变换域数字水印技术的研究

Research on Transform Domains Digital Watermarking

王慧琴[1,2] 王志雄[1] 李人厚[1]

(西安交通大学系统工程研究所)[1] (西安建筑科技大学信控学院)[2]

Abstract Watermarking is a new technique for information security. We give a relative unambiguous definition of digital watermark and propose a general watermarking framework (GWF) defined by the seven-tuple. We discuss the current primary watermarking mechanism: transform domain watermarking. The key aspects of watermarking involving types of attacks and requirements of watermarking are presented. At last, we also discuss the unsolved problems about watermarking schemes and give the ideas of potential solutions.

Keywords Watermarking, Copyright protection, Steganography, GWF

随着信息技术和计算机网络的飞速发展,数字信息的存取和共享更加快捷、方便,同时也使侵权和非法盗版等行为更加难以觉察和认证[1],盗用者通过非法手段获取网络中的传输数据、修改数据内容、生产和再传输复制品等,将会给被盗用者造成巨大的经济损失。因此,保护图像、文本、音频和视频等数字媒体的知识产权就变得十分迫切。目前对多媒体数字版权保护的研究,主要集中在基于密码学、数字签名、数字水印的技术和方法,其中近几年刚刚发展起来的数字水印技术,被认为是一种极具潜力的数字版权保护方法[2-5]。数字水印技术主要解决数字媒体版权保护中设置识别标记的问题,其基本思想是将含有作者电子签名、日期、商标、使用权限等的数字信息作为水印信号,嵌入到图像、文本、视频和音频等数字媒体中,并且在需要时,能够通过一定的技术检测手段抽取出水印,以此作为判断媒体的版权归属和跟踪起诉非法侵权的证据。目前对该领域的数字水印技术的研究,主要集中在空间域和变换域两个方面。从综合性能分析,空间域数字水印方法应用领域较窄,对一些攻击的抵抗性较差。变换域的数字水印方法对有损压缩和其他的信号处理具有较强的免疫力,因此更具优越性,目前占据了主要地位。

1 数字水印的关键技术

1.1 概念

目前对数字水印没有统一规范的定义,作者通过对已有的水印研究算法和方案[1-5]的分析认为:数字水印是指利用一定的算法,在多媒体数据中嵌入具有不可知觉性、鲁棒性、抗检测性的,含有认证敏感信息的数字编码的技术,数字水印按其特性分为可见水印、不可见——脆弱水印、不可见——鲁棒水印;按水印内容可划分为有意义水印和无意义水印。大多数的数字水印算法都非常类似,主要包括水印信号的设计、水印嵌入和水印提取三个过程。

1.2 数字水印信号的设计

数字水印信号通常设计为使用各种概率密度函数(pdf)生成的伪随机信号。概率密度函数可以是高斯型、单极型、双极型的。

一般情况使用几种典型的伪随机序列来设计数字水印信号,如由移位寄存器产生的M序列。这种序列具有周期性和规律性,可以人为地产生和复制。由于移位寄存器的输出是由初始状态和反馈逻辑直接决定,那么任取一段输出是不可能得到其他的输出。

1.3 数字水印的嵌入和提取

水印嵌入过程就是将水印信号叠加或自适应叠加在图像的灰度(亮度)或者色彩信息上,其过程可发生在空间域或者变换域上,常用的变换域有离散傅立叶变换(DFT)域、离散余弦变换(DCT)域、分块DCT域、小波变换(DWT)域、分形(the fractal)域等。

通常利用信号的相关性实现水印的提取,如相关接收器或匹配过滤器等。在水印验证过程中,通过计算检测到的水印信号与已知水印信号的相似度,可以判

王慧琴 博士生,**王志雄** 博士生,主要研究方向为网络安全、图像处理、数字水印、智能控制理论与方法等。**李人厚** 教授,博士生导师,主要研究方向为智能控制理论与方法、网络安全、数字水印、多媒体CSCW应用、智能制造系统调度等。

断数据中是否含有已知的水印信号。另外，为了进一步提高安全性，通常要将数字水印算法与加密/解密算法相结合，利用密码和密钥来禁止水印的非法提取。这样，即使非授权用户可以提取出水印，但是在没有密钥的情况下，也无法读出水印信息。

数字水印的嵌入和提取过程分别如图1和图2所示。

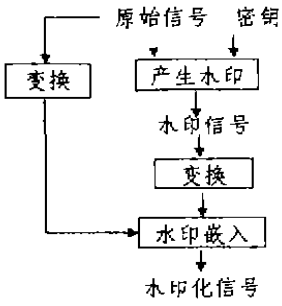

图1 水印的嵌入过程

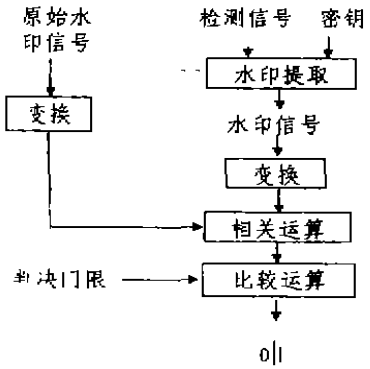

图2 水印的提取过程

首先定义嵌入多媒体作品中的数字水印信号 W 为：

$$W=\{w(k)\mid w(k)\in U,k\in W^d\}$$

其中 $U=\{0,1\}$ 或者 $U=\{-1,1\}$，W^d 中的 d 的值可取1、2或3，分别对应于音频、图像和视频三种情况下的水印域。在此基础上我们可以用一个七元组 (X,W,W^*,K,g,t,D) 定义水印的整体框架(GWF)，其中 X 表示需要被保护的数字作品 X 的集合；W 表示原始水印信号集合；K 表示水印密钥集合；g 表示用 K 和 X 生成水印的算法：

$$g:X\times K\to W,W=g(X,K) \tag{1}$$

t 是把水印 W 嵌入数字作品 X_0 中的水印嵌入算法；$X_w=t(X_0,W)$，X_w 表示 X_0 的水印版，水印的嵌入规则具体可表示为：

$$\begin{cases}X_w=X_0+\alpha W & \text{加法规则}\\ X_w=X_0+\alpha WX_0 & \text{乘法规则}\end{cases} \tag{2}$$

变量 X 在空域或时间域指采样的强度或振幅，在变换域指变换系数的大小。参数 α 为根据不同情况而变化的比例因子，具体可由试验确定。D 表示水印探测算法，W^* 表示实际从数字作品中提取出的水印信号，这样我们可以用相关系数来衡量 W^* 和原始水印 W 之间的相关性：

$$C_r=(W^*,W)=\frac{W^*\cdot W}{\sqrt{W^*\cdot W^*}} \tag{3}$$

预先设定一个阈值 C_t，进行比较运算，如果 $C_r>C_t$，可以判定水印 W 存在于数字作品 X 中，否则不存在。此过程用二值函数表示如下：

$$D:X\times K\to\{0,1\}$$
$$D(X,W)=\begin{cases}1 & \text{如果 W 存在于 X 中}\\ 0 & \text{否则}\end{cases} \tag{4}$$

在 GWF 中，g 决定水印的唯一性、有效性及不可逆性；t 对水印的不可觉察性和鲁棒性有影响；D 决定算法的可靠性和计算效率；而水印的安全性则取决于 K。

2 数字水印算法的应用需求分析

①鲁棒性：所谓鲁棒性是指水印信号在经历多种无意或有意的信号处理后，仍能保持完整性或仍能被准确鉴别的特性。可能的信号处理过程包括信道噪声滤波、数/模与模/数转换、重采样、剪切、位移、尺度变化以及有损压缩编码等。

②知觉透明性：数字水印的嵌入不应引起数字作品的视觉/听觉质量下降，即不向原始载体数据中引入任何可知觉的附加数据。

③安全性：为了提高对非法探测和非法水印提取的对抗能力，保持水印的可靠性和完整性，数据水印系统需要利用一种或多种加密体系。含有标准版权的敏感信息以密码的形式隐藏在载体信号中，利用密钥控制水印的探测和提取，这样，只有拥有密钥的合法用户，才能实现水印的提取和恢复。

3 数字水印算法设计中需要考虑的因素

①有源或无源提取：水印的提取分有源提取和无源提取两种。如图2所示，必须利用原始数据的水印提取方法叫有源提取，反之是无源提取。有源提取比较容易实现，而且可使水印信号鲁棒性更好，但在很多情况下原始数据很难得到，或者即使能得到，也会因为容量太大(例如音频/视频信号)而无法实际利用。因此，设计有效的无源提取技术是很必要的。

②内嵌信息量(水印的位率)和内嵌的强度(水印的能量)的要求：一般情况下，嵌入到载体数据中的水印信号，是含有序列号或作者签名的编码，信息量较

少，但有些情况下，比如语音信号，需要嵌入的水印信息量较大，相应的水印算法就需要具有内嵌信息量大、强鲁棒性等特点。另外，对水印强度的要求，使水印的鲁棒性和知觉透明性成为一对矛盾的特性，鲁棒性要求增大信号的内嵌强度，而这将使图像的视觉质量下降，因此，水印算法必须折衷考虑内嵌强度和图像质量的要求。

4 常见的攻击方法和对策

与密码学类似，数字水印也是一个对抗性的研究领域，非法侵权者利用各种技术手段试图去除、削弱、伪造、篡改水印，达到非法侵权的目的。所以，必须在充分掌握水印攻击技术的基础上，才能对水印算法做更加深入的研究和进行安全性测试[17]，目前常见的攻击方法有：

①主动攻击：也称波形攻击或噪声攻击。是指通过对整个水印化数据进行操作，试图削弱嵌入的水印信号，而不是试图识别和分离出水印信号的攻击，常见的攻击方法有：线性和非线性滤波、各种压缩算法（JPEG、MPEG）、附加噪声、位移、剪切、像素域量化、A/D、D/A 转换等。

对策：在人类视觉特性决定的最大容许范围内，增加嵌入的力度；或者采用冗余嵌入技术。两种方法都会增加水印的强度，从而抵抗主动攻击。

②同步攻击：或称毁坏提取攻击。这种方法通过破坏水印的相关性，使得原始水印不能被恢复，或使得水印提取器不能正常工作从而达到攻击的目的。常见的有几何变形攻击法，比如缩放、空间移位、旋转、剪切、拼接、像素置换、重采样等任何形式的几何变形变换。

对策：因为大多数水印提取算法需要知道嵌入水印的确切位置，所以同步攻击很难防御。目前有效的对策是在嵌入水印的同时嵌入水印参照物。那么在提取过程中，先根据水印参照物的变化获得同步攻击的变换步骤，然后应用反转变换获得水印的完整恢复。

③二义性攻击：又称反向水印攻击、假水印攻击或死锁，通过伪造原始数据或者在水印化数据中嵌入伪造水印，使版权产生二义性。

对策：为了解决水印的主权问题，可以采用数字时间戳（由可信的第三方提供的）方法。在水印信号设计中考虑单向函数的使用，可以构造出不可逆的水印，从而解决水印的版权问题。

④消除攻击：该方法通过分析水印化数字信号，或分别对原始水印信号和载体信号的估测，试图把水印化数据分解成载体信号和水印信号，然后丢弃代表版权信息的水印信号，达到盗版的目的。典型的有共谋攻击、非线性滤波操作、基于图像同步模型的压缩攻击等。

对策：为了减少消除攻击可利用的样品，应该限制提供的水印化数字作品的数量，另外，在水印信号设计中使用随机密钥进行加密也可以有效地增加消除攻击的计算复杂度，导致消除攻击不可实现。

5 数字水印算法的研究现状及发展趋势

目前对数字水印技术的研究，主要集中在空间域和变换域两个方面。

空间域方法是指通过直接改变像素的亮度或彩色光带，或在这两者之上叠加一个调制信号的方式嵌入水印信号[7,9]。常见的空间域方法有：文档结构微调法；最低有效位算法（LSB）；Patchwork 方法；纹理块映射编码（Texture Block Coding）等。

变换域数字水印方法是指将水印嵌入到多媒体数据的变换域上，一般为了降低算法的计算复杂度，变换域的选择要结合图像和视频压缩标准。常见的方法有：

①基于 DCT 变换域数字水印：这是目前针对图像研究得最多的一种数字水印方法，主要思路是：在图像的 DCT 变换域上，选择人眼感知最重要的频谱成分叠加水印信息。该方法常利用人眼视觉系统（HVS）的视觉掩蔽特性折衷水印的不可知觉性和鲁棒性要求，达到应用效果。DCT 水印算法可以有效抵御有损压缩和一些利用信号失真破坏水印的攻击方法[10]。因此，常常用于数字视频作品（如 VCD、DVCD）的版权保护上。

②基于小波变换的数字水印：小波变换将图像在独立的频带和不同空间方向上进行分解，能更好地与人类视觉特性（HVS）相匹配，建立视觉门限模型，定义视觉掩蔽函数，该方法的优点是，在满足水印信号的知觉透明性要求的前提下，可以尽可能提高其鲁棒性[11]。算法的应用领域十分宽广，因此在各类数字作品的版权保护中都可以得到应用。

③基于直接序列扩频水印算法：扩频水印方法与扩频通信类似[12]，将一般算法生成的水印信号再通过扩频调制后，叠加到原始数据上。从频域上看，水印信息分布于整个频谱上，无法通过一般的滤波手段恢复，这样可以更好地抵抗消除攻击。因此，该算法将会逐步应用于版权保护技术中，并占据重要位置。

利用人眼视觉掩蔽特性，根据不同的视觉敏感度，嵌入不同强度的自适应水印技术[13-15]，将是未来的研究重点。特别值得指出的是，利用小波变换的良好局部特性，结合人类视觉特性的自适应数字水印算法，会是一个很有前途的研究方向。

随着水印技术的深入研究和推广应用，其他领域的先进技术将会被不断引入，例如混沌理论、分形理论

等[16]结合图像编码中的各种压缩算法、音视频编码技术，将产生更加高效安全的数字水印算法。数字作品的版权保护将变得越来越安全有效。

6 数字水印技术中的难点问题和解决思路

①密钥丢失或被盗：一般情况下，属于同一所有者的所有作品中的数字水印是由同一密钥生成的，一旦该密钥丢失或被盗，盗用者就可能将所有该作品中的水印去掉，这是十分危险的。可能的解决办法是利用各种生物认证技术，如指纹、视网膜等管理密钥、甚至直接生成密钥，加强密钥的安全性，这会增加水印提取的计算复杂度。

②所有权死锁（二义性问题）：水印算法应该可以无二义性地定义被保护作品的所有权，但是盗版者可能在含有水印的作品中加入自己的水印，或伪造一个与原水印有很好相关性的水印，导致现有的技术无法判定哪一个是标明版权的最初原始水印，这样就形成版权死锁。可能的解决方法目前集中于第三方认证系统的建设上。利用一个公众认可的认证机构，提供第三方认证信息，在水印信号设计算法中使用，从而在版权验证时可以明确指出作品的原始版权水印。

③水印的标准问题：目前数字水印技术还没有一个统一的标准。从市场经济角度看，水印技术标准化意味着相应产品的垄断。1998年，美国版权保护技术组织（CPTWG）成立了数据隐藏小组（DHSG），着手制定版权保护水印的技术标准[18]。知识产权是一个敏感的问题，只有深入开展水印技术的研究，尽快制定我国的版权保护水印标准，才能使我们在未来可能的国际知识产权纠纷中取得主动权。

结束语 数字水印技术是一种横跨信号处理、数字通信、密码学、计算机网络等多学科的全新的信息安全技术，它在多媒体版权保护中的重要作用，必将使之成为信息媒体能够顺应数字化潮流的重要基础和保障[19]。随着全球网络化的发展，各种数字信息内容流通量的不断增加，数字水印技术的应用也必将进一步普及，并最终形成一门颇具特色的独立技术学科。

参 考 文 献

1 Hartung F, Kutter M. Multimedia Watermarking Techniques. Proc of IEEE, 1999, 87(7)

2 Schneck P B. Persistent Access Control to Prevent Piracy of Digital Information. Proceedings of IEEE, 1999, 87(7)

3 Kundur D, Hatzinakos D. Digital Watermarking for Telltale Tamper Proofing and Authentication. Proceeding of IEEE, 1999, 87(7)

4 Voyatzis G, Pitas I. The Use of Watermarks in the Protection of Digital Multimedia Products. Proceeding of IEEE, 1999, 87(7)

5 Yeung M M. Digital Watermarking. Editorial for Communications of the ACM, 1998

6 Xia X, Boncelet C, Arce G. A multiresolution watermark for digital images. In: Proc. IEEE Int Conf. Image Processing, VOL. I, Santa Barbara, CA, Oct. 1997

7 Nikolardis N, Pitas I. Robust image watermarking in the spatial domain. Signal Processing, 1998, 66(3)

8 Swanson M D. Multimedia Data-Embedding and Watermarking Technology. Proceedings of the IEEE, 1998, 86(6)

9 Brassil J T, et al. Copyright Protection for the Electronic Distribution of Text Documents. Proceedings of IEEE, 1999, 87(7)

10 Barni M, et al. A DCT-domain system for robust image watermarking. Signal Processing, 1998, 66(3)

11 Kundur D, Hatzinakos D. A Robust Digital Image Watermarking Method using Wavelet-based Fusion. ICIP 1997

12 Cox I J, et al. Secure spectrum watermarking for multimedia, IEEE Trans. Image Processing, 1997, 6

13 Podilchuk C I, Zeng W. Perceptual watermarking of still images. In: Proc. of Workshop Multimedia Signal Processing, Princeton, NJ, June 1997

14 Delaigle J F, et al. Watermarking algorithm based on a human visual model. Signal Processing, 1998, 66(3)

15 Tao, Dickinson B. Adaptive watermrking in the DCT domain. In: Proc. Int. Conf. Image Processing (ICIP), Lausanne, Switzerland, Sept. 1996

16 Bas P, Chassery J-M. Using fractal code to watermark images. In: Proc. Int Conf Image Processing (ICIP), Chicago, IL, 1998, I

17 Barnett R, Pearson D. Attack Operators for Digitally Watermarked Images. IEEE Proc.-Vis Images Signal Process, 1998, 145(4)

18 Qiao L, Nahrstedt K. Watermarking schemes and protocols for protecting rightful ownership and customer's right. J. Visual Commun. Image Representation, 1998, 9(3)

19 Ruanaidh J J K O, Dowling W J, Boland F M. Phase watermarking of digital images. In: Proc. Int. Conf. Image Processing (ICIP), 1996, 3(Sept.)

20 Ruannidh J J K O, Pun T. Rotation, scale and translation invariant digital image watermarking. In: Proc. IEEE Int. Conf Image Processing, 1997

21 Cox I J, et al. Watermarking as Communications with Side Information. Proceedings of IEEE, 1999, 87(7)