数字水印系统的鲁棒性和常见的攻击

吴崇明，　王晓丹
(空军工程大学 导弹学院，陕西 三原　713800)

摘　要:数字水印作为一种新型实用的信息隐藏技术,近年来已经引起了极大的关注并得到迅速发展。针对数字水印系统的鲁棒性与可能的攻击方法进行了讨论,分析指出了对于数字水印系统的鲁棒性及其度量目前待解决的问题。

关键词:数字水印;系统;鲁棒性
中图分类号:TP391　　**文献标识码**:A　　**文章编号**:1009-3516(2002)01-0090-05

数字水印的思想来源于信息隐藏(information hiding)或更严格地称为信息伪装(steganography)[1-2],是一种十分贴近实际应用的信息隐藏技术。1990年Tanaka等[3-4]将不可感知的信息隐藏在视听数据中,1993年Tirkel等[5]正式提出了水印(watermarking)一词,到了1995至1996年数字水印已经引起了极大的关注[6]。随着计算机网络、通讯、多媒体技术的发展,如何在网络环境中提供有效的版权保护(copyright protection)和信息安全(information security)手段成为迫切需要,从而使得数字水印发展迅速。但是到目前为止仍存在许多需要研究解决的公开问题,满足实际用途的各种数字水印系统也需要开发。

鲁棒性是水印技术的一个核心问题,一个实用的水印算法应该能够抵抗各种无意或有意的攻击。

本文将针对静止图像数字水印系统的鲁棒性与可能的攻击及有关问题进行讨论,在下面的讨论中,我们将首先介绍数字水印技术的基本框架和有关的参数,然后重点讨论数字水印的鲁棒性问题和对现有水印技术的各种攻击方法,以及鲁棒性度量问题,最后指出对于数字水印系统鲁棒性及其度量目前存在的问题和今后的研究方向。

1　数字水印框架及参数

1.1　数字水印的原理/基本框架

通用的数字水印算法包括两个基本方面:水印的嵌入(embedding)和水印的检测或提取(recovery)。水印可由多种模型构成,如随机数字序列、数字标识、文本及图像等。

1)水印的嵌入:即将水印信号在原始信号上进行调制。为了能够成功地提取水印信号,算法必须对无意和有意的攻击具有鲁棒性。

设I为已知数字图像、W为水印信号、密钥为K(通常是随机数发生器的种子)。I_w为加水印后的图像。

水印的嵌入过程可定义为一种映射:$I\times K\times W\rightarrow I_w$。如果用函数形式表达,可以表示为:$I_w=E(I,KW)$,其中$E$为编码函数。

框图如图1所示:

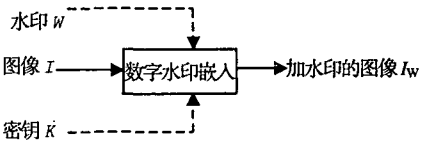

图1　数字水印嵌入框图

2)水印的检测:即从接受的信号中提取水印,或者判定水印是否存在。水印检测是水印算法中最重要的步骤。检测过程的输出为恢复的水印W或是某种置信度测量,表示已知水印在所观察图像中出现的可能性有多大。

收稿日期:2001-06-07
作者简介:吴崇明(1965-),男,江苏淮安人,讲师,硕士,主要从事计算机网络及网络安全技术研究.

框图如图 2 所示:

设 $\hat{I}_w$ 为待检测的图像,D 为解码函数,则有:

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

或

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

其中 W^* 为提取出的水印,K 为密码,函数 C 做相关检测,δ 为决策阈值。这种形式的检测函数是创建有效水印框架的一种最简便方法,如假设检验[7]、或水印相似性检验[8]。

检测器的输出结果可以作为版权保护的潜在证据,这要求水印的检测过程和算法应完全公开。对于假设检验的理论框架,可能的错误有以下两类:

第一类错误:检测到水印但水印实际不存在。这类错误用误识率 P_{fa} 衡量。

第二类错误:没有检测到水印但水印实际存在。这类错误用拒识率 P_{rej} 表示。

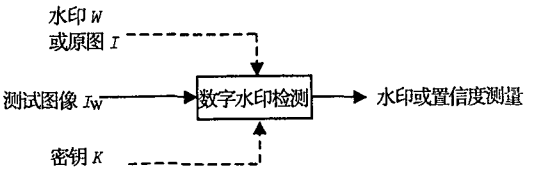

图 2　数字水印检测框图

总错误率为:$P_{err} = P_{fa} + P_{rej}$,且当 P_{rej} 变小时检测性能较好。但是检测的可靠性则只与误识率 P_{fa} 有关。

1.2　数字水印的重要参数

对于数字水印系统,需要考虑以下一些重要参数及对系统性能的影响:

1)内嵌信息量。直接影响水印的鲁棒性。欲内嵌的信息越多,水印的鲁棒性越低,即内嵌信息量的增加会引起鲁棒性的下降。需内嵌的信息量取决于应用。

2)内嵌强度(即水印的能量)。通常需在内嵌强度和图像质量之间进行折衷。鲁棒性的增加要求增大内嵌强度,而这将使图像的视觉质量严重下降。

3)图像尺寸和特性。图像尺寸正比于鲁棒性。尽管很小的图像没有很大的商业价值,但却要求水印软件能从中恢复出水印。图像的特性对水印的鲁棒性也是一个冲击。通常,对自然图像具有较高鲁棒性的方法,当应用于合成图像时,其鲁棒性会惊人地减少。一个好的数字水印系统应适用于广泛的图像尺寸范围以及不同类型的图像。

4)秘密信息(如密钥)。尽管秘密信息并不对图像的视觉精度或水印的鲁棒性构成冲击,但它在系统安全性方面起着重要的作用。密钥空间取值范围必须足够大,以使攻击耗时而不可能。

5)编解码器运算量。算法的运算量将决定算法是否实用。对于实时应用(如电视广播监控等),编解码器运算量必须足够小,以保证能够与实时任务同步。

2　数字水印系统的鲁棒性与常见的攻击

数字水印算法的鲁棒性反映水印算法经受各种攻击的能力。

鲁棒性直接依赖于内嵌强度,而内嵌强度与图像退化(即水印的显著性)相关。一个好的数字水印系统,理论上应该使得加入水印后的宿主图像具有较强的鲁棒性和最小的视觉失真。

攻击目的是想改变数据,使内嵌于其中的水印标记无法辨认,即降低检测水印的可能性。有效的水印算法必须具有鲁棒性,即数字水印必须很难被清除,或者任何企图破坏水印的操作都将导致图像质量的严重下降。

2.1　对数字水印可能的攻击

对含有水印图像的常见攻击方法分为无意(unintentional)的攻击和有意(intentional)的攻击两大类。

1)无意的攻击

无意的攻击包括那些能保持感官相似性的处理操作,如:图像压缩、滤波、图像量化与图像增强、噪声污染、扫描与复印、尺寸变化等。可以分为信号处理变换和几何变换两大类。

①信号处理变换。常见的信号处理变换包括:

JPEG 压缩

JPEG 是最常见的图像有损压缩算法之一。任何数字水印系统都必须对一定程度的压缩具有鲁棒性。

一般图像的主要能量均集中于低频分量上,压缩算法会压缩掉原图中视觉上不显著的信息,通常为图像的高频分量,而水印的不可见性要求水印要驻留于图像不显著的视觉信息中。所以经过图像压缩后,水印所在的高频分量会被当作冗余信息而清除掉。

为使水印算法对 JPEG 压缩具有鲁棒性,已有很多文献提出将水印嵌入图像的视觉最显著处(即低频分量中)或设计具有低通性的水印[8~9],虽然这可能会降低图像的质量。

平滑、锐化滤波

包括线性及非线性滤波。常用的低通滤波器有中值、高斯和均值滤波器等。图像中的水印应具有低通特性,应该无法删除图像中的水印,事实上当前很多针对水印的攻击行为都是用滤波完成的[10~11]。

锐化滤波是对某些水印算法的有效攻击,因为它在检测由水印添加所引起的高频噪声时非常有效。例如,一个基于 Laplacian 算子的攻击可以表示为:$I = I - \alpha\nabla^2(\nabla^2 I - I)$。

图像量化、亮度、对比度增强技术

一些常规的图像操作,如图像在不同灰度级上的量化、亮度与对比度的变化、直方图修正与均衡,均不应对水印的提取和检测有严重的影响。

添加噪声

水印系统应对有意添加的噪声或由图像处理操作(如打印、扫描等)而产生的噪声具有一定的鲁棒性。

②几何变换。几何变换的目的只是改变图像的外观,并不降低图像的质量。但却可能使水印变得不可检测,故水印对几何运算的鲁棒性也就非常重要。常见的几何运算有:

剪切:剪切图像的一部分,将使水印不能分布和复制到整个图像,从而使检测失败。

旋转:它会使图像的水平特征重新排列。小角度的旋转常与剪切相结合,它通常不会改变图像的商业价值,但可能降低水印的可检测性。

尺度变换:在对打印图像进行扫描、或将高清晰度图像作为网络出版等用途时会遇到尺度变换的情况。尺度变换通常分为均匀和非均匀两种,均匀尺度变换其水平和垂直方向尺度的变换因子相同,而非均匀尺度变换其水平和垂直方向具有不同的尺度变换因子。常见的水印算法大多仅对于均匀尺度变换具有鲁棒性。

删除几行或几列:删除几行或几列是对某些版权标记系统的一种基本的攻击。它对任何在空间域中直接使用扩频技术的水印算法是一种非常有效的攻击。

广义上的几何变换通常是尺度变换、旋转、剪切等的结合。尺度变换、旋转等几何变换与 JPEG 压缩结合进行也是常见的攻击。很多水印算法对上述这些几何操作都非常脆弱,容易被去掉[11],因此研究水印对几何失真的鲁棒性是人们所关注的[12]。

此外,还存在有随机非线性不可感知的几何变换等。

2)有意的攻击

以下列举出几种对水印系统进行的直接有意的攻击。

伪造水印的抽取:攻击者对于特定的产品 X 生成一个信号 W,使得水印检测算子输出一个肯定结果,而 W 是一个从来不曾嵌入产品 X 中的水印信号(攻击者将它作为自己的水印)[13]。

伪造的肯定检测:攻击者运用一定的程序找到某个密钥 K 能够使水印检测程序输出肯定结果,并用密码表明对产品的所有权。但是,在当水印本身能够以很高的确定度检测时,该攻击方法就不再可行[13]。

统计学上的水印抽取:攻击者通过统计估计的方法来去除水印图像中的水印[14]。

多重水印:攻击者并不是试图去除已有的水印,而是在水印图像中嵌入他自己的水印,从而使得攻击者和原始产品所有者都能用自己的密码检测出自己的水印。这时原始产品所有者必须在发布其产品前保存一份自己的加水印的产品,用以检测发布出去的产品是否被加了多重水印。

尽管到目前为止提出的水印算法可以分别抵抗一些基本的图像操作,但还没有哪一种水印算法能成功地对付上述所有可能的攻击。

2.2 鲁棒性的度量与检测

1)鲁棒性的度量与检测

为了定量描述已知水印系统对特定攻击的鲁棒性,可将攻击程度连续地增加,直到水印不再能可靠地提取。通常,在水印的鲁棒性和水印显著性之间应有一折衷。

对于水印系统的鲁棒性,应该有统一的检测水印鲁棒性的准则。现有水印算法中具有的各种各样的鲁棒性,并没有使用相同的检测准则,这不利于不同算法间的比较。

为了检测水印算法的鲁棒性,需要对加水印后的图像施加各种攻击行为。可以使用一些常用的工具产生基本的处理操作,如:旋转、尺寸变化、重采样、有损压缩等。但是,对于一些攻击的组合或随机几何失真等的产生却较复杂,且没有统一的标准。

所以,由于缺乏统一的鲁棒性准则及存在的各种攻击的多样性,因此对于各种水印算法、数字水印系统的性能评价是一个很困难的问题。必须提供一个统一的、公正的平台,该程序系统能在相同的比较条件下对数字水印系统的性能进行评价、比较。

StirMark 是目前网上出现的一个水印鲁棒性测试软件 < http://www. cl. cam. ac. uk/~fapp2/watermarking/ stirmark/ > ,它可以产生多种可能的攻击行为。给定一个添加水印后的图像,它可用不同的参数模拟一些可能的攻击。对输出图像再进行水印提取操作,进而可以判定该水印算法的鲁棒性。设计者称 StirMark 可以破坏目前绝大多数水印算法嵌入的水印或使之失败[15-16],因此可以将它作为一个测试水印算法鲁棒性的标准工具。

对于水印算法性能的测试,应该使用多个不同内容、不同类型的图像来进行。为公正地评价不同水印算法的性能,应使用同一组测试样本图像进行。目前在图像处理研究中使用较多的图像数据库,如 USC - SIPI 图像数据库 < http://sipi. usc. edu/services/database/Database > 中的一些常用图像:lenna、baboon、pepers 等,常被在水印算法和算法鲁棒性测试中使用。

2)鲁棒性与应用

对于数字水印系统的鲁棒性进行评价,另一个重要的问题是必须要结合其应用背景。

通常,水印必须能够经得起如前所述常见的信号处理变换和几何变换操作。在许多应用中,要求水印算法对所有可能的处理都具有鲁棒性是没有必要的。

例如,在电视广播监控中,所加的水印通常只需能够经得起信号发送前后的 D/A、A/D 变换、有损压缩、少量的水平或垂直平移,而不需要对不可能发生的如旋转、尺度变化、高通滤波等操作具有鲁棒性。在隐蔽通信中,如果宿主媒体的传送无须压缩,则只需要对那些试图检测水印是否出现的消极攻击具有鲁棒性,因为隐蔽通信的特点决定了应该避免水印被发现。用于简单认证的水印,其作用仅是表明媒体是否被改动过,水印应是易碎的。而对于那些在水印的嵌入和检测出期间可能会受到不可预见的一些处理或攻击的应用,如版权保护、数据认证、拷贝控制等,水印算法必须对各种可能发生的试图去除水印或使得水印无法检测到的主动攻击具有鲁棒性。

所以,应用领域不同,对于数字水印鲁棒性的要求也不同。使用一个标准来评价应用于不同领域的水印算法是不恰当的。因此,针对不同的应用,水印算法鲁棒性的衡量就应采用不同的准则。

3　总结

本文着重讨论了数字水印系统的鲁棒性与常见的攻击。

鲁棒性是水印技术的一个核心问题,如何设计能抵抗各种攻击的水印算法仍是一个急待解决的问题。针对不同的应用,衡量鲁棒性准则的确定也是一个急需解决的问题。同时,衡量这些鲁棒性所用的测试图像也因数字水印系统的不同而不同。

数字水印技术作为一种新兴的应用技术,一个尚未完善的学科领域,对研究者提出了严峻的挑战。有关水印鲁棒性的研究将是今后较长一段时间内数字水印技术的研究重点。

参考文献:

[1] Johnson N F, Jajodia S. Exploring steganography: seeing the unseen[J]. IEEE Computer, 1998, 31(2): 26 - 34.

[2] Aderson R J, Nakamura Y, Matsui k. The steganographic file system[A]. Proceeding of Information Hiding98[C]. 1998. 73 - 82

[3] Tanaka K, Nakamura Y, Matsui k. Embedding secret information into a dithered multilevel image[A]. Proceeding of 1990 IEEE Military Commun. Conf. [C]. 1990. 216 - 220.

[4] Tanaka K, Nakamura Y, Matsui k. Embedding the attribute information into a dithered image[J]. Syst. Comput. 1990, 21(7): 79 - 87.

[5] Trikel Ramkin G. Electronic water mark[A]. Proceeding of DICTA [C]. 1993. 666 - 672.
[6] Hartung F, Kutter M. Multimedia watermarking Techniques[J]. Proceedings of the IEEE, 1998,87(7): 1079 - 1107.
[7] Nikolaidis N, Pitas I. Robust image watermarking in the spatial domain[J]. Signal Processing, 1998,66(3):385 - 403.
[8] Cox I J, Killian J, Leighton F T, et al. Secure spread spectrum watermarking for multimedia[J]. IEEE Trans on Image Processing, 1997,6(12):1673 - 1687.
[9] Nikolaidis N, Pitas I. Copyright protection of images using robust digital signatures [A]. Proceeding of ICASSP'96[C]. 1996. 2168 - 2171.
[10] Barnett R, Pearson D E. Frequency mode LR attack operator for digitally watermarked images[J]. Electronics Letter, 1998,34(19):1837 - 1839.
[11] Cox I J. Public watermarks and resistance to tampering[A]. Proceeding of ICIP97[C]. 1997. 26 - 29.
[12] Oruanaidh J, Pun T. Rotation, scale and translation invariant spread spectrum digital image watermarking[J]. Signal Processing, 1998,66(3):303 - 317.
[13] Craver S, Zhu B, Tewfik A. Resolving rightful ownership with invisible watermarking techniques: limitations, attacks, and implications[J]. IEEE Journal on Selected Areas in Communications, 1998,16(4):573 - 586.
[14] Swanson M D. Robust audio watermarking using perceptual masking[J]. Signal Processing, 1998, 66(3):337 - 355.
[15] Kutter M, Petitcolas F. A fair benchmark for image watermarking systems[A]. Proceeding of Electronic Imaging'99 Security and Watermarking of Multimedia Contents[C]. 1999. 3657 - 3671.
[16] Petitcolas F, Anderson R. Attacks on copyright marking systems[A]. Proceeding of Information Hiding 98[C]. 1998. 218 - 238.

(编辑:田新华)

Robust of Watermarking System and the Possible Attacks

WU Chong - ming, WANG Xiao - dan
(The Missile Institute, Air Force Engineering University , Sanyuan 713800, China)

Abstract: Watermarking is a new practical information hiding technique, and has grown rapidly. The robust of watermarked image , the possible attacks and the measure of robust are mainly discussed in this paper. The existing problems are pointed out in the end of this paper.

Key words: digital watermarking; robust; system

简 讯

《空军工程大学学报》在中国科学引文数据库 2000 年的期刊评价与调整中被列为中国科学引文数据库扩展库的来源期刊。

中国科学引文数据库是在国家自然科学基金委员会和中国科学院共同资助下建成的一个大型综合性的多功能期刊引文数据库。它以国内出版的数、理、化、天、地、生、农林、医药卫生、工程技术等领域的核心期刊和优秀期刊作为来源期刊。

中国科学引文数据库先后被中国科学院院士主席团、国家自然科学基金委员会、国家重点实验室办公室、国家青年科学家奖组委会等重要的管理部门指定为查询库。