ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.17, No.2, February 2006, pp.315−324
DOI: 10.1360/jos170315
© 2006 by *Journal of Software*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

一种抗 JPEG 压缩的半脆弱图像水印算法*

李 春[1,2], 黄继武[1,3+]

[1](中山大学 信息科技学院,广东 广州 510275)

[2](中国电子科技集团公司第七研究所,广东 广州 510310)

[3](广东省信息安全技术重点实验室,广东 广州 510275)

A Semi-Fragile Image Watermarking Resisting to JPEG

LI Chun[1,2], HUANG Ji-Wu[1,3+]

[1](School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China)

[2](China Electronics Technology Group Corporation No.7 Research Institute, Guangzhou 510310, China)

[3](Guangdong Key Laboratory of Information Security Technology, Guangzhou 510275, China)

+ Corresponding author: E-mail: isshjw@mail.sysu.edu.cn

Li C, Huang JW. A semi-fragile image watermarking resisting to JPEG. *Journal of Software*, 2006,17(2): 315−324. http://www.jos.org.cn/1000-9825/17/315.htm

Abstract: Semi-Fragile watermark has attracted attention due to its important role in content authentication for multimedia. In order to differentiate incidental attacks and malicious attacks, semi-fragile watermark must be robust against content-protection image processing. Semi-fragile watermark resisting to JPEG compression while maintaining high detection ability to tamper has been the emphasis in this area. In this paper, a semi-fragile watermarking algorithm to resist JPEG is proposed based on the fact that most of the coefficients of high frequency have the same relative energy relations after JPEG compression. The experimental results demonstrate that the proposed algorithm has the advantages such as simple computation complexity, big embedding capability, good robustness to JPEG, and exact location for tamper.

Key words: semi-fragile watermarking; digital watermarking; JPEG; wavelet coefficient

摘 要: 半脆弱水印因为在多媒体内容认证方面的重要作用而受到人们密切的关注.为了能够区分偶然攻击与恶意篡改,半脆弱水印需要对一般的内容保护图像操作有一定的鲁棒性.由于 JPEG 压缩应用的普遍性,在保持较高的对篡改检测能力的情况下,提高抗 JPEG 压缩性能一直是半脆弱水印的重要问题.根据图像相邻小波高频系数之间的大小关系在 JPEG 压缩之后大多数没有发生变化这一事实,提出了一种新的抗 JPEG 压缩的半脆弱水印算法.实验结果表明,该算法计算简单,嵌入容量大,有很好的抗 JPEG 压缩性能,同时对篡改的定位也很精确.

关键词: 半脆弱水印;数字水印;JPEG;小波系数

中图法分类号: TP309 文献标识码: A

* Supported by the National Natural Science Foundation of China under Grant Nos.60133020, 60325208, 60403045 (国家自然科学基金); the Natural Science Foundation of Guangdong Province of China under Grant No.04205407 (广东省自然科学基金)

Received 2005-01-19; Accepted 2005-08-15

© 中国科学院软件研究所 http://www.jos.org.cn

随着多媒体技术和网络技术的飞速发展及广泛应用,对图像、音频、视频等多媒体内容的保护已成为迫切需要解决的问题.对多媒体内容的保护一般可以分为如下两个方面:一是版权保护;二是内容完整性(真实性)保护,即认证(或称为“篡改提示”).用于版权保护的数字水印要求有很强的鲁棒性和安全性,用于多媒体内容真实性鉴定(或篡改提示)的水印,一般称为易损水印或脆弱水印,这种水印同样是在内容数据中嵌入不可见的信息[1].

在实际应用中,并不需要脆弱水印对所有的修改都非常敏感.对恶意篡改高度敏感和对内容保护操作鲁棒的半脆弱水印更能适应实际应用的要求.一般地,一个半脆弱水印应该满足 3 个基本要求:对恶意篡改的高度敏感性和对内容保护操作的鲁棒性、不可见性、安全性[2,3].半脆弱水印不但要对恶意的攻击特别敏感,又要对一些常规的图像操作(如 JPEG 压缩、加噪等)有一定的鲁棒性,从而将偶然攻击与恶意篡改区分开来.

由于 JPEG 压缩应用的普遍性,抗 JPEG 压缩一直是半脆弱水印的重要研究内容及难点.Ho 与 Li 利用基于相关信号及系数大小提取的图像特征不受 JPEG 压缩影响的特点,将水印序列嵌入到量化后的 DCT 系数中,以达到抗 JPEG 压缩的目的[4].在文献[5]中,Zhang 等人根据 JPEG 压缩过程中的不变参量进行水印生成和嵌入调制,利用小波特性对图像篡改区域进行定位.Zhou 等人通过在水印序列中加入纠错编码(ECC)来恢复由于 JPEG 压缩所造成的误码[6].Hu 等人在文献[7]中提出了基于人类视觉特性 HVS 的易碎水印,并通过对水印的形态学处理,使检测结果更加准确.在文献[8]中,Hu 为了减少运算量和提高算法的安全性,提出了基于提升格式参数化整数小波变换的脆弱水印方案,但这个方案不具备抗 JPEG 压缩的功能.尽管已有不少抗 JPEG 的脆弱水印被提出来,但已有算法的抗 JPEG 性能还不能令人满意.

为了在提升格式整数小波变换的基础上提出一种有效的抗 JPEG 压缩的脆弱水印算法,我们考察了经 JPEG 压缩前后的图像小波系数.通过大量实验发现,图像高频子带相邻小波系数大小关系,在较高质量因子的 JPEG 压缩前后变化不大,并以此为根据提出了一种抗 JPEG 的半脆弱水印算法.实验结果表明,该算法在具有十分精确的篡改定位检测能力的同时,对 JPEG 压缩也有较好的鲁棒性.此外,算法计算量较小.

本文第 1 节描述高频小波系数大小关系在压缩前后的变化.第 2 节给出水印算法的嵌入、提取及对攻击类别的判断过程.第 3 节给出了算法的实验结果.最后对所提出的算法进行了总结.

1 JPEG 压缩对图像高频小波系数大小关系的影响

Lie 与 Chang 提出利用变换域小波系数的大小关系来实现在音频上的稳健水印[9].在图像水印算法方面,Lin 和 Chang 利用图像不同 DCT 系数块相应位置上系数的大小关系提出的水印算法对 JPEG 压缩有一定的鲁棒性[10].但是,目前利用 DWT 系数之间的大小关系实现图像脆弱水印抗 JPEG 压缩的算法还未见报道.为证明这一方法的可行性,下面先考察图像小波系数在 JPEG 压缩前后大小关系的特征.图 1 是本文实验所用的测试图像.

Fig.1 Images used for testing the relationship between the neighboring wavelet coefficients

图 1 用于测试 JPEG 压缩对图像相邻小波系数大小关系影响的图像(512×512)

将原始图像与经 JPEG 压缩后的图像分别进行一次参数化整数小波提升,将得到 HH_1,HL_1,LH_1,LL_1 这 4 个子带.在两幅图的 LH_1 子带分别用 1×2 的不重选窗口滑动截取小波系数,得到$[a_1\ a_2]$,计算 $d=a_1-a_2$.若原始图像的 $d>g$(g 为一预先设定的阈值,$g\geq 0$),而压缩之后图像的小波系数满足 $d\geq 0$,则视为两者的大小关系没有发生改变;同理,若原始图像小波系数之差 $d<-g$,而压缩之后的图像小波系数满足 $d<0$,则也视为大小关系没有发生变

© 中国科学院软件研究所 http://www.jos.org.cn

化.在 HL_1 子带也做同样的实验,但是窗口改成 2×1(即在纵向截取系数),得到 $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$,计算 $d=a_1-a_2$.判断大小关系是否发生变化的准则与 LH_1 子带相同.最后,用大小关系没有发生变化的系数对除以系数对的总个数,分别得到经 JPEG 压缩后,LH_1 子带与 HL_1 子带的相邻高频系数大小关系没有发生变化的系数对占系数对总数比例的多少,见表 1(其中,LH_1 子带用横向滑动窗,HL_1 子带用纵向滑动窗).若在 LH_1 子带中纵向截取系数,HL_1 子带中横向截取系数,可得到表 2 的数据(其中,LH_1 子带用纵向滑动窗,HL_1 子带用横向滑动窗).表 1 与表 2 中的小数表示大小关系没有发生改变的系数对与系数对总数的比值.

作为低频带的 LL_1 子带,理论上对于 JPEG 压缩有更好的鲁棒性.但是,由于通过修改相邻小波系数大小关系来嵌入水印的方法对图像的降质比较明显,所以,虽然水印嵌入到 LL_1 子带能够获得更好的抗 JPEG 性能,但却造成了嵌入后图像的严重失真,从而失去应用价值,故本文重点考察 HL_1 和 LH_1 子带系数之间的大小关系.

由表 1 和表 2 可以得出如下结论:

(1) 经 JPEG 压缩之后,LH_1 子带的小波系数在水平方向上大小关系没有发生变化的系数对所占的比例均高于垂直方向上的系数对,相反地,在 HL_1 子带的小波系数在垂直方向上大小关系没有发生变化的系数对的比例均高于水平方向上的系数对,这是因为图像经过小波分解之后,LH_1 子带主要表现了图像水平方向上的细节分量,而 HL_1 子带则表现了图像垂直方向上的细节分量,而 JPEG 压缩对图像的细节分量影响不大,所以使得 LH_1 子带在横向的系数大小关系对 JPEG 压缩得没有纵向系数那么敏感,HL_1 子带则相反.

(2) 经 JPEG 压缩之后,LH_1 子带小波系数在水平方向上大小关系未改变的系数对与 HL_1 子带小波系数在垂直方向上大小关系未改变的系数对相差不大.

(3) 经 JPEG 压缩之后,只要在选择相邻小波系数时所取方向合适,就可以发现绝大多数相邻小波系数的大小关系没有发生变化,且未发生变化的系数对所占比例随着阈值 q 的增加而增加.本文的水印算法将基于高频小波相邻系数间的大小关系在 JPEG 压缩之后绝大部分未发生变化这一现象来嵌入和提取水印.

Table 1 The relationship between the neighboring wavelet coefficients in high frequency subbands of the images under JPEG

表 1 JPEG 压缩对图像小波相邻高频系数大小关系的影响

Image \ Quality	q	90		80		70		60		50	
		LH_1	HL_1	LH_1	HL_1	LH_1	HL_1	LH_1	HL_1	LH_1	HL_1
Lena	0	0.786	0.770	0.718	0.711	0.679	0.666	0.651	0.632	0.636	0.613
	1	0.844	0.828	0.773	0.766	0.731	0.720	0.681	0.662	0.665	0.640
	2	0.894	0.878	0.829	0.823	0.787	0.774	0.715	0.696	0.697	0.671
	3	0.928	0.923	0.842	0.848	0.790	0.785	0.751	0.729	0.732	0.701
Baboon	0	0.916	0.928	0.864	0.897	0.827	0.868	0.798	0.838	0.771	0.821
	1	0.942	0.950	0.891	0.917	0.854	0.890	0.819	0.857	0.790	0.839
	2	0.962	0.967	0.913	0.936	0.876	0.909	0.839	0.876	0.808	0.857
	3	0.978	0.982	0.932	0.953	0.892	0.924	0.857	0.894	0.828	0.874
Boat	0	0.834	0.785	0.766	0.735	0.729	0.710	0.703	0.689	0.686	0.677
	1	0.882	0.843	0.818	0.799	0.784	0.770	0.733	0.731	0.715	0.715
	2	0.921	0.899	0.871	0.857	0.840	0.832	0.766	0.782	0.745	0.762
	3	0.949	0.943	0.879	0.892	0.836	0.861	0.802	0.838	0.779	0.812
Goldhill	0	0.862	0.867	0.794	0.824	0.747	0.791	0.716	0.767	0.689	0.749
	1	0.900	0.907	0.833	0.861	0.786	0.828	0.741	0.796	0.710	0.780
	2	0.932	0.939	0.871	0.894	0.824	0.861	0.768	0.828	0.734	0.809
	3	0.958	0.963	0.888	0.923	0.834	0.888	0.795	0.856	0.757	0.835
Lake	0	0.843	0.858	0.780	0.804	0.740	0.768	0.711	0.741	0.688	0.724
	1	0.888	0.899	0.822	0.844	0.784	0.810	0.742	0.775	0.716	0.754
	2	0.928	0.931	0.865	0.880	0.827	0.846	0.773	0.808	0.745	0.785
	3	0.956	0.963	0.892	0.911	0.841	0.868	0.800	0.836	0.776	0.811
Milkdrop	0	0.770	0.777	0.697	0.726	0.664	0.687	0.636	0.658	0.611	0.638
	1	0.829	0.838	0.754	0.781	0.713	0.744	0.669	0.693	0.638	0.671
	2	0.883	0.893	0.811	0.836	0.770	0.796	0.701	0.732	0.668	0.703
	3	0.925	0.934	0.827	0.872	0.772	0.820	0.731	0.768	0.697	0.733

© 中国科学院软件研究所 http://www.jos.org.cn

Table 2 The relationship between the neighboring wavelet coefficients in high frequency subbands of the images under JPEG after changing the slip windows

表 2 改变滑动窗后得到的 JPEG 压缩对高频相邻小波系数大小关系的影响

Image \ Quality	g	90		80		70		60		50	
		LH_1	HL_1	LH_1	HL_1	LH_1	HL_1	LH_1	HL_1	LH_1	HL_1
Lena	0	0.730	0.702	0.643	0.640	0.616	0.613	0.594	0.594	0.587	0.585
	1	0.788	0.767	0.699	0.698	0.669	0.669	0.618	0.614	0.608	0.603
	2	0.851	0.832	0.768	0.774	0.738	0.741	0.646	0.641	0.630	0.628
	3	0.876	0.845	0.753	0.751	0.705	0.702	0.675	0.671	0.654	0.652
Baboon	0	0.882	0.896	0.812	0.834	0.763	0.792	0.732	0.760	0.707	0.736
	1	0.912	0.920	0.843	0.862	0.795	0.821	0.751	0.779	0.724	0.753
	2	0.938	0.942	0.871	0.885	0.823	0.848	0.770	0.798	0.742	0.770
	3	0.959	0.961	0.887	0.898	0.830	0.852	0.789	0.816	0.759	0.786
Boat	0	0.778	0.709	0.692	0.660	0.656	0.634	0.627	0.622	0.615	0.611
	1	0.832	0.779	0.753	0.722	0.714	0.694	0.651	0.658	0.636	0.642
	2	0.884	0.846	0.815	0.791	0.770	0.754	0.678	0.697	0.660	0.678
	3	0.909	0.883	0.805	0.803	0.753	0.768	0.710	0.743	0.688	0.718
Goldhill	0	0.816	0.788	0.713	0.715	0.666	0.674	0.628	0.649	0.611	0.634
	1	0.858	0.836	0.759	0.762	0.707	0.716	0.646	0.674	0.625	0.656
	2	0.898	0.881	0.799	0.805	0.752	0.760	0.665	0.702	0.638	0.676
	3	0.925	0.913	0.805	0.819	0.734	0.766	0.684	0.729	0.654	0.701
Lake	0	0.798	0.800	0.709	0.713	0.667	0.669	0.639	0.646	0.621	0.634
	1	0.848	0.844	0.751	0.760	0.709	0.713	0.659	0.669	0.640	0.656
	2	0.891	0.884	0.798	0.802	0.753	0.752	0.685	0.695	0.661	0.677
	3	0.926	0.923	0.818	0.820	0.752	0.755	0.708	0.719	0.679	0.699
Milkdrop	0	0.697	0.695	0.618	0.625	0.590	0.598	0.574	0.583	0.562	0.570
	1	0.763	0.763	0.672	0.679	0.634	0.647	0.592	0.606	0.580	0.590
	2	0.825	0.827	0.734	0.738	0.694	0.703	0.611	0.627	0.598	0.609
	3	0.851	0.854	0.712	0.727	0.665	0.683	0.633	0.650	0.612	0.627

2 抗 JPEG 压缩的半脆弱图像水印算法

2.1 水印嵌入

本文的算法选用的载体图像为 $m\times m$,嵌入的水印是$(m/4)\times(m/4)$的二值图像.为了提高系统的安全性,实际嵌入到图像中的水印图像是原始水印图像经过一种由密钥控制的置乱加密算法置换之后得到的.嵌入时先将载体图像做一次参数化整数小波提升,然后在小波系数的 LH_1 子带用 2×2 的互不重叠的滑动窗截取小波系数,由于 LH_1 子带有 256×256 个系数,故每 4 个小波系数对应于 1 个水印图像素,所以,每一个滑动窗就对应于水印图像相应位置的像素.计这个 2×2 的滑动窗矩阵为 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.若对应的水印 bit 为 0,则要求 $a_{11}-a_{12}>g$ 且 $a_{21}-a_{22}>g$(g 为预先设定的阈值,$g\geq 0$);若不满足这个条件,则分别增加 a_{11},a_{21},同时分别减少 a_{12},a_{22},直到满足为止.若对应的水印 bit 为 1,则要求 $a_{11}-a_{12}<-g$ 且 $a_{21}-a_{22}<-g$;若不满足这个条件,则分别减少 a_{11},a_{21},同时分别增加 a_{12},a_{22},直到满足为止.若水印选择嵌入在 HL_1 子带,仍然用 2×2 的滑动窗截取小波系数,同样计为 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.若对应的水印 bit 为 0,则要求 $a_{11}-a_{21}>g$ 且 $a_{12}-a_{22}>$g;若不满足这个条件,则分别增加 a_{11},a_{12},同时分别减少 a_{21},a_{22},直到满足为止.若对应的水印 bit 为 1,则要求 $a_{11}-a_{21}<-g$ 且 $a_{12}-a_{22}<-g$;若不满足这个条件,则分别减少 a_{11},a_{12},分别增加 a_{21},a_{22},直到满足为止.从上述水印嵌入的方法可推知,一对相邻的小波系数要么不必作任何修改,要么需要作修改,而修改之后系数之差一定是$\pm(g+1)$.

2.2 水印嵌入区域的选择

本算法给水印的嵌入提供了两个区域,LH_1 子带和 HL_1 子带.由第 2 节的实验数据可知,只要选取相邻小波系数的方向合适(LH_1 子带在水平方向上选取,HL_1 子带在垂直方向上选取),小波系数对大小关系受 JPEG 压缩

© 中国科学院软件研究所 http://www.jos.org.cn

的影响就相差不大,在选择嵌入区域时主要考虑嵌入区域对嵌入后图像的 PSNR 的影响.由第 2.1 节的水印嵌入方法可知,PSNR 的高低与修改的小波系数的多少及修改量有关.由于置乱后的水印 bit 值的分布基本上是 0,1 各占一半,所以水印嵌入到 LH_1 或 HL_1 子带所修改的小波系数对的多少并不会有太大差别,这样,对系数的修改量就成了影响 PSNR 的主要因素.相邻小波系数的差值越小,在整个嵌入过程中对小波系数的修改量就越小,嵌入后图像的 PSNR 也就越高.为此,定义两个量 *avr_LH* 与 *avr_HL*,定义如下:

$$avr\_LH = \frac{2\sum_{y=0}^{H-1}\sum_{x=0}^{\frac{W}{2}-1}[p(2x,y)-p(2x+1,y)]^2}{H\times W} \tag{1}$$

$$avr\_HL = \frac{2\sum_{x=0}^{W-1}\sum_{y=0}^{\frac{H}{2}-1}[p(x,2y)-p(x,2y+1)]^2}{H\times W} \tag{2}$$

其中 H 与 W 分别是小波子带的高与宽,$p(x,y)$代表坐标(x,y)处的小波系数.这两个量反映了 LH_1,HL_1 两个子带相邻小波系数的均方差,由前面的分析可知,也反映了嵌入后图像 PSNR 的高低.本算法在嵌入水印之前先计算出这两个量,然后将水印嵌入到较小量的子带中,从而可以获得更高的 PSNR.几种常见图像 *avr_LH* 与 *avr_HL* 的测试结果见表 3.

Table 3 The average variance of neighboring wavelet coefficients of test images

表 **3** 测试图像相邻小波系数的均方差

Image	Lena	Baboon	Boat	Goldhill	Lake	Milkdrop
avr_LH	112.229 6	544.563 7	356.649 6	217.793 4	204.904 0	113.178 8
avr_HL	41.880 6	1 672.2	98.452 5	165.427 3	276.399 1	89.611 6

分别选用 512×512×8bits 的测试图像,在相同的 $g(g=0)$下分别在 LH_1 子带及 HL_1 子带嵌入 128×128 的二值水印图像,统计其信噪比,得到的数据见表 4.

Table 4 PSNR for embedding watermark into different sub-bands of high frequency of test images

表 **4** 嵌入测试图像不同高频子带得到的水印图像的 PSNR

Image	Lena	Baboon	Boat	Goldhill	Lake	Milkdrop
PSNR (LH_1)	41.132 2	35.027 6	36.632 0	38.423 0	38.146 5	39.863 1
PSNR (HL_1)	44.220 4	30.057 5	41.843 9	39.487 4	37.251 3	43.113 4

对照表 3 与表 4 可以发现,将水印嵌入到相邻系数均方差较小的子带中可以获得较高的 PSNR.

2.3 水印提取

在提取水印时,应首先确定水印是嵌入在哪个子带.由第 2.1 节可知,在嵌入水印前后,相邻系数之差要么不改变,要么为$\pm(g+1)$.也就是说,嵌入水印的子带其相邻小波系数之差的绝对值要么不变,要么为 $g+1$.这样,当 $g+1\le\min\{avr_LH, avr_HL\}$时,嵌入水印的操作必然减小了嵌入的子带其相邻小波系数之差的平均值.本来嵌入水印的子带其相邻系数之差平均值就小,嵌入水印后差值更小.所以,提取水印时在将待检测图像做一次参数化整数小波提升之后就计算 *avr_LH* 与 *avr_HL*,两者中较小的说明水印是嵌入在相应子带中的.

表 5 给出了测试图像在嵌入前后相应子带相邻小波系数平均差的结果$(g=0)$.可见,嵌入水印的子带其相邻小波系数平均差确实较嵌入前有所减小.

确定水印的嵌入子带后,在相应子带用 2×2 的滑动窗截取小波系数,记为 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.对于 LH_1 子带,若 $a_{11}+a_{21}>a_{12}+a_{22}$,则对应的水印 bit 为 0,反之为 1;对于 HL_1 子带,若 $a_{11}+a_{12}>a_{21}+a_{22}$,则对应的水印 bit 为 0,反之为 1.所得到的水印图像经过一个反置乱算法即可得到原始的水印图像.

© 中国科学院软件研究所 http://www.jos.org.cn

Table 5 Alteration of average variance of neighboring wavelet coefficients of test images after embedding watermarking

表 5 嵌入水印对子带相邻小波系数均方差的影响

Image	Lena	Baboon	Boat	Goldhill	Lake	Milkdrop
Embedded sub-band	HL_1	LH_1	HL_1	HL_1	LH_1	HL_1
The average variance of near wavelet coefficients before embedding watermarking	41.880 6	544.563 7	98.452 5	165.427 3	204.904 0	89.611 6
The average variance of near wavelet coefficients after embedding watermarking	25.926 4	276.699 9	52.719 2	79.552 1	81.149 5	57.864 9

2.4 图像安全检测

2.4.1 篡改判别准则

由文献[11]可知,可由水印差图中稀疏点与稠密点的比值来判断攻击类别,其中水印差图的定义为

$$D(i,j)=\left|W(i,j)-W'(i,j)\right| \tag{3}$$

差值图像的实际意义是比较两幅二值图像之间的差异,如果相应像素点的像素值相等,则在差值图像上像素值为 0,即表现为白色点;反之则为 1,表现为黑色点.

文献[11]给出了稀疏点的定义,按照这个定义计算水印差图中稀疏点与稠密点分别的个数,并定义:

S={水印差图的像素点总数};

S_{dense}={差图中稠密点的像素点总数};

S_{sparse}={差图中稀疏点的像素点总数};

$S_{total}=S_{dense}+S_{sparse}$; $\lambda=S_{total}/S$; $\delta=S_{sparse}/S_{dense}$.

判断步骤是:首先计算λ,若λ=0,则说明没有篡改;若λ>0,则计算δ,若$\delta\geq T$(T 为检测方设定的一个非负的门限值),则认为是偶然攻击,若$\delta<T$,则认为是恶意篡改.阈值 T 的选取依据是最小均方误差准则,目的是使得提取水印达到视觉上可以接受的程度[8].T 的值取得越高,对偶然攻击的判断条件就越严格,整个系统的安全性也就越高.通过对具有不同纹理图像的实验,阈值确定为 0.5 时效果比较理想.

2.4.2 形态滤波

判断了攻击的类别之后,对于偶然攻击,需要对提取得到的水印进行形态学上的处理,使其在视觉上给出尽可能好的效果;对于恶意篡改,则需要对差图进行处理,使篡改区域更加紧凑,从而尽可能准确地定位篡改区域.

对于恶意篡改的定位,文献[11]给出了效果不错的结构元素与滤波顺序,本文采用该方法.

对于偶然攻击,需要对提取到的水印进行处理.在对水印图进行滤波的时候,考虑到在绝大多数有意义的二值水印图像中,黑色点与白色点的分布都各自较为集中,所以本算法给出一个新的稀疏点定义,计为稀疏点 II.首先给出四连通与八连通点的定义:

在二值图像中,当前像素点为 $x(i,j)$,则 $x(i-1,j)$,$x(i,j-1)$,$x(i,j+1)$,$x(i+1,j)$称为 $x(i,j)$的四连通点,而 $x(i-1,j-1)$,$x(i-1,j+1)$,$x(i+1,j-1)$,$x(i+1,j+1)$称为 $x(i,j)$的八连通点.

下面给出稀疏点 II 的定义:在一幅二值图像中(不包括边界点),对于任一像素点,若满足以下两个条件中任意一个,即可视为稀疏点 II:

(1) 该像素点没有一个四连通点与之灰度值相同;

(2) 该像素点有且仅有一个四连通点与之灰度值相同,但没有与该像素点灰度值相同的八连通点.

在对受偶然攻击提取的水印图的滤波中,我们需要去掉的就是这些稀疏点 II.所以,首先对水印图做循环滤波,直到没有这样的稀疏点 II 存在为止.为使最后得到的水印图在视觉上有更好的效果,也采用形态滤波的方法,先在行列方向上分别做一次膨胀,然后分别在行列方向上做一次腐蚀,膨胀与腐蚀均选用长度为 2 的结构元素.进行形态滤波处理以后,再将处理过的水印图滤波去除稀疏点即得到最终的水印图.

2.5 水印安全性

为了增强水印的安全性,在水印嵌入前对水印进行置乱.置乱算法基于文献[12],因为该方法是建立在色彩

© 中国科学院软件研究所 http://www.jos.org.cn

空间变换的基础上的,不会将攻击分散,所以对提取水印时进行篡改定位十分方便.置乱时,发方先随机选取一个密钥 Key,然后通过哈希函数得到和水印图像高度 m 与宽度 n 具有相同长度的字符串口令 $P_H=\{P_H^1,P_H^2,\ldots,P_H^m\}$与 $P_W=\{P_W^0,P_W^1,\ldots,P_W^n\}$,将水印图像置乱后进行嵌入.同时,发方用收方的公钥加密 Key,将加密后的 Key'通过安全通道传给收方,收方收到 Key'用自己的私钥解密得到 Key,然后利用相同的哈希函数可以得到 P_H 与 P_W,从而将收到的水印图像去乱.对于本文的水印图像,字符串口令的可能性共有 $[(2^{128})^{128}]^2=2^{32768}$ 种.用穷举法来搜索字符串口令几乎不可能.可见,水印的安全性是有保证的.水印的安全基于 Key 与收方私钥的保密.

3 实验结果

3.1 阈值g的选取

由第 2.3 节可知,当 $g+1\le\min\{avr_LH,avr_HL\}$时,嵌入水印的操作必然减小了嵌入的子带其相邻小波系数之差的平均值,使得我们可以通过计算待检测图像 HL_1 与 LH_1 子带相邻小波系数的平均差值来判断水印究竟嵌入在哪个子带.所以,在选取阈值 g 时应注意满足 $0\le g\le\sqrt{\min\{avr\_LH,avr\_HL\}}-1$.

选用 512×512×8bits 灰度图像 lena,baboon,milkdrop 在 g 分别为 0,1,2,3,4,5 的情况下嵌入水印.嵌入水印后的图像与原图在视觉上没有任何区别,如图 2 所示.

(a) Original images

(a) 原始图像

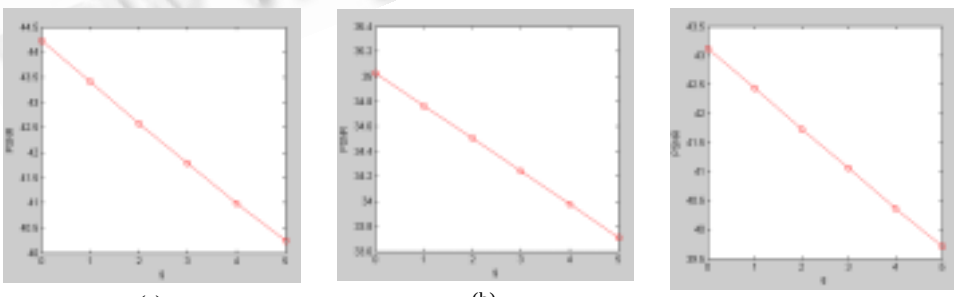
(b) Watermarked images, g=2

(b) 嵌入水印后的图像,g=2

Fig.2

图 2

由实验结果可知,嵌入水印后图像的 PSNR 随着 g 的增加而减少,且大致呈线性关系,如图 3 所示.其中,图 3(a)~图 3(c)所用载体图像分别为 Lena,Baboon,Milkdrop.



(a) (b) (c)

Fig.3 The relationship between PSNR and g

图 3 PSNR 与 g 的关系

© 中国科学院软件研究所 http://www.jos.org.cn

对嵌入水印后的 Lena 图像分别进行压缩因子为 90,80,70,60,50 的 JPEG 压缩,对压缩之后的图像提取水印,按照第 2.4.1 节的方法计算δ,并给出对攻击的判断类别(T=0.5).表 6 为实验数据,其中 Y 表示被认为是偶然攻击,N 表示被认为是恶意攻击.

Table 6 Watermark data extracted from watermarked Lena after JPEG Watermark data extracted from watermarked Lena after JPEG

表 **6** 对嵌入水印后的 Lena 图像经压缩后提取水印得到的数据

g \ Quality factor / δ	90	80	70	60	50
1	2.065 2	1.633 0	1.004 1	0.565 6	0.429 5
2	4.439 7	3.419 2	2.249 3	1.145 0	0.7176
1	Y	Y	Y	Y	N
2	Y	Y	Y	Y	Y

由表 6 的实验结果可知,当 T 取 0.5,g=2 时,水印即可达到抗质量因子为 50 的 JPEG 压缩的要求.如图 4 所示,图 4(a)为原始水印,图 4(b)~图 4(f)分别为当 g=2 时,经质量因子 90,80,70,60,50 压缩后提取的水印经形态学处理之后的水印图像.

(a) (b) (c) (d) (e) (f)

Fig.4 Robustness of the watermark resisting to JPEG

图 4 水印抗 JPEG 压缩的性能

由实验结果可知,水印抗 JPEG 压缩的能力随着 g 的增大而提高.

3.2 算法对恶意篡改的定位能力

由实验结果可知,无论是大面积的篡改还是小块的篡改,都能够被判断为恶意攻击并给予准确地定位,如图 5 所示.

(a) Embedded watermark image(g=1), the area in broken line will be replaced by the original image

(a) 水印图像(g=1),虚线区域内的部分将以原始图像内容替换

(b) Image after tamper

(b) 篡改结果图

(c) Extracted watermark (δ=0.0942)

(c) 水印提取(δ=0.0942)

(d) Final difference image

(d) 最后得到的差图

Fig.5 Location for tamper

图 5 篡改探测图

© 中国科学院软件研究所 http://www.jos.org.cn

3.3 与其他算法的比较

为验证本文算法的有效性,将本文的算法与文献[6,13]作了比较.所用到的载体图像均为 Lena,比较结果见表 7.

Table 7 Performance comparison with other algorithms

表 7 与其他算法的性能比较

	Size of host image	Size of watermark	Precision of orientation for tamper	PSNR (db)	Quality factor of resisting to JPEG
This paper (g=1)	512×512	128×128	Very exactly	43.42	60
Reference[6]	256×256	1024	Commonly	44	70
Reference[13]	512×512	64×64	Relatively exactly	43.56	65

由比较可知,在 PSNR 相差不多的情况下,本文算法的嵌入容量远远超过了文献[6,13]的方案,均是二者的 4 倍,这可以大幅度地改进篡改检测的精度,且抗 JPEG 压缩性能也有所提高.文献[13]中用到的水印图像尺寸偏小,实用应用价值不大[14].

4 结 论

本文的主要思路在于利用 JPEG 压缩对图像小波系数大小关系的影响提出了一种半脆弱水印算法.所提出的脆弱水印系统具有以下特点:

(1) 根据数字图像小波分解的高频系数之间的大小关系在 JPEG 压缩之后大多数没有发生变化这一事实,完成水印的嵌入及提取.实验表明,该算法能有效地抗 JPEG 压缩.

(2) 采用参数化整数小波分解图像,运算量小.水印嵌入之前先经过置乱,从而使攻击者在不知道密钥和小波分解参数的情况下很难获得任何关于水印的信息.

(3) 针对有意义的水印信息处理提出了一种新的滤波规则.实验表明,该规则对于大多数有意义水印信息都能明显改善滤波效果.

下一步的研究工作将致力于利用参数小波分解图像,以进一步提高算法的安全性.

References:

[1] Hua XS, Shi QY. Research on fragile watermarking problems. Journal of Image and Graphics, 2001,6(11):1089−1095 (in Chinese with English abstract).

[2] Lin ET, Delp EJ. A review of fragile image watermarks. In: Proc. of the ACM Multimedia and Security Workshop. Orlando: ACM Press, 1999. 25−29.

[3] Fridrich J. Methods for tamper detection in digital images. In: Proc. of the ACM Workshop on Multimedia and Security. Orlando: ACM Press, 1999. 19−23.

[4] Ho CK, Li CT. Semi-Fragile watermarking scheme for authentication of JPEG images. In: Proc. of the Int'l Conf. on Information Technology: Coding and Computing 2004. Piscaterway: IEEE Press, 2004. 7−11.

[5] Zhang J, Zhang CT. Semi-Fragile watermarking for JPEG2000 image authentication. ACTA Electronica Sinica, 2004,32(1): 157−160 (in Chinese with English abstract).

[6] Zhou X, Duan XH, Wang DX. A semi-fragile watermark scheme for image authentication. In: Proc. of the 10th Int'l Conf. on Multimedia Modelling. Piscaterway: IEEE Press, 2004. 374−377.

[7] Hu JQ, Huang JW, Huang DR. A DWT-based fragile watermarking tolerant of JPEG compression. In: Proc. of Int'l Workshop on Digital Watermarking 2002. LNCS 2613, Berlin: Springer-Verlag, 2003. 179−188.

[8] Hu JQ. Fragile watermarking: Methods and applications [Ph.D. Thesis]. Guangzhou: Sun Yat-Sen University, 2003 (in Chinese with English abstract).

[9] Lie WN, Chang LC. Robust and high-quality time-domain audio watermarking subject to psychoacoustic masking. In: Proc. of the 2001 IEEE Int'l Symp. on Circuits and Systems. Piscaterway: IEEE Press, 45−48.

© 中国科学院软件研究所 http://www.jos.org.cn

[10] Lin CY, Chang SF. A robust image authentication method distinguishing JPEG compression from malicious manipulation. IEEE Trans. on Circuits and Systems of Video Technology, 2001,11(2):153−168.

[11] Hu JQ, Huang JW, Huang DR, Shi YQ. Image fragile watermarking based on fusion of multi-resolution tamper detection. Electronics Letters, 2002,38(24):1512−1513.

[12] Yan WQ, Zou JC, Qi DX. A novel digital image scrambling method. Journal of North China University of Technology, 2002,14(1):1−7 (in Chinese with English abstract).

[13] Sun R, Sun H, Yao TR. A SVD- and quantization based semi-fragile watermarking technique for image authentication. In: Proc. of the 6th Int'l Conf. on Signal Processing. Beijing, 2002. 1592−1595.

[14] Katzenbeiseer S, Petitcolas FAP; Wu QX, *et al*., Trans. Information Hiding Techniques for Steganography and Digital Watermarking. Beijing: Posts & Telecom Press, 2001 (in Chinese).

附中文参考文献:

[1] 华先胜,石青云.易损数字水印若干问题的研究.中国图像图形学报,2001,6(11):1089−1095.

[5] 张静,张春田.用于 JPEG2000 图像认证的半脆弱性数字水印算法.电子学报,2004,32(1):157−160.

[8] 胡军全.脆弱水印的方法及其应用[博士学位论文].广州:中山大学,2003.

[12] 闫伟齐,邹建成,齐东旭.一种基于 DES 的数字图像置乱新方法.北方工业大学学报,2002,14(1):1−7.

[14] Katzenbeiseer S, Petitcolas FAP;吴秋新,等,译.信息隐藏技术——隐写术与数字水印.北京:人民邮电出版社,2001.

李春(1982－),男,四川巴中人,硕士,工程师,主要研究领域为多媒体信息处理.

黄继武(1962－),男,博士,教授,博士生导师,主要研究领域为多媒体信息安全.

© 中国科学院软件研究所 http://www.jos.org.cn