基于图像内容的二次水印模型及其实现

高智勇
(武汉大学，武汉 430079)

摘要：提出了一种基于图像内容的二次水印的理论模型。通过提取能够唯一标识原始图像的信息摘要 MD5，将其和图像水印依次嵌入到原始图像中，使得嵌入的水印具有与原图相关和不相关的双重意义；通过有实际内容的图像水印来实现水印系统的鲁棒性；通过 MD5 与原始图像的关联性来抵抗解释攻击。实验结果表明，提出的方法简单可行，内嵌数据量小，与图像文件关联性强，具有较强的鲁棒性。

关键词：数字水印；解释攻击；MD5(消息摘要算法第五版)；二次水印
中图分类号：TS853^{+}.6；TS865 **文献标识码**：A **文章编号**：1001-3563(2011)21-0094-04

Research on the Theory Model of Quadratic Watermarking Based on Image

GAO Zhi-yong
(Wuhan University, Wuhan 430079, China)

Abstract: A quadratic watermark theoretic model based on image content was proposed. The model extracted MD5 value firstly, which is the unique identification of information abstract of the origin image. Then the MD5 value was embedded into the origin image together with image watermark, which makes the embedded watermark include both contents relevant and irrelevant to the content of the origin image. The robustness of the watermark system was realized by using image watermark of practical content; the explanation attack was resisted with the association of MD5 value and the origin image. Experiments results showed that the method proposed is simple, feasible, low embedded data flow, high relevancy to the image file, and high robustness.

Key words: digital watermark; interpretation attack, MD5; quadratic watermark

互联网时代的多媒体知识产权作品所面临的安全性问题，逐渐成为制约信息化的瓶颈。单纯借助密码学的信息安全机制不能解决所有问题——多媒体内容一旦成功解密，便失去了任何保护，因此，研究在数字化和信息化环境下对各种数字作品，特别是对多媒体作品实施保护的技术，就成了水印研究的最初动力。

目前，大多数数字水印技术都是基于水印的各种物理攻击而进行研究。实际上，针对数字水印的攻击方式除了鲁棒性攻击以外，还有解释攻击和表示攻击[1]。解释攻击的实施是指攻击者只需在同一个嵌入了水印的图像中再嵌入另一个水印，使得该水印具有与所有者嵌入水印相同的强度，这使一个图像中出现了 2 个水印，从而导致了所有权的争议；然而，防范解释攻击或者当解释攻击已经实施后再来进行自我验证的过程则相当复杂。操作起来简单易行的水印攻击方式，长期以来都是图像版权维护领域的一个亟待解决的难题。

1 数字水印的研究现状

目前一些学者着力于研究解释攻击的攻击过程和原理，以及一些非对称的数字水印模型，想以此来抵抗解释攻击。Cox I J[2]等人提出在数学上非常容易找到的不对称变换，应用于数字水印却并不可行；龚理等人[3]提出在载体层到中间层的底层映射中实现数字水印鲁棒性，在中间层到应用层的上层映射中实现不对称性的分层数字水印模型，在理论上很好地解决了非对称性数字水印算法十几种并不存在的技术性难题。

基于分层理论以及图像内容的水印与原图之间的相关性，笔者提出一种基于图像内容的二次水印模

收稿日期：2011-08-26

作者简介：高智勇(1971－)，男，辽宁人，硕士，武汉大学讲师，主要从事印刷包装工程的教学与研究。

型,通过构造一个原图特征信息,并使其具有继承性和可检测性,然后将其与鲁棒性图像水印一同嵌入到原图中,使得水印的鲁棒性放在了水印的不相关层,将水印的抗解释攻击性集成在它的二次分层结构和与原图的相关性中,实现上简单可行,且内嵌数据流很小,在检测或查询版权归属时并不需要建立并搜索庞大的数据库,在实际操作中具有可行性。

2　基于图像内容的二次水印模型

解释攻击的实质是依照原始作品、原始水印和发布水印作品之间的关系,通过发布的水印作品构造一个新的伪造作品、伪造水印和发布水印作品之间的关系[4]。在设计针对解释攻击的解决方案时需要寻求一种图像特征信息能够同时具有继承性和唯一识别性,笔者试图寻找一种能够对图像进行唯一标识的特征信息,并将其人为的添加到发布的水印作品中,由于其信息量小,这种特征信息在原图中可以重复嵌入很多次,只要嵌入位置合适,在攻击者攻击原图但是一定保留原图使用价值的前提下,攻击者的攻击不可能破坏每个嵌入的特征信息,从而使得图像特征信息便具有了可检测的继承性,又因为这一信息能够唯一标识原图,因此这一实验模型能够抵抗解释攻击。

2.1　二次水印嵌入

设 I 代表原始图像,大小为 $M\times N$,J 代表水印图像,大小为 $P\times Q$,其中 M 和 N 分别是 P 和 Q 的偶数倍,将水印 J 嵌入图像 I,具体算法分为下列 4 步。

1) 将 I 分解成 $(M/8)\times(N/8)$ 个尺寸为 8×8 的方块 B;同时,将 J 也分解为 $(M/8)\times(N/8)$ 个尺寸为 $(8P/M)\times(8Q/N)$ 的方块 V。

2) 对每一个方块 B 做 DCT 变换处理,即:$DB=DCT(B)$。

3) 对每个处理过的方块 DB 和方块 V 嵌入水印,$s(i)$为从 DB 的中频部分甄选出的水印嵌入位置,$1\leqslant i\leqslant(8P)/M\times(8Q/N)$;$t(i)$代表水印 V 的位置坐标,$1\leqslant i\leqslant(8P)/M\times(8Q/N)$,$DB'(s)=A\times V$,其中 A 代表加权的系数,用 $DB'(s)$来代替 DB,就可获得嵌入了水印的图像 DBC。

4) 对上述获得的每一个分块水印图像 进行 $IDCT$ 变换,即 $IDBC=IDCT(DBC)$,然后将各个子块合并成一个图 I',最终得到了嵌入水印后的新图像,流程见图 1。

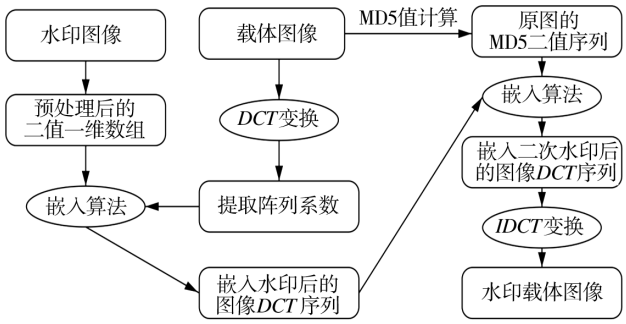


图 1　基于图像内容的二次水印的嵌入过程

Fig. 1 The embedding process of quadratic watermarking based on image content

2.2　MD5 值计算

MD5(Message Digest Algorithm 5)是计算机安全领域广泛使用的一种散列函数,用以提供消息的完整性保护,下面将分别对 MD5 具体计算方法进行介绍。

假设输入的数据是一个字节串,其中每个字节用 8 个比特表示。生成 MD5 的具体过程如下。

1) 第 1 步,补位。

计算 MD5 时首先要对参与运算数据进行补位,以使输入的数据长度符合一定要求,即除以 64 的余数是 56,也就是数据长度扩展到 $\mathrm{LEN}=K\cdot64+56$ 个字节,其中 K 属于整数。进行补位的具体方法是:补一个 1,然后其余补 0,直到满足上面提到的要求。在这一步里一共补充的字节数根据输入数据的不同限制在 0～63 之间。

2) 第 2 步,附加输入数据的长度信息。

将数据的原始长度表示成一个长度为 64 的整数(共 8 个字节),并将这 8 个字节的数据长度信息按照低位在前、高位在后的方式添加到经过补位处理的数据后面。此时,再次添加信息后的数据长度变为:$\mathrm{LEN}=K\cdot64+56+8=(K+1)\cdot64$。

3) 第 3 步,对 MD5 参数进行初始化。

在这一步中,通过 4 个长度为 32 的整数变量(A,B,C,D)来计算输入数据的信息摘要,这 4 个变量首先被初始化为 4 个十六进制数(低位字节在前),其具体数值如下:

word A:01234567

word B:89abcdef

word C:tedcba98

word D:76543210

4) 第 4 步,定义 4 个基本的按 MD5 位操作函数。

X,Y,Z 分别为长度为 32 的整数。

$F(X,Y,Z)=(X \text{ and } Y)\text{or}(\text{not}(X)\text{and } Z)$

$G(X,Y,Z)=(X \text{ and } Z)\text{or}(Y \text{ and not}(Z))$

$H(X,Y,Z)=X \text{ or } Y \text{ or } Z$

$I(X,Y,Z)=Y \text{ or}(X \text{ or not}(Z))$

再定义 4 个依次用于 4 轮变换的函数。

预设 Mj 代表信息的第 i 个子分组($0\leqslant i\leqslant 15$), S 代表循环向左移 S 位,那么 4 种变换可表示为:

$FF(a,b,c,d,Mj,S,ti)$表示 $a=b+(a+(F(b,C,d)+Mj+ti)<<<S)$;

$GG(a,b,c,d,Mj,S,ti)$表示 $a=b+(a+(G(b,c,d)+Mj+ti)<<<S)$;

$HH(a,b,c,d,Mj,S,ti)$表示 $a=b+(a+(H(b,c,d)+Mj+ti)<<<S)$;

$II(a,b,c,d,Mj,S,ti)$表示 $a=b+(a+(I(b,C,d)+Mj+ti)<<<S)$。

5) 第 5 步,变换输入的数据。

进行数据处理,N 代表总字节数,然后每 64 个字节划分为一组,每组数据作一次循环,每次循环依次经过上文提到的 4 轮变换。进行变换操作的 64 字节数据用 16 个长度为 32 的整数数组 $M[0\cdots 15]$表示。其中数组 $T[1\cdots 64]$代表一组常数; $T[i]$代表 4294967296 · abs(sin(i))的 32 位整数部分,i 的单位为弧度,取值范围为 1～64。

6) 第 6 步,输出结果。

A,B,C,D 连续存放,大小为 16 字节,共 128 位。最后以十六进制方式依次输出这 16 个字节。

2.3 水印图像预处理－Arnold 变换

利用图像置乱技术,可以达到使原图变得杂乱无章的目的,使得置乱后的图像无法传达原图所包含的内容。通过这种方式,即使图像信息被非法劫取,盗用者也只是获得一堆乱码信息。将图像置乱技术应用到数字水印系统中,可以大大提高嵌入水印的隐蔽性,并增加整个系统的安全性。较为常用的图像置乱方法有以下几种: Arnold 变换、Fibonacci 变换、Hilbert 变换、幻方变换以及仿射变换[5]。

采用的置乱算法是 Arnold 变换。对于一个 $N\times N$ 的数字图像来说,其二维 Arnold 变换定义为:

$$\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}11\\10\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}(\text{mod}N) \tag{1}$$

式中: $x,y\in\{0,1,2,\cdots,N-1\}$代表变换前的像素位置; x',y'代表变换后的像素位置; mod 代表模运算。图像以数字信息表示,是一个二维的矩阵。

实施 Arnold 变换之后图像信息的二维矩阵会转置重排,在视觉上表征为图像信息的杂乱无章,内容上和原始图像相比发生了很大的改变,但是对于合法用户,可以利用 Arnold 变换的周期性,通过重复进行的 Arnold 变换,经过一定的次数之后将这些无意义的内容还原成有意义的原始图像[6]。

2.4 MD5 检测与提取

1) 对载体水印图像进行 DCT 变换。

2) 从载体水印图像的 DCT 阵列系数的第 1 位开始,逐位与标记信息的二值序列进行按位与异或运算,并将运算结果赋给一个变量,当变量值不为 0 时,则从 DCT 阵列系数的下一位重新开始,如此循环。当此标记比对次数的变量 num 值为 8 时,便将这一位之后的 $32\times 8=256$ 位二值序列依次赋值给一个长度相同的数组并保存,一直运算到载体水印图像的 DCT 阵列系数的最后 1 位[7]。

3) 对于每一个赋值的数组,统计每一个相同位 0,1 值出现的次数,将次数较多的那一位 0,1 值赋给一个新数组,这个数组即为最终提取的 256 位的 MD5 二值序列。

4) 将此 256 位的 MD5 二值序列转为 16 进制,获得长度为 32 位的 MD5 值,即为 MD5 值提取结果。

2.5 水印图像的提取

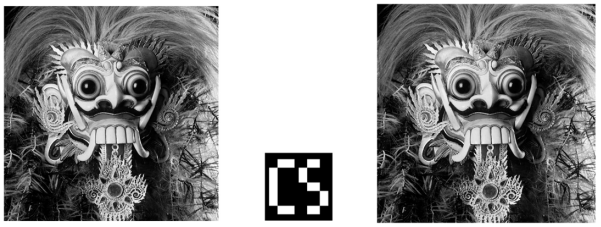
设图像 D 代表已经嵌入了水印信息的载体图像,水印信息检测过程为[8]:(1)将 D 分解为$(M/8)\times(N/8)$个尺寸为 8×8 的方块 BD;(2)对每个子块 BD 实施二维 DCT 变换, $DBD=DCT(BD)$;(3)提取 MD5 值,得到提取 MD5 值后的载体水印图像的 DCT 阵列系数;(4)对每个经过变换的分块 DBD 按照式 $V'=\frac{1}{A}\times DBD$ 得到变换后的子块图像 V';(5)把经过以上步骤所获得的所有 V',合并成一个水印整图的置乱序列图。

3 实验与分析

3.1 水印嵌入

实验用的载体图像为 512×512 的 24 位图像,见图 2a;水印图像是内容为“CS”的 128×128 的 8 位图像,见图 2b;水印嵌入后的图像见图 2c。

理论上 32 位的图像 MD5 值能够唯一标识的图



a 载体图像　b 二值水印图像　c 嵌入水印后图像

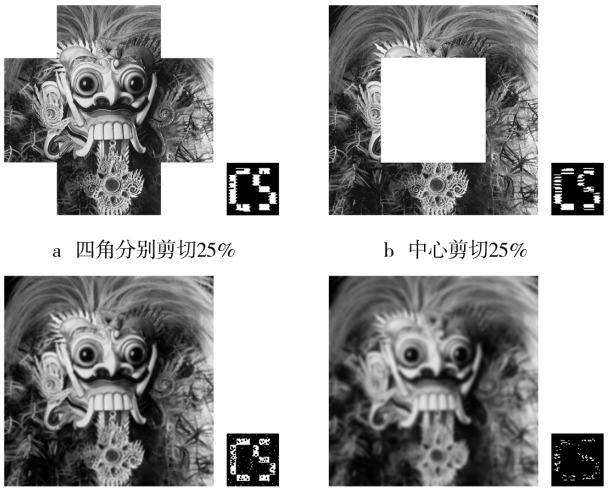
图 2　水印嵌入实验

Fig. 2 Watermark embedding experiment

像个数为 32 的阶乘，但由于数字图像产生的速度过快，在现有的研究阶段中，图像的 MD5 值最高已扩展到 40 多位。考虑到嵌入信息量的大小和系统实现时的简便，选取了 32 位的 MD5 值进行嵌入，根据水印图像的二值性选择不同的嵌入系数。

3.2　实验结果分析

水印攻击实验见图 3，水印鲁棒性分析见表 1。



a 四角分别剪切25%　b 中心剪切25%

c 高斯模糊（半径为3）　d 高斯模糊（半径为3）

图 3　水印攻击实验

Fig. 3 Watermark attack experiment

表 1　水印鲁棒性分析

Tab. 1 Analysis on robustness

攻击类型	PSNR	水印检测结果	MD5 提取/嵌入	MD5 检测结果
压缩 20%	45.314	Yes	61/64	吻合
压缩 40%	53.264	Yes	41/64	吻合
压缩 60%	57.612	Yes	32/64	吻合
旋转 15%	29.357	Yes	44/64	吻合
旋转 30%	27.843	Yes	35/65	吻合
旋转 45%	23.482	Yes	32/64	吻合
四角各剪切 25%	47.524	Yes	35/64	吻合
中心剪切 25%	56.314	Yes	34/64	吻合
高斯模糊(半径 3)	27.562	Yes	42/64	吻合
高斯模糊(半径 6)	22.223	Yes	31/64	吻合

4　结语

实验结果表明，该水印系统对于常规的物理攻击具有一定的鲁棒性，对于图像压缩和旋转的鲁棒性较强，对于剪切和高斯噪声攻击的鲁棒性较差。由于当 MD5 嵌入个数为 64 时，在各种攻击方式下，基本可以保证正确的 MD5 检测个数接近于嵌入 MD5 个数的一半，因此检测出来的 MD5 结果能与原图的 MD5 吻合，能起到抗解释攻击的效果，因此，基于图像内容的二次水印，可以使嵌入的水印具有与原图相关和不相关的双重意义，能够有效保证水印系统的鲁棒性并能较好地抵抗解释攻击，且实现上简单可行。

较之于一次水印，二次水印具有更好的抗解释攻击能力。即便是攻击者继续利用解释攻击的原理来伪造一种类似的关系，即伪造并注册一个与伪造原图相关的数据，并证明这个数据在原图中存在，但是这些相关信息所标记的原图和伪造原图的关系是相互的，因为伪造原图保留有的，原图一定保留，伪造原图信息改动的，原图一定不保留，从而使得伪造一个类似单向的从伪造原图指向原图的 MD5 特征信息，并将其嵌入到原图中，是不可能的。

参考文献：

[1] 周军. 利用双水印技术对抗解释攻击[J]. 信息技术与网络服务，2006(4)：20－21.

[2] COX I J，MILLER M L，BLOOM J A. 数字水印[M]. 王颖，译. 北京：电子工业出版社，2003.

[3] 龚理. 数字水印技术研究与应用[D]. 长沙：湖南大学，2004.

[4] 王莉，洪亮. 探析包装防伪印刷[J]. 包装工程，2006，27(5)：301－303.

[5] 詹斌. 基于互联网的数字水印技术研究[D]. 重庆：重庆大学，2009.

[6] MAKHLOUFI A，OULED Zaid A，BOUALLEGUE A. Wavelet Domain Watermark Embedding Strategy Using TTCQ Quantization[C]. 14th International Conference on Image Analysis and Processing，2007.

[7] 黎冠英. 用于图像处理的数字水印算法改进研究及仿真[D]. 上海：华东师范大学，2010.

[8] 秦峰. 彩色图像数字水印研究[D]. 大连：大连理工大学，2006.