传统的加密技术能够在数据传输过程中起到保护作用，但数据被接受或者解密之后将不再受到保护，同时加密的信息内容容易引起好事者的注意，激发他们的破译热情。为克服传统的加密技术的管理不便、不能定位篡改、不能区分篡改强度等缺点，迫切需要寻求另外一种途径。数字水印作为信息安全领域的一个新方向，甚至被认为是多媒体内容保护的最后一道防线，目前得到了学术界的广泛的研究，其发展也相当迅速。因此，本论文的研究课题不仅对信息安全领域的发展有一定的理论意义,同时具有重要的现实意义。

二、国内外研究现状分析

## 2.1 研究背景

### 2.1.1 信息加密技术

传统的加密算法一般是基于文本数据设计的，它把一段有意义的数据流(即明文)转换成看起来没有意义的数据(即密文)，如DES。该技术通过将明文加密为密文，使得在网络传递过程中非法拦截者无法从中获得信息，从而达到保密的目的。

传统的加密方法一直被认为是通信领域主要信息安全手段而受到极大重视，但近年来人们逐渐意识到其对多数字信息的保护和完整性认证具有一定的局限性。首先，加密算法只用在通信的信道中，密文数据因其不可理解性妨碍了数字信息的传播：其次，数字信息经过加密后容易引起攻击者的好奇和注意，并有被破解的可能；而一旦被破解，其内容完全透明，版权所有者就失去了对盗版的任何控制权；另外，密码学中的完整性是通过数字签名方式实现的，它并不是直接嵌入到数字信息当中，因此无法察觉信息在经过加密之后的再次传播中内容的改变；最后，传统的加密手段无法对数字信息的修改进行局部定位，也无法恢复。因此，基于加密技术的多媒体保护和完整性认证有其严重的缺陷，亟待寻求一种新的技术弥补其不足。

### 2.1.2 数字水印技术

随着数字媒介的出现，数字作品的知识产权保护及认证需求导致了数字水印的出现。因此可以说数字水印是信息时代的特有产物。数字水印技术是指用信号处理的方法在数字媒体中嵌入隐蔽的标记，这种标记不影响原内容的价值和使用，并且只有通过专用的检测器或阅读器才能提取[1]。在多媒体中加入数字水印可以确立版权所有者信息、认证多媒体来源的真实性、识别购买者、提供关于数字内容的其它附加信息、确认所有权和跟踪侵权行为[2]。由于数字水印技术为解决开放性网络上的版权保护、来源认证、篡改认证、网上银行、用户跟踪等一系列问题提供了崭新的解决思路，因此一经提出便在多媒体信息安全领域得到迅速的发展。

数字水印技术除了应具备信息隐藏技术的一般特点外，还有着其固有的特点和研究方法。例如，从信息安全的保密角度而言，隐藏的信息虽然被破坏掉，而系统可以视为安全的，因为密码信息并未泄露；但是，在数字水印系统中，隐藏信息的丢失，即意味着版权信息的丢失，从而也就失去了版权保护的功能，也就是说，这一系统就是失败的。由此可见，数字水印技术必须具有较强的鲁棒性、安全性和透明性。

### 2.1.3 典型数字水印系统模型[2]

图2-1为水印信号嵌入模型，其功能是完成将水印信号加入原始数据中；图2-2为水印信号恢复模型，其负责从水印数据中提取出水印信号；图2-3为水印信号检测模型，用以判断某一数据中是否含有指定的水印信号。

密钥

水印信息

载体数据

水印嵌入算法

水印载体数据

图2-1 水印信号的嵌入

水印载体数据

密钥

原始载体数据

水印提取算法

水印信息

图2-2 水印信号的恢复

水印载体数据

密钥

原始水印

原始载体数据

水印检测算法

水印存在与否？

图2-3 水印信号的检测

图2-2和2-3中的虚框表示在提取或判断水印信号时，原始载体数据不是必要的。

### 2.1.4数字水印的分类及应用

多媒体数据存在形式繁多，如语音、图像、视频或者文本等。水印信息的存在形式也是多样化的，如一段标识、序列号、伪随机数或者是图像、音频、视频数据等。因此，数字水印有多种分类方法。在本文中原始数据主要选择数字图像。从不同的角度对数字水印进行分类，有下面几种主要的分类方法[3]：

1．按用途不同分为：图像认证水印、版权保护水印、隐藏标识水印、防复印水印等。

2．按载体不同分为：文档水印、图像水印、音频水印、视频水印等。

3．按水印抵抗攻击的能力不同分为：鲁棒水印和脆弱水印。依据脆弱水印鉴别篡改的能力不同，脆弱水印又分为完全脆弱水印和半脆弱水印。有的文献中将完全脆弱水印和半脆弱水印统称为脆弱性水印。

4．按嵌入位置的不同分为：空域水印和变换域水印。空域水印是直接修改图像的像素（如直接修改像素的最低位）；变换域水印是对图像进行各种变换后再进行水印的嵌入，如 DCT（离散余弦变换)、DWT（离散小波变换）、DFT（离散傅里叶变换）等。

5．按加水印后是否可见分为：可见性水印和不可见性水印。

6．按检测水印时是否需要原始图像分为：私有水印和公有水印。私有水印检测时需要原始图像的参与，公有水印检测时不需要原始图像的参与（也称为盲水印检测）。

目前数字水印技术得到了广泛的应用，主要包括以下几个方面[3]：

1．版权保护：即数字作品的所有者可用密钥产生一个水印，并将其嵌入原始数据，然后公开发布他的水印版本作品。当该作品被盗版或出现版权纠纷时，所有者即可利用图2-2或图2-3的方法从盗版作品或水印版本作品中获取水印信号作为依据，从而保护所有者的权益。对这种应用领域来说，水印技术必须有较好的鲁棒性、安全性、透明性和水印嵌入的不可逆性。

2．篡改提示：当数字作品用于法庭、医学、新闻和商业时，常需要确定它们的内容是否被修改、伪造或者特殊处理过。为实现该目的，通常可以将原始图像分成多个独立块，再将每个块加入不同的水印。同时可通过检测每个数据块中的水印信号，来确定作品的完整性。与其他水印不同的是，这类水印必须是脆弱的，并且检测水印信号时，不需要原始数据。

3．加指纹：为避免未经授权的拷贝制作和发行，出品人可以将不同用户的ID或序列号作为不同的水印（指纹）嵌入做平的合法拷贝中。一旦发现未经授权的拷贝，就可以根据此拷贝所恢复出的指纹来确定它的来源。对这种应用领域来说，水印技术除了具有版权保护的应用中的特性外，还必须具有防止串谋攻击（多拷贝攻击）等功能。

4．标题和注释：即将作品的标题、注释等内容（如，一副照片的拍摄时间和地点等）以水印形式嵌入做平中，这种隐式注释不需要额外带宽，而且不易丢失

5．使用控制：这种应用的一个典型的例子是DVD防拷贝系统，即将水印信息加入DVD数据中，这样DVD播放机可通过检测DVD数据中的水印信息而判断其合法性和可拷贝性。从而保护制造商的商业利益。

2.2 图像认证水印技术概述

图像认证水印技术是传统密码学与隐蔽通信相结合的产物，它涉及到密码学、通信理论以及数字图像处理等方面的技术。认证水印是把标志信息作为水印嵌入到数字多媒体内容中（保证嵌入的信息不影响感官质量为前提），当需要对多媒体内容进行认证时，提取出水印信息以鉴定数字多媒体内容是否真实完整。基于数字水印的图像认证系统的一般框架实现包括：认证水印的产生、嵌入和检测认证。

图像认证属于盲认证，即在认证时无法获得原始信息，往往通过比较一些与原始图像有关的认证信息来判断图像完整性和真实性。根据传输这些信息采用的方法不同，可将图像认证技术分为基于数字签名的图像认证技术和基于数字水印的图像认证技术。基于数字签名的图像认证技术只是简单的将这些认证信息附加在原始图像之后，而基于数字水印的图像认证技术克服了基于数字签名的图像认证技术的缺点，借助信息隐藏的方法将认证信息隐藏在原始图像本身，通过对原始图像进行一定的修改，攻击者很难发现认证信息的存在和提取认证信息。通过认证水印技术可以判断数字图像是否受到篡改以及篡改发生的位置，甚至恢复被篡改的内容。因此，当前图像完整性认证方案研究的热点和难点都集中在脆弱水印技术上[1]。

### 2.2.1 认证水印技术要求

认证水印技术是近几年研究的新课题，作为数字水印的一个分支它具有数字水印系统的一般要求，也有一些区别于鲁棒水印的特殊要求和评价标准，了解这些特点后才能提出有效的、不可见性好的脆弱性水印算法。

作为数字水印技术的一种，用于数字图像完整性保护和内容认证的脆弱水印的嵌入原理与一般水印的嵌入原理基本相同，从数字信号处理的角度上看都是对载体图像的调制过程。脆弱水印除了具有一般水印系统的基本特征如不可感知性、安全性等特点外，还应能够完成认证功能[4]。具体来说，脆弱水印应满足以下特殊要求[1]：

（1）检测篡改。脆弱水印的最基本的要求就是具有认证功能，即载体图像遭到篡改时，认证水印系统能有效地检测篡改并指出数字图像被篡改的程度及定位篡改发生的位置，甚至有些应用场合要求系统能够恢复出被篡改的图像内容。

（2）脆弱性和鲁棒性。为了鉴别数字图像的内容是否发生篡改，则要求数字水印为脆弱性的水印。即一旦检测不出水印的存在，则说明载体图像内容已经被正常的信号处理操作改变或被篡改。实际的应用场合，大部分数字图像要经过一定的压缩处理后才在网络上传输，以便提高信息传输量，这就要求脆弱性的数字水印也具有一定的鲁棒性。

（3）不可感知性。同鲁棒性水印相同，脆弱性水印也要求人眼很难觉察到嵌入水印前后载体图像的变化，以便保证数字图像的使用价值。

（4）盲检测。水印检查时，使用原始的宿主信号，更有利于检测和提取信息，如鲁棒性水印用于数字媒体的版权保护，检测时允许使用原始数据，但是检测时用到的原始宿主信号容易暴露给恶意的攻击者，而且在某些应用中，并不能获得原始载体图像信息。脆弱水印的功能就是用来验证待检测图像的完整性、真实性，如果接收方有原始图像，那对图像真实性进行验证就显得多余。对于认证系统来说，盲检测这一点是必须的。因此，除了极个别的方案外，目前主要研究的是盲检测水印技术。

（5）更高的安全性。脆弱性水印与鲁棒性水印的最大差别就在于脆弱性水印对各种攻击保持着高度的敏感性，脆弱性水印主要应用于对各种重要数据的精确认证和内容的模糊认证，例如医疗过程中的影像，或者是用于法庭举证的图片、视频等等。在这种场合下，脆弱水印的安全性就成为非常重要的问题。为了增加水印的安全性，很多研究者采用对水印信息加密的方式，利用一个控制水印信息的密钥来保证水印不被攻击者获取。密钥的选择方式有很多种，根据载体图像的特征选择最合适的控制密钥能大大提高水印的安全性。

（6）更强的敏感性。图像认证系统要求能够检测出被保护图像上的篡改操作。目前针对数字图像的各种攻击和处理可分为两类：

①无意的攻击。采用压缩编码标准如 JPEG、JPEG2000 等对图像进行编码；数字图像处理中的图像增强等操作。

②有意的攻击。保护数字图像的真实性、完整性是图像认证系统的主要功能，因此存在某些设法篡改图像内容但不会破坏水印信息以求通过认证的“伪认证”攻击。

### 2.2.2 认证水印所受攻击

用于数字多媒体内容完整性、真实性认证的水印信息本身具有一定脆弱性，常见的对鲁棒水印的攻击方法，如简单攻击，删除攻击，混淆攻击对认证水印系统的效果微乎其微[2]。图像认证系统最需要抵抗的是试图篡改图像的内容却不损坏水印信息的“伪认证”攻击，为保证认证水印嵌入时的安全性，通常使用密码算法或密钥，使攻击者无法提取水印信息，但当同一密钥嵌入多幅图像后，攻击者会对这些数字图像进行大量分析，破译密钥，进而对图像进行恶意篡改操作，却仍可通过认证[1]。目前，关于数字水印认证技术的攻击算法报道还较少，根据已有的文献来看，对图像水印认证系统可能实施的有效伪认证攻击有以下几种：

1．恒值攻击（Constant-average Attack）：恒值攻击是指在不改变水印基础上任意改变像素的值保证图像块的像素均值不变，例如一个2×2的图像块的四个像素的像素值为[50 60; 60 60]，将图像块像素信息修改为[20 90; 90 20]，图像块的均值不变。

2．内容攻击（Content Attack）：内容攻击和恒值攻击有些类似。也是在保持不改变水印的前提下，对非水印部分任意修改。有些算法的水印产生与非水印层的无关，因此，只要不去修改水印，被篡改的图片就可以通过验证，不会发现图片已经被篡改了。

3．拼贴攻击（Collage Attack）主要常见于基于分块的脆弱水印算法中，它利用水印算法在多幅图像中嵌入同一水印信息时存在的安全漏洞，对于嵌入相同水印信息的多幅图像，在保持图像分块相对位置保持不变的情况下，将属于不同图像的图像块进行拼贴得到篡改后的图像。如果攻击者在进行图像拼贴时，精心地选择图像块，保证拼贴后图像一定的视觉质量，那么几乎所有的基于独立分块的脆弱水印算法都将无法检测到篡改。

4．量化攻击（Vector Quantization Attack）也是一种针对基于独立分块的脆弱水印算法的攻击方法。这种攻击的前提条件是每个图像块中嵌入的水印信息与其他图像块的内容无关，这样如果两个图像块中嵌入的水印信息相同，就可以将它们的内容互换而不会被检测到篡改。预防量化攻击最有效的方法是设计脆弱水印算法时保证每个图像块生成的水印信息依赖于其他图像块的内容。

5．统计分析攻击：当采用同一密朗在不同图像中嵌入同样认证信息时，攻击者通过对大量取证图像统计分析，有可能分析出水印嵌入规律。

### 2.2.3 认证水印算法研究现状

脆弱数字水印从1994年被提出至今，已有近20年的发展历史，在这期间，脆弱水印凭借其表现出的巨大的社会经济潜力，越来越受到人们的重视。IHW（国际信息隐藏学术研讨会）、IWDW（国际数字水印研讨会）、ICIP（IEEE 图像处理国际会议）等国际重要学术会议定期报道数字水印技术的最新发展，1999 年国内召开了第一届 CIHW 信息隐藏学术会议，之后每年召开一次，对国内数字水印技术的发展进行报道。英国剑桥大学、美国麻省理工学院、日本NEC研究所等著名研究机构最早开始数字水印相关领域的研究，并取得不少的成果。国内对于脆弱水印的研究是从2003年开始的，中科院自动化研究所、北京大学、清华大学等一批国内科研机构和院校也逐步投入到脆弱水印的研究领域。

由于脆弱水印的类型错综复杂，很难一类地分清，下面从脆弱水印算法的功能角度，对脆弱水印算法的国内外发展进行综述。

1．像素级篡改定位的脆弱水印技术发展：1995年Walton[5]率先提出的LSB方法是最早出现的像素级篡改定位的空域脆弱水印方法，其主要思想是计算载体图像像素点的最高七个比特位的校验和，在将其嵌入对应的最低有效位中。文献[6]使用二值图像作为水印内容，将图像尺寸和最高七个比特位进行Hash运算，对运算结果与二值水印进行异或操作，操作结果经公钥加密后嵌入到LSB位。文献[7]也是以二值图像为水印，通过将水印的某些特征映射为随机数0或者1，形成一个查找表LUT（Look Up Table），通过利用LUT对像素进行量化实现水印信息的嵌入。文献[8]将文献[7]的算法从空间域推广到变换域，使得算法可以与JPEG压缩标准相兼容。文献[9]将图像中的像素点与其周围的上下和左右各*k*个点，共计2*k+*1个像素点结合起来生成认证信息，最后嵌入到图像最低位。然而需要指出的是基于像素级脆弱水印算法存在以下几个问题：嵌入容量大，算法漏警率高，且容易受到各种伪认证攻击，因此，目前基于像素级的脆弱水印算法正在逐渐减少。

2．块级篡改定位的脆弱水印技术发展：1998年，Wong等在文献[10,11]提出一种基于独立分块的脆弱水印方法。该算法将图像划分为互不重叠的独立小块，每个小块都会生成相应的水印并嵌入到图像块自身中，从而实现了篡改的分块定位。1999年，Fridrich在文献[12]中首次提出一种DCT变换域的分块自嵌入水印。该算法首先将载体图像划分为互不重叠的大小为 8×8 的图像块，将一个子块最高七位的DCT系数按照特定的码长进行量化编码，并采用固定偏移方式嵌入到另一个子块的LSB平面位。文献[13,14]为了克服固定偏移方式的缺点，提出一种基于混沌的DCT域分块自嵌入水印方法。该算法利用混沌系统的极端敏感性和伪随机特性，实现水印信息的随机偏移嵌入，进一步扩大了密钥空间，提高了算法的安全性。文献[15] 提出了一种针对彩色图片检测的算法。算法将彩色图片分成不重叠的4×4小块，每一个小块的最重要的6位用来产生验证数据，采用块截断编码的方式来生成恢复数据。为了抵御量化攻击，采用一个2维变换公式来生成块映射。在文献[16]中，Zhang根据像素将图像分为不重叠的2×2的小块，之后对每一个小块进行DCT变换。同样的，Zhang的算法也将水印分成两种数据，恢复数据和验证数据。恢复数据由DCT变换后的DC系数生成，验证数据则是根据这些恢复数据生成，并采用logistic映射生成块映射。

3．篡改可恢复的脆弱水印技术发展：为了提高图像的利用价值，有些图像认证的应用场合要求脆弱水印具备篡改恢复能力。篡改恢复可以分为精确恢复和近似恢复。精确恢复要求将被篡改的数据完全恢复到原始状态。由于图像认证只要求恢复后的篡改内容与原始内容大体一致，允许恢复后的数据与原始数据存在一定的偏差，因此对于图像认证只需实现近似恢复即可。脆弱水印的近似恢复方法一般可以分为自嵌入、纠错码恢复和盲恢复三种。自嵌入是将图像自身特征内容作为水印信息，纠错码恢复是在原始图像中嵌入错误纠正码，盲恢复是首先判断篡改类型，然后对篡改进行逆向操作。在三种方法中自嵌入水印是最为重要的一种篡改恢复方法。

自嵌入水印最早由Fridrich[12]于1999年提出，之后出现了许多改进的自嵌入水印算法[17-20]。文献[19]在文献[12]的基础上，重新设计了DCT量化系数的编码位长，进一步提高了算法的安全性。文献[20]进一步改进自嵌入方法，每个图像块中嵌入的水印信息包含了子块的认证信息及其偏移子块的压缩信息，水印信息同样嵌入到子块的LSB位平面，但最多修改LSB位平面中一半的数据，因此含水印图像的质量较好。文献[21,22]对文献[19]中分层的分块认证方法进行了改进，利用分层检测的方法对水印进行篡改检测与定位。文献[23]在分析文献[21]的基础上提出一种图像篡改检测与恢复的双水印算法，该算法中每个图像块的生成的水印信息都有一个拷贝版本，通过将两份相同的水印嵌入到不同的图像块中，大大提高了篡改恢复的质量，在篡改量较大的情况下仍然能够保持较高的恢复质量。不过，该算法存在较为严重的安全性问题，该算法不能抵御多种常见的“伪认证”攻击，比如拼贴攻击，恒值攻击等。文献[24]改进了文献[11]的算法，通过量化的方式为原始图像创建了一个量化表，用来恢复使用，并且与文献[23]，拥有更好的恢复效果。但是算法同样存在安全漏洞。

## 2.3 参考文献

[1] 孙圣和，陆哲明，牛夏牧等. 数字水印技术及应用[M].北京：科学出版社，2004.

[2] 张小华. 基于数字水印的图像认证技术研究[D].西安：西安电子科技大学，2004.

[3] 王炳锡，陈琦，邓峰森. 数字水印技术[M].西安：西安电子科技大学出版社，2003.

[4] 张弘，王颖.脆弱性数字图像水印算法比较研究[J].计算机工程，2008, 34(3):174-177.

[5] Walton S. Information Authentication for a Slippery New Age. Dr.Dobbs Journal, 1995,20(4):18-26.

[6] Wong P W. A Public Key Watermark for Image Verification and Authentication. Proceedings of the IEEE International Conference on Image Processing[C], Chicago, Oct.1998:455-459.

[7] Yeung M, Mintzer F. Invisible watermarking for image verification[J]. Journal of Electronic Imaging, 1998,7(3):578-591.

[8] Wu M, Liu B. Watermarking for image authentication[A]. In: Proceedings of the IEEE International Conference on Image Processing [C], Chicago, USA, 1998, 2: 437-441.

[9]谢建全, 阳春华. 一种像素级的图像篡改认证算法[J]. 计算机应用, 2007, 27(6):1337-1342.

[10] Wong P W. A watermark for image intergrity and ownership verification[C]//Proceeding of the IS & T PIC Conference, Oregon, Portland, 1998.

[11] Memon N, Wong P. Secret and public key authentication watermarking schemes that resist quantization attack[C]//Proceeding of the SPIE International Conference on Security and Watermarking of Multimedia Contents Ⅱ. San Jose, USA, 2000, 3971:417-427.

[12] Fridrich J, Goljan M. Images with self-correcting capabilities[A].In Proceedings of International Conference on Image Processing[C], Kobe, Japan 1999, 3: 25-28.

[13] 和红杰，张家树. 基于混沌置乱的分块自嵌入水印算法[J]. 通信学报，2006, 27(7): 80-85.

[14] 和红杰，张家树. 基于混沌的自嵌入安全水印算法[J]. 物理学报，2007，56(6): 3092-3100.

[15] Huang S C, Jiang C F. A color image authentication and recovery method using block truncation code embedding[J]. Journal of Marine Science and Technology, 2012, 20(1): 49-55.

[16] Zhang J, Zhang Q, Lv H. A novel image tamper localization and recovery algorithm based on watermarking technology[J]. Optik-International Journal for Light and Electron Optics, 2013, 124(23): 6367-6371.

[17] 和红杰，张家树. 区分不同篡改的脆弱水印算法[A]. 第十二届全国图像图形学术会议[C], 2005, 140-144.

[18] 和红杰，张家树，田蕾. 能区分图像或水印篡改的脆弱水印方案[J]. 电子学报，2005,

33 (9): 1557-1561.

[19] 张鸿宾, 杨成. 图像的自嵌入及窜改的检测和恢复算法[J]. 电子学报, 2004, 32(2): 196-199.

[20] 钱振兴, 程义民, 王以孝等. 一种图像自嵌入方法[J]. 电子学报, 2006, 34(7): 1347-1350.

[21] Lin P L, Hsieh C K, Huang P W. A hierarchical digital watermarking method for image tamper detection and recovery[J]. Pattern Recognition, 2005, 38(12):2519-2529.

[22] 刘泉, 江雪梅. 用于图像篡改定位和恢复的分层半脆弱数字水印算法[J]. 通信学报, 2007, 28(7):104-110.

[23] Lee T Y, Lin S D. Dual watermark for image tamper detection and recovery[J]. Pattern Recognition, 2008, 41(11):3497-3506.

[24] Yang C W, Shen J J. Recover the tampered image based on VQ indexing[J]. Signal Processing, 2010, 90(1): 331-343.

三、研究目标、内容及关键技术

3.1 预期目标

论文将结合武汉天喻软件公司数据安全防护的技术研究需求，研究以图像为研究对象，基于信息安全领域的混沌技术，提出一种基于数字水印的图像认证算法。该算法应具有以下特性：

* 定位精度良好，同时具有较低的虚警率，漏警率；
* 不可感知性高，要求人眼很难觉察到水印嵌入前后图像的变化，以便保证数字图像的使用价值；
* 安全性高，能够抵御目前主流的“伪认证”攻击；
* 盲检测，对于认证算法来说，盲检测这一点是必须的；
* 修复效果较好，即使篡改面积较大，修复效果能可接受；
* 敏感性强，能够检测出被保护图像上的任何无意有意的篡改操作。

3.2 研究内容

论文将在分块的基础上向图像中嵌入水印信息。分块的目的是为了控制定位精度，因此块的面积越小代表精度越高，但是块的面积过小，嵌入的水印容量也就越少。首先将图像分成合适的大小，但是块与块之间的独立不能抵御“伪认证”攻击中的VQ攻击和拼贴攻击，因此需建立块与块之间一一映射关系。建立联系的方式是首先根据块自身的特点生成验证数据和恢复数据。然后将验证数据嵌入于自身块当中，而将恢复数据嵌入与之对应的映射块当中。检测的时候，先检验验证数据，验证数据通过了检验才可以使用嵌入该块当中恢复数据。此外，恢复数据也可以用来验证产生恢复数据的块是否被篡改过。因此，算法具有多级检测功能。

3.3 关键技术

本课题的关键技术如下：

1．图像分块与块映射的生成

图像分块的算法不仅与定位精度相关，还与修复效果相关，因此，分块大小必须适中。如果块的面积过大，仅仅因为一个像素点的改变而将整块标记为篡改，此时虚警率就会过高，影响整个算法的效果。如果块的面积过小，导致水印容量过小，嵌入的水印有限，此时漏警率就会过高，同样会影响整个算法的效果。因此块的大小选择非常重要。在确定了分块的大小之后，下一步就是块与块之间一一映射。映射的产生主要分为两种，线性映射和非线性映射。为了提高水印的安全性，现在算法较多都是采用非线性映射。

2．水印的产生

由于算法中涉及到篡改检测和修复技术，因此我们将水印分为验证数据和恢复数据。验证数据的产生主要来自于对图像自身的内容编码，如果验证数据跟图像内容不相关，则很容易遭受伪认证攻击中的内容攻击和恒值攻击。恢复水印是当某个块被检测为篡改时，在使用恢复数据近似恢复被篡改块的信息，但是，使用前有个条件，就是含有恢复水印那个块需先证明这个块没有受到攻击，才能说明嵌入其中的恢复水印也没有遭受篡改，这样才能将其用来恢复被篡改的块。

3．水印嵌入和提取算法

对水印的嵌入算法需满足以下几个要求。首先，水印不可感知性高，嵌入后，人眼不能观察到有任何一丝明显的变化，且不影响图像的正常使用。其次，算法要安全和可靠。水印不仅要能够抵御常见的“伪认证”攻击，而且在水印嵌入后，不受攻击的情况下，能够100%通过验证。最后，算法要满足盲检测的条件，即在不需要原始图像，水印的情况下，即可检测图像是否被篡改过。

4．混沌技术在数字图像认证保护中的应用

混沌系统在数字水印技术中的应用研究主要从两方面进行考虑，对于脆弱水印算法，主要是利用混沌对于初值的敏感性和长期不可预测性，生成与图像内容或特征有关的混沌序列作为水印信号，或者利用混沌映射将水印嵌入原始图像。对于半脆弱水印算法，主要利用混沌映射对水印信号进行预处理或者直接利用混沌序列的低频特性产生水印信号。目前，混沌系统在数字图像认证领域的应用还处于初级阶段，如何将混沌系统具有的显著优点应用于数字图像的认证保护还有很多问题需要探讨。构造具有通用性的基于混沌系统的水印生成模型，且提出兼具良好的篡改检测能力以及一定修复能力的基于混沌系统的数字水印技术是一个有待解决的问题。

四、研究方法和技术路线

4.1 置乱和混沌技术的应用

置乱技术一直是作为图像加密的一种较为常见的技术，通过对图像上的像素点进行打乱来对图像进行加密。本课题将采用置乱技术用于生成非线性块映射，以此来增强算法的安全性能。

混沌普遍具有以下4个性质：对初始值的极端敏感性：对初始值的任何微小改动，都会引起后来结果的巨大变化，这就是所谓的“差之毫厘，谬以千里”。不可长期预测性：处于混沌状态的系统是不能长期预测的。遍历性：当时间趋于无穷时，混沌系统的轨迹会遍历到当前空间的每一个点。伪随机性：由确定性系统产生看似随机的结果。倘若将混沌技术运用到水印算法当中，进一步扩大了密钥空间，提高了算法的安全性。

置乱技术和混沌技术的应用，极大的增加了水印的安全性能和抗攻击性。

**4.2水印的生成**

现有的二值文本图像脆弱水印算法中水印信息主要有以下三种：

1、随机序列：文献是通过密钢生成一定长度的与图像内容无关的二值随机序列作为水印信息嵌在图像的空域中。

2、二值图像：文献，是把包含所有者信息或者特殊内容的二值图像嵌在载体图像中。这些二值图像虽能一定程度上表征所有者信息，但由于完全与图像自身内容无关，为攻击者提供便利，造成了较大的安全隐患。

3、结合图像内容生成水印：为提高算法的安全性，文献则结合图像内容生成二值水印信息。

为了抵抗伪认证攻击，提高水印的定位精度以及提供较好的恢复效果，因此，毫无疑问，肯定要结合图像内容来产生，具体如何生成，还需考虑通过实验来不断的调整。

4.2 算法实现技术

算法采用matlab语言编程实现，因此，在掌握matlab语言的一些基础应用之外，还必须学习有关图像处理方面的知识，最后做成一个水印嵌入和检测系统。

4.3实验验证方案

本课题按所设计算法对图像嵌入水印后，还要做以下几个工作来验证算法可行性及合理性：

1、检测效果：对图像进行任意的局部篡改，通过评价指标如虚警率和漏报率等对检测效果进行评价，并与现有的算法进行比较，确保算法具有较好的定位效果。

2、恢复效果：篡改定位后，对篡改区域进行恢复，并对恢复后的图片与原图进行比较和评价，在与现有的恢复算法进行比较，以此来判断恢复算法的恢复效果。

3、伪认证攻击分析：针对现有的攻击方法，对含水印图像进行各种攻击，再对攻击后的图片进行检测，并对攻击后的检测效果进行评价，以此来判断算法对攻击的承受能力。

五、预期的成果和可能的创新点

**预期的成果如下：**

该水印系统最终集成到intkey或CAD软件中，可抵抗预期攻击，并实现精确定位，产生一定的经济效益。

**可能的创新点：**

* 多层检测机制

对水印进行分类，分成验证数据和恢复数据，恢复数据也可以当作验证数据，形成多层检验机制，增强算法的抗攻击性和定位精度。

* 水印的加密算法

将混沌系统与图像的特征相结合生成混沌密钥，并对水印进行加密，且不同图像的混沌密钥是不同的，这将极大增强水印的抗攻击性能，同时混沌系统的使用，极大的增加了密钥的范围，使得破解密钥难度加大。

六、计划进度安排

* 2015.09– 2015.11 基础理论学习，资料查阅
* 2015.12– 2016.03 研究分块算法
* 2016.04– 2016.07 研究水印嵌入和提取算法
* 2016.08– 2016.10 水印评估及改进
* 2016.11– 2016.12 系统实现
* 2017.01– 2017.04论文撰写

七、可行性分析

指导老师对于数字水印已有多年的研究，积累了丰富的研究经验，实验室的师兄也是在做与水印相关的工作，本人目前阅读了一些国内外的文献，为之后的研究打下了一些基础。

八、目前演究进展

前期调研已经完成，并提交报告

研究的总体框架已经形成

九、导师评语

九、导师评语