3.2、数字水印基本理论

信息安全中心
钮心忻、杨榆、雷敏

数字水印的提出

- 水印
  - 存在于纸张、纸币中，用于标识真伪
- 数字水印
  - 对数字产品标识真伪
    - 数字图书馆、网络音频和视频、数字地图等

数字水印系统三要素

- 数字水印
- 水印嵌入算法
- 水印检测算法

数字水印构成方式

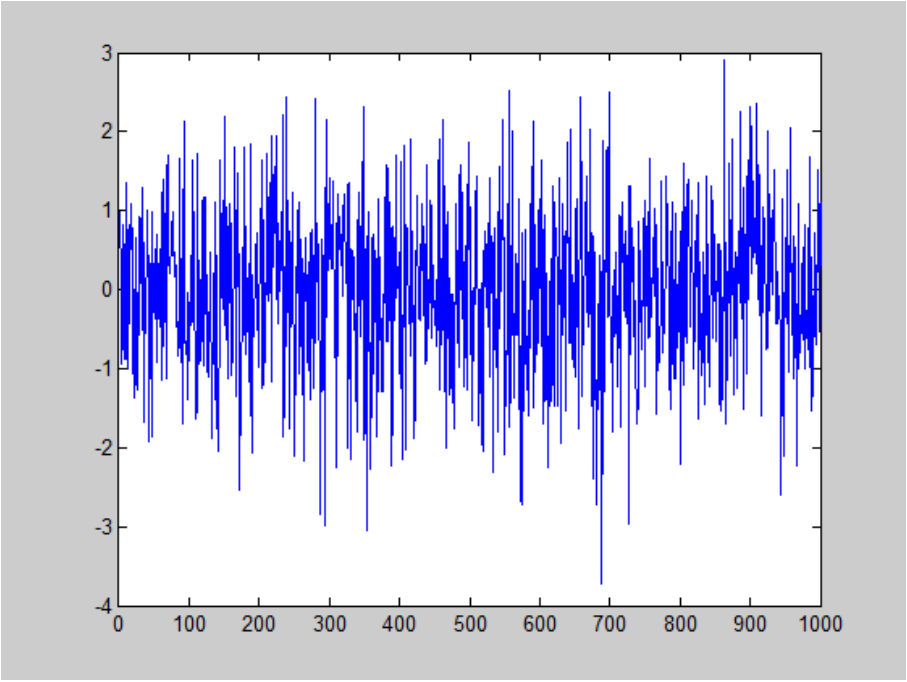
- 有意义水印
  - 文本信息，例如：“微软版权所有”
  - 图片信息，商标、徽标、标识等，例如：

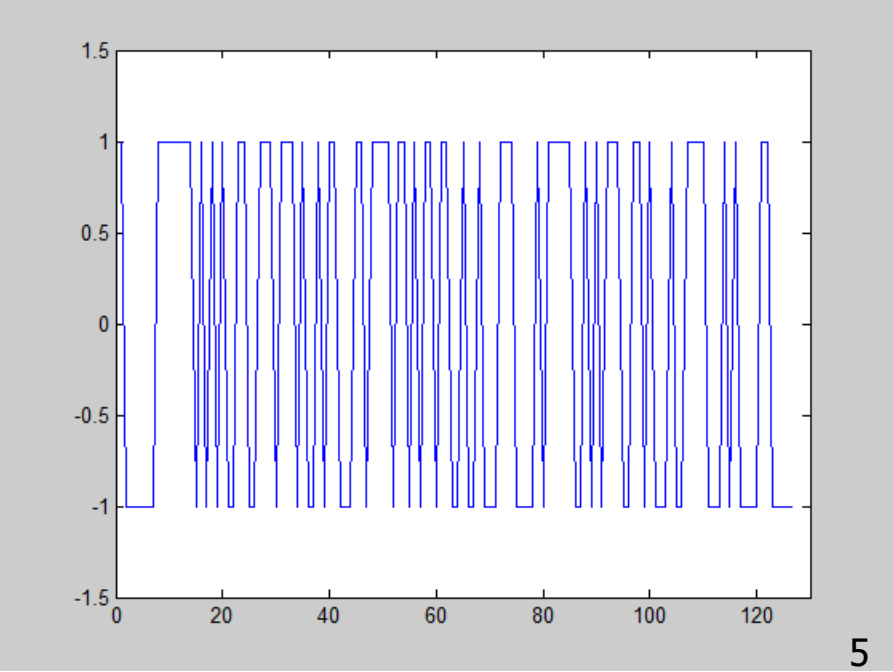
数字水印构成方式

- 无意义水印
  - 伪随机序列，随机噪声等

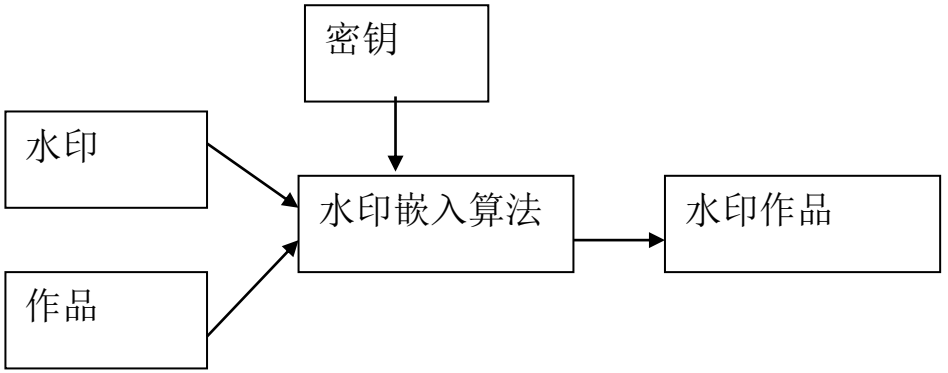
随机噪声作为水印

伪随机序列作为水印

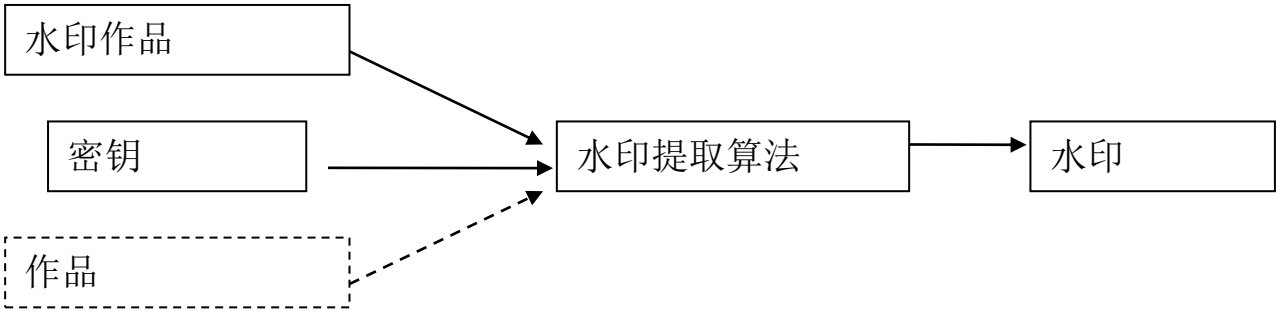




数字水印嵌入和提取模型



水印嵌入模型



水印提取模型

水印嵌入模型

- 设 I 为作品，W 为水印，K 为密钥，处理后的水印为

$$\tilde{W} = F(I, W, K)$$

- 水印嵌入过程：设水印嵌入函数 E，作品 I 和水印 $\tilde{w}$，嵌入水印后的水印作品为

$$I_W = E(I, \tilde{W})$$

水印提取模型

- 水印提取过程：设水印提取函数 D ，水印提取的两种形式：
  - 提取水印信息：如文字、徽标
  - 0-1判决：判定水印存在与否

$$W^* = D(\hat{I}_W, I, K)$$

$$C(W, W^*, K, \delta) = \begin{cases} 1, & W\text{存在} \\ 0, & W\text{不存在} \end{cases}$$

水印算法性能指标

- 水印的应用一般认为在广播监视、版权标记、版权跟踪、内容认证、拷贝控制等方面，不同应用对水印算法性能有不同的要求。
- 根据应用的需求调整水印算法性能，使各个性能指标在调整的过程中获得妥协的平衡。
- 某个性能指标的改善，一般是通过牺牲其他特性的性能得到的。

水印算法性能指标

- 安全性
  - 水印系统抵抗恶意攻击的能力。
- 稳健性（健壮性、鲁棒性）
  - 水印系统抵御常规处理的能力。
- 透明性（保真性、不可感知性）
  - 算法对载体感官质量的影响程度，作品在被算法处理前后的相似程度越高，透明性越好。
- 容量
  - 在作品中能够嵌入的最大有效载荷比特数。
- 计算量
  - 嵌入算法与提取算法的计算成本。

水印算法性能指标——安全性

- 水印算法/系统的安全性
  - 假设攻击者知道系统部分知识，并对系统进行了恶意攻击。在这种情况下，若数字水印能够被准确提取和判断，并为版权保护或者完整性保护提供清晰的结论，则称系统是安全的。
- 水印攻击类型
  - 非授权嵌入、非授权提取、非授权去除。

水印算法性能指标——稳健性

- 稳健性
  - 算法承受常规处理的能力。
- 常规处理
  - 滤波、去噪、格式转换、打印-扫描、重采样、有损压缩等等
  - 几何失真
    - 旋转、平移、缩放（RST：Rotation，Translation，Scaling）
    - 抖动（随机去除若干行、列...）

水印算法性能指标——透明性

- 随着嵌入水印信息量的增加，水印作品的感官质量必然下降
- 水印算法透明性的评价方法
  - 主观
  - 客观

水印透明性的主观度量

- 以图像为例
- 观察者对图像进行观测，给出评价
- ITU-R Rec.500图像质量度量

等级级别	损　害	质　量
5 4 3 2 1	不可察觉 可察觉，不让人厌烦 轻微的让人厌烦 让人厌烦 非常让人厌烦	优 良 中 差 极差

水印透明性的客观度量

差分失真度量	
平均绝对差分	$AD = \frac{1}{XY}\sum_{x,y}\left\|p_{x,y} - \tilde{p}_{x,y}\right\|$
均方误差	$MSE = \frac{1}{XY}\sum_{x,y}\left(p_{x,y} - \tilde{p}_{x,y}\right)^2$
L^p—范数	$L^p = \left(\frac{1}{XY}\sum_{x,y}\left\|p_{x,y} - \tilde{p}_{x,y}\right\|^p\right)^{1/p}$
拉普拉斯均方误差	$LMSE = \sum_{x,y}\left(\nabla^2 p_{x,y} - \nabla^2 \tilde{p}_{x,y}\right)^2 / \sum_{x,y}\left(\nabla^2 p_{x,y}\right)^2$
信噪比	$SNR = \sum_{x,y} p_{x,y}{}^2 / \sum_{x,y}\left(p_{x,y} - \tilde{p}_{x,y}\right)^2$
峰值信噪比	$PSNR = XY \max_{x,y} p_{x,y}{}^2 / \sum_{x,y}\left(p_{x,y} - \tilde{p}_{x,y}\right)^2$

水印透明性的客观度量

相关失真度量	
归一化互相关	$NC = \sum_{x,y} p_{x,y} \tilde{p}_{x,y} / \sum_{x,y} p_{x,y}^{2}$
相关质量	$CQ = \sum_{x,y} p_{x,y} \tilde{p}_{x,y} / \sum_{x,y} p_{x,y}$

水印透明性的客观度量

其它	
全局西格马信噪比	$GSSNR = \sum_b \sigma_b^{\,2} / \sum_b \left(\sigma_b - \tilde{\sigma}_b\right)^2$ 其中， $\sigma_b = \sqrt{\frac{1}{n}\sum_{\text{块}b} p_{x,y}^2 - \left(\frac{1}{n}\sum_{\text{块}b} p_{x,y}\right)^2}$
直方图相似性	$HS = \sum_{c=0}^{255} \left\| f_I(c) - f_{\tilde{I}}(c) \right\|$

水印透明性的客观度量

1. NMSE（Normalised MSE,标准均方误差）

$$NMSE = \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)-R(i,j)]^2 / \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)]^2$$

2. SC （Structural content,结构化内容）

$$SC = \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)]^2 / \sum_{i=1}^{M}\sum_{j=1}^{N}[R(i,j)]^2$$

3. AAD （Absolute average difference 绝对平均差）

$$AAD = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|I(i,j)-R(i,j)|$$

水印透明性的客观度量

4. CQ（Correlation quality,相关性）

$$CQ = \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)\ R(i,j)] / \sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j)$$

5. NCC (Normalised cross-correlation,归一化相关系数)

$$NCC = \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)\ R(i,j)] / \sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j)]^2$$

6. PMSE （Peak MSE,峰值均方误差）

$$PMSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j) - R(i,j)]^2 / 255^2$$

水印透明性的客观度量

7. IF (Image fidelity,图像保真度)

$$IF = 1 - \left(\sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j) - R(i,j)]^2 / \sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j)]^2 \right)$$

8. PSNR （Peak Signal Noise Ratio,峰值信噪比）

$$PSNR = 10 \log_{10} \left[M \times N \times 255^2 \Big/ \sum_{i=1}^{M} \sum_{j=1}^{N} (I(i,j) - R(i,j))^2 \right]$$

数字水印的分类

- 从作品类型上分类
- 从透明性上分类
- 从嵌入方式上分类
- 从检测方法上分类
- 从稳健性上分类

从作品类型上分类

- 图像水印
  - 图像是使用最多的一种多媒体数据，也是经常引起版权纠纷的一类载体。
    - 彩色/灰度图像，卡通，设计图，二值图像（徽标、文字），等
- 视频水印
  - 保护视频产品和节目制作者的合法利益。
- 音频水印
  - 保护MP3、CD、广播电台的节目内容等 。

从作品类型上分类

- 软件水印
  - 是镶嵌在软件中的一些模块或数据，通过它们证明该软件的版权所有者和合法使用者等信息。
  - 软件水印分为静态水印和动态水印两类
    - 静态水印：不依赖于软件的运行状态，可以在软件编制时或编制完成后被直接加入。
    - 动态水印：依赖于软件的运行状态，通常时在一类特殊的输入下才会产生，水印的验证也是在特定的时机下才能完成 。
- 文档水印
  - 确定文档数据的所有者。

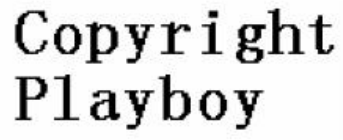
从透明性上分类

- 可见水印（可察觉水印）
  - 如电视节目上的半透明标识，其目的在于明确标识版权，防止非法的使用，虽然降低了资料的商业价值，却无损于所有者的使用。
- 不可见水印（不可察觉水印）
  - 水印在视觉上不可见，目的是为了将来起诉非法使用者。不可见水印往往用在商业用的高质量图像上，而且往往配合数据解密技术一同使用。

从透明性上分类

- 案例：可见和不可见水印

(a)原始图像



(b)水印信息

(c)以不可见方式嵌入水印之后的图像

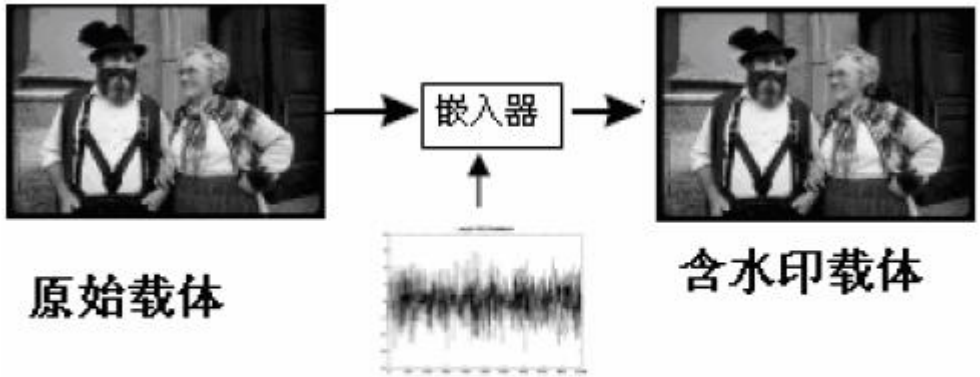
(d)以可见方式嵌入水印之后的图像

从嵌入方式上分类

- 空间域水印
  - LSB方法
  - 拼凑方法
  - 文档结构微调方法
- 变换域水印
  - DCT变换，小波变换，傅立叶变换，Fourier-Mellin变换或其它变换

从嵌入方式上分类

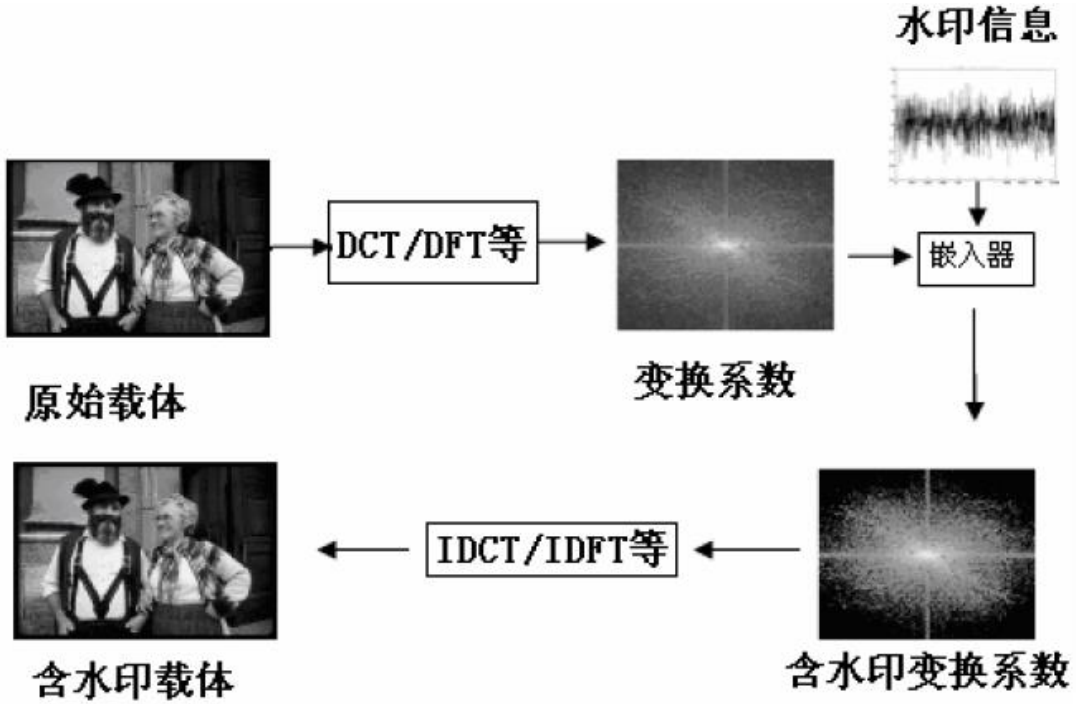
- 案例：空域水印示意



（a）时空域水印嵌入算法

从嵌入方式上分类

- 案例：变换域水印示意



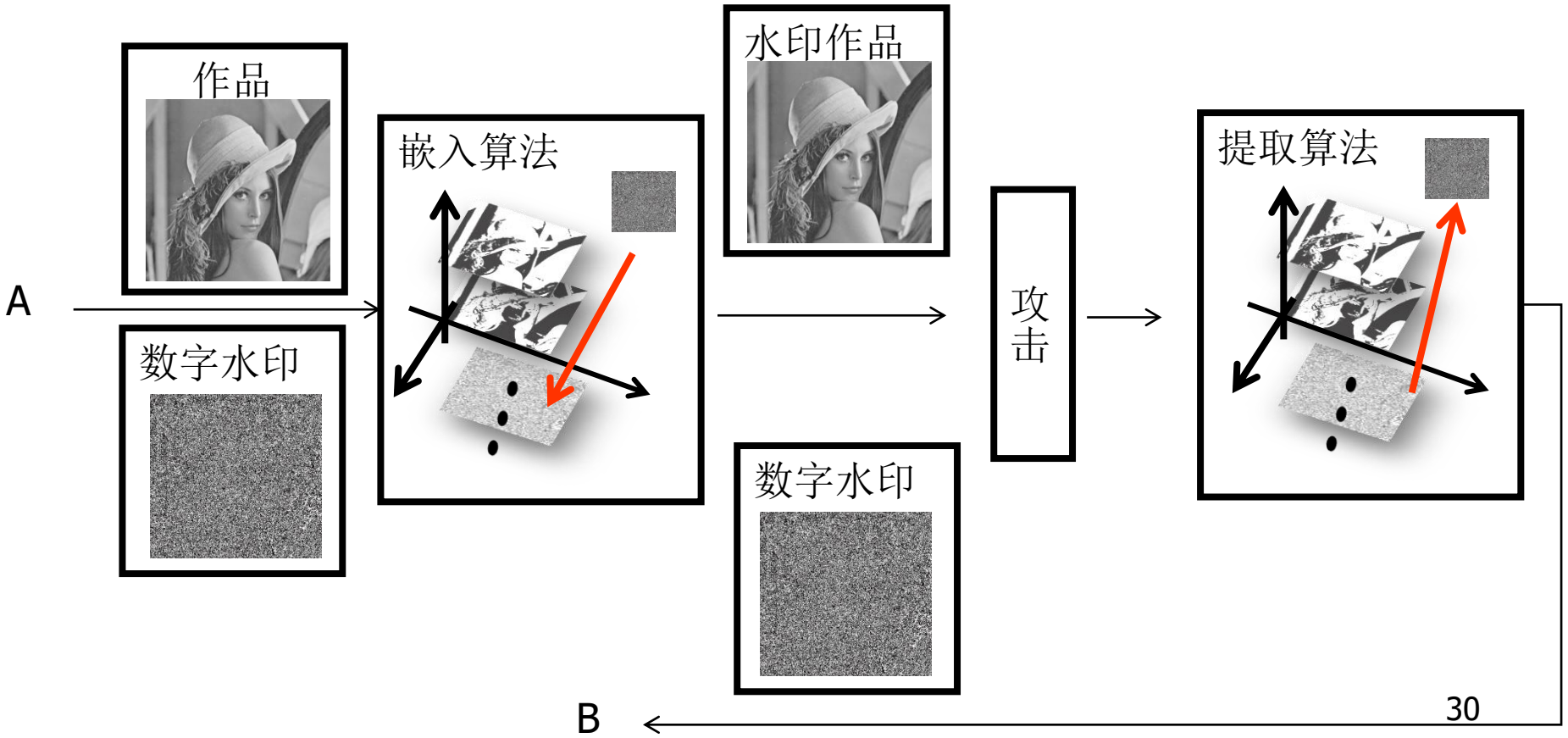
（b）变换域水印嵌入算法

从检测方法上分类

- 非盲水印和盲水印
  - 非盲水印（私有水印）：水印检测时需要原始载体
  - 盲水印（公开水印）：水印检测时无需原始载体
- 私钥水印和公钥水印
  - 私钥水印：水印加载和检测使用同一密钥
  - 公钥水印：水印加载和检测使用不同的密钥（同密码学中的公钥密码）

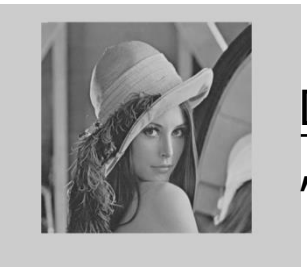
从检测方法上分类

- 案例：基于LSB的盲水印系统



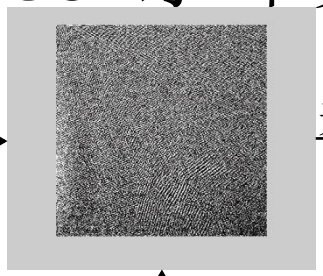
从检测方法上分类

- 基于Cox水印算法非盲水印系统



水印载体

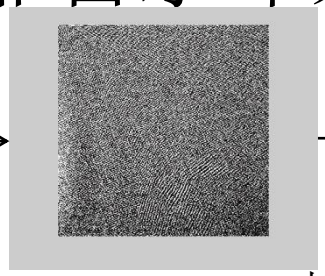
DCT(DFT ,DWT)



关键系数位置

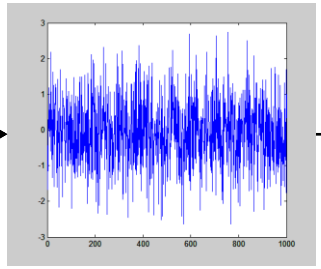
$p_i = \underset{i}{argmax}(c_i^c)$

关键系数



$c_i^w = (1+\alpha)c_i^c$

恢复水印

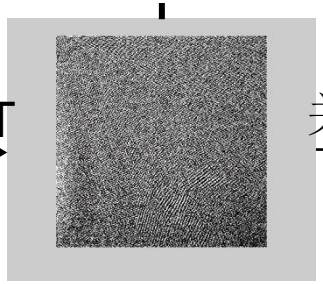


判决结果

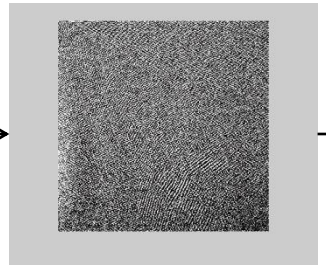
根据提取所得水印与原始水印相关性判决

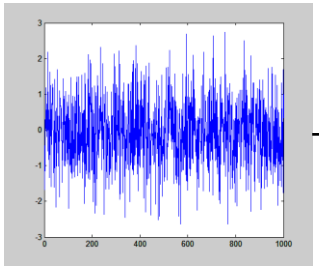
原始载体

DCT(DFT ,DWT)



关键系数





原始水印

从稳健性上分类

- 健壮性数字水印
  - 要求水印能够经受各种常用的操作。
  - 只要载体信号没有被破坏到不可使用的程度，都应该能够检测出水印信息。
- 脆弱性数字水印（完全脆弱性/半脆弱性）
  - 要求水印对载体的变化很敏感，根据水印的状态来判断数据是否被篡改过
  - 特点：载体数据经过很微小的处理后，水印就会被改变或毁掉。
  - 主要用于完整性保护。
  - 与稳健性水印的要求相反。

健壮性水印例

放大和旋转后的含水印图像
（size=280×280, θ =2度）

提取的水印

健壮性水印例

原始图像

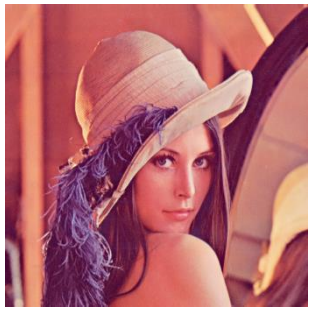
嵌入水印后图像

打印扫描后的图像

- 水印仍能正确提取

半脆弱水印例

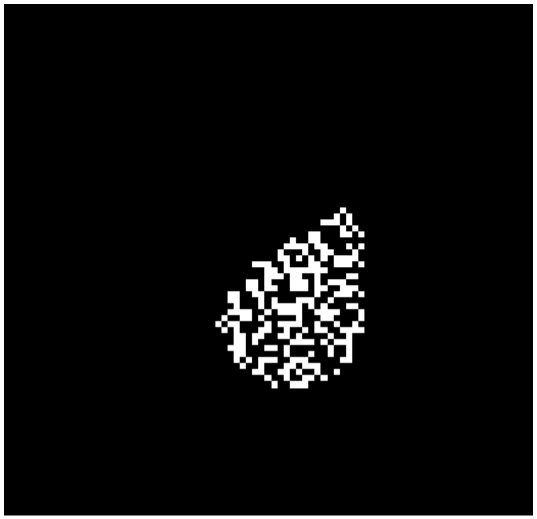
原始图像

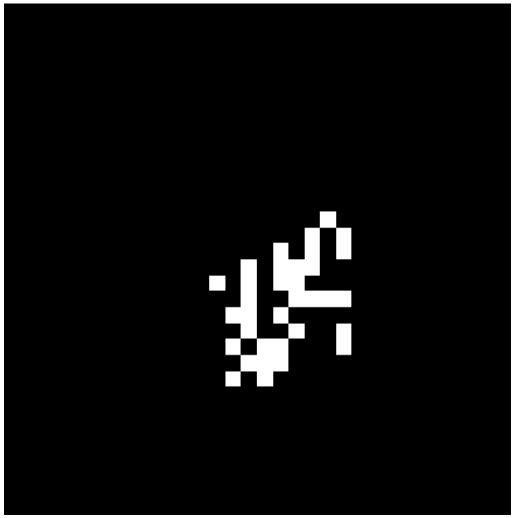


含水印图像

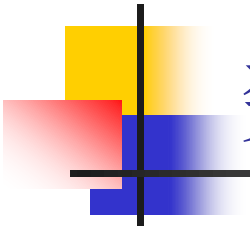
受攻击水印图像

篡改检测
第一层





篡改检测
第二层



数字水印的性能评价

- 水印容量、可感知性、健壮性三者之间的平衡

隐写术与数字水印的区别

	隐写术	数字水印
用途	用于保密通信	用于版权标识
前提	一般不知有信息隐藏	可以公布有水印存在
主要攻击	隐写分析	对水印的非授权删除
主要考核	透明性	稳健性
安全性含义	不显著改变载体对象统计特性	能够抵抗恶意攻击
载体（作品）和信息	1、信息与载体无关 2、载体可选	1、水印与作品相关 2、作品不可选

基础练习题

- 1、请简单介绍数字水印的构成方式。
- 2、请简介数字水印算法性能指标有哪些？水印算法安全性与隐写算法安全性有何区别？
- 3、水印算法可分为可见和不可见水印算法，这种分类方式是：（ ）

A、根据载体分类； B、根据嵌入方式分类；
C、根据透明性分类；D、根据鲁棒性分类

基础练习题

- 4、下面类别，哪一个不是根据稳健性分类：（）

A、脆弱水印； B、半脆弱水印； C、变换域水印；D、鲁棒水印

- 5、关于盲水印，下列说法正确的是：

A、盲水印指嵌入水印后，载体中的水印是不可见的。

B、盲水印指提取水印时，算法需要使用原始载体。

C、盲水印指嵌入水印后，载体中的水印是可见的。

D、盲水印指提取水印时，算法不需要使用原始载体。

附录

基础练习题解答

基础练习题

- 1、请简单介绍数字水印的构成方式。
  - 答：数字水印可以是有意义信息，例如版权声明文本，机构、组织或公司的Logo。还可以是无意义信息，例如随机数序列。

基础练习题

- 2、请简介数字水印算法性能指标有哪些？水印算法安全性与隐写算法安全性有何区别？
  - 答：数字水印指标有安全性、稳健性、透明性、容量和计算量等。
  - 隐写算法安全性描述算法对载体统计特征的影响程度，安全的隐写算法不显著改变载体统计特征。
  - 数字水印算法安全性描述算法抵抗恶意攻击的能力，安全的水印算法能抵抗恶意攻击。

基础练习题

- 3、水印算法可分为可见和不可见水印算法，这种分类方式是：（ ）

A、根据载体分类； B、根据嵌入方式分类；C、根据透明性分类；D、根据鲁棒性分类

答：选（C），根据透明性，水印算法可以分为可见和不可见水印算法两类。

基础练习题

- 4、下面类别，哪一个不是根据稳健性分类：（）

A、脆弱水印； B、半脆弱水印； C、变换域水印；D、鲁棒水印

答：选(C)。根据稳健性，水印算法分为稳健，半脆弱和脆弱水印三类。

基础练习题

- 5、关于盲水印，下列说法正确的是：

A、盲水印指嵌入水印后，载体中的水印是不可见的。

B、盲水印指提取水印时，算法需要使用原始载体。

C、盲水印指嵌入水印后，载体中的水印是可见的。

D、盲水印指提取水印时，算法不需要使用原始载体。

答：选（D）。按提取水印时，是否需要原始载体，水印算法可分为盲水印和非盲水印算法两类。