

一种基于直方图统计特征的灰度图像水印算法

王树梅¹, 范勇²

(1. 徐州师范大学 计算机科学与技术学院, 江苏 徐州 221116;

2. 徐州师范大学 教育技术中心, 江苏 徐州 221116)

摘要:直方图是数字图像的一个重要全局特征的表示方法,利用直方图的这个性质在空域里实施水印的嵌入和提出。在水印嵌入过程中,首先根据载体图像的特征得到其相应的直方图,即在相应灰度级下各个像素值出现的几率大小;然后计算出直方图相邻元素的距离,找出其最大值;再根据最大值和经过预处理过的二值水印来调整直方图的布局;最后得到含水印图像。在水印提取过程中,根据含水印图像和原始图像的直方图特点,找到嵌入水印的像素值,从而提取出水印。实验结果证明,该算法具有较好的抗几何攻击性质。

关键词:数字水印;直方图;二值水印;统计特征

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2010)01-0177-04

A Watermarking Algorithm of Gray Image Based on Histogram Statistical Characteristics

WANG Shu-mei¹, FAN Yong²

(1. School of Computer Science and Technology, Xuzhou Normal University, Xuzhou 221116, China;

2. Educational Technology Center, Xuzhou Normal University, Xuzhou 221116, China)

Abstract: Digital image histogram is an important global features, histogram will be used to embed binary watermark in the space domain. In the watermark embedding process, firstly, the corresponding histogram of image is generated according to the original images' characteristics, which shows the probability of the corresponding pixel in gray scale; Secondly, the distances between adjacent elements of histogram are calculated, and found their maximum; Then, the histogram is adjusted according to the maximum and pretreated watermark; Finally, the watermark image is generated. In the watermark extracting process, in accordance with the histogram of watermark image and the original image, and the pixels embedded watermark are found, thereby extracting watermark. The experimental results show that the algorithm has good anti-geometric nature of attacks.

Key words: digital watermarking; histogram; binary watermark; statistical characteristics

0 引言

当前,图像水印算法已经从空域转向了频域^[1-4],因为频域中的水印算法的安全性较强,还有就是鲁棒性也比较好。但是对于频域水印隐藏算法的难点有两个,一是如何避免由频域转化到空域过程中所引入的误差,二是如何解决基于频域水印图像的篡改定位问题。虽然采用整数变换可以避免取整误差,但溢出误差和定位问题却较难解决。空域水印隐藏技术解决了这两个难题,所以已经有很多的专家在这方面做了很多的研究工作。Hosinger^[5]给出了一种基于模加法的

信息隐藏技术,但该算法的缺点在于当图像像素点的灰度值接近于0或255时,往往会引起较大强度的失真,导致胡椒盐(salt-and-pepper)噪声的出现。Fridrich^[6]给出了一种基于无损压缩LSB比特位的水印算法,其水印的容量取决于隐藏位置和无损压缩算法的选取。Macq给出了一种基于修改的拼贴法(patchwork algorithm)的无损水印隐藏算法。但以上算法的共同缺点在于容量有限,不能嵌入大量的水印信息。文中给出一种基于灰度直方图的无损水印算法,该算法的优点是:水印容量大、失真小,且易重构。

将直方图与水印技术相结合的方法可以分为两类:修改载体直方图到指定直方图的水印算法和保持原始直方图特性的水印算法。文献[7]给出了两种基于直方图规范化的水印嵌入算法,其中一种方法中采用的水印信息是一种指定的直方图,通过直方图规范

收稿日期:2009-05-14;修回日期:2009-08-30

基金项目:徐州师范大学自然科学基金(XLB08016)

作者简介:王树梅(1972-),女,山东单县人,博士,讲师,研究方向为数字水印技术与图像处理。

化方法修改载体图像直到其具有指定的直方图。这种方法有两个优点:第一,能够较好地保持图像的视觉效果,而且如果采用均匀直方图,在嵌入水印的同时也增加了图像的对比度;第二,能够抵抗不改变载体灰度分布的几何攻击等。但这种算法也有自身的缺点,即对压缩等攻击的鲁棒性较差,因为这些攻击对图像的灰度分布影响较大。文献[8]则选择了一类特殊的具有周期性并且覆盖所有灰度级的水印,从而减少了某些灰度的出现频率,使得载体图像的视觉效果能够保持得更好。这两种算法都是修改原始载体直方图到指定直方图的水印算法。此外,在文献[9,10]中也提出了利用直方图特性来嵌入水印信息的方法,其中在文献[9]中提出了一种在彩色图像中应用直方图特性嵌入水印信息的技术,在文献[10]中给出了一种基于时间误差模型、空间误差模型和直方图误差模型的水印算法。但这些算法都是直接根据直方图峰值的多少来调节像素值大小,没有充分体现像素本身的特点。文献[11]根据直方图统计信息,找出空缺灰度级,从而在空缺处嵌入水印。这样做虽然可以增加嵌入水印的信息量,但因为改变了原有图像灰度级数量,从而影响载体图像的质量。

文中的主体思想仍然是根据灰度图像直方图的统计特征来嵌入水印,这里没有根据指定的直方图来修改原始图像直方图,而是根据直方图本身的统计特征和二值水印信息来调制原始图像的像素值的,从而达到修改原始直方图的目的。该算法一个比较大的优点就是保持了原始图像的统计特征,失真较小,抗几何攻击性较强。但是检测时需要原始图像。

1 直方图

在数字图像处理中,一种简单而有效的工具是灰度直方图。灰度直方图是数字图像中的一个重要的全局特征,它反映了不同灰度值 d 的面积或在图像中占的比例。直方图与像素的位置无关,不涉及图像的形状问题。理论上,各个灰度值在图像中出现的概率是相等的,但对于一幅图像,各个灰度出现的概率并不相同,有的出现频率非常小,甚至为 0,而有的灰度出现的频率却非常高。统计结果表明,区间 $[0,255]$ 两端的灰度出现的概率往往小于中间的灰度出现的概率,所以可以说灰度直方图描述了图像由中心灰度值(图像的平均灰度值)滑向两端灰度值的变化过程。充分利用这些在图像中出现频率最高和最低的灰度可以隐藏信息。

直方图是图像的一种统计表达。对一幅图像,其灰度统计直方图反映了该图中不同灰度级出现的统计

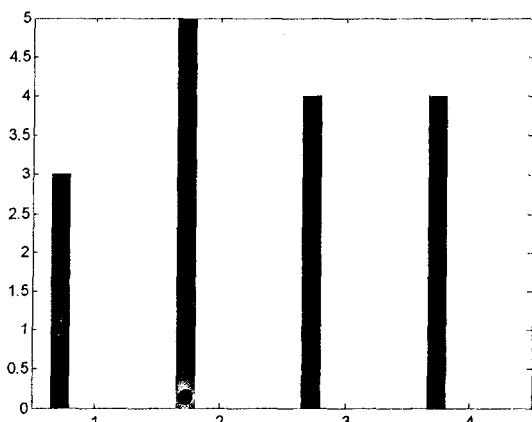
情况。严格地说,图像的灰度统计直方图是一个 1-D 的离散函数,可写成:

$$h(k) = n_k \quad k = 0, 1, \dots, L-1 \quad (1)$$

其中 n_k 是图像 $f(x, y)$ 中具有灰度值 k 的像素的个数。直方图提供了原图中各种灰度值分布的情况,也可以说给出了一幅图像所有灰度值的整体描述。例如一幅图像如图 1(a), 则其灰度统计直方图可表示为图 1(b), 其中横轴表示不同的灰度级, 而纵轴表示了图像中各个灰度级像素的个数。

0	1	2	3
1	2	3	1
2	3	1	0
3	1	0	2

(a)



(b)

图 1 图像和直方图

2 水印的嵌入和提取过程

2.1 水印预处理

为了加强水印信息的安全性,在水印嵌入之前需要对水印利用密钥进行置乱,以得到难以辨认的信息,只有密钥的人才可以知道真正的信息意义是什么。通常有很多将水印置乱的办法,在这里利用 Logistic 方程作为产生混沌序列的发生器模型。因为由 Logistic 方程产生的混沌序列具有两个优点:一是迭代不重复性,二是初值敏感性。

Logistic 方程的形式为:

$$x_{n+1} = \mu x_n (1 - x_n) \quad (2)$$

只要给定合适的 μ 值,就能产生满足混沌特性的序列。在这里利用 Logistic 方程将水印图像的行和列进行混沌排序,从而达到使得水印像素值置乱的目的。

2.2 水印算法

● 嵌入过程:

(1) 读入文件,得到原始图像的数据矩阵 $f(i, j)$, $i = 1:M, j = 1:N, M$ 和 N 为图像的行数和列数;

(2) 计算出每个像素值出现的次数,即图像中每个像素出现的频率,用数列 $H(s_k)$,用(1)式表示;

(3) 计算出 $H(s_k)$ 中元素之间的距离,得到数列 $D(s_k)$,即灰度级为 $k+1$ 的像素个数与灰度级为 k 的像素个数之间的差,用式(3)表示:

$$D(s_k) = |n_{k+1} - n_k| \quad k = 0, \dots, L-2 \quad (3)$$

(4) 在 $D(s_k)$ 中,找出最大的值 D_{\max} ,并记录其灰度级为 L ;

$$D_{\max} = \max(D(s_k), k = 0, \dots, L-2) \quad (4)$$

(5) 扫描每一个像素值,将水印嵌入到灰度级为 L 的像素值中,具体嵌入办法如下:

$$f(i, j) = \begin{cases} f(i, j) + 1 & \text{if } |H(f(i, j) + 1) - H(f(i, j))| = D_{\max} \\ & \text{且 } w(k) = 1 \\ f(i, j) - 1 & \text{if } |H(f(i, j)) - H(f(i, j) - 1)| = D_{\max} \\ & \text{且 } w(k) = 0 \\ f(i, j) & \text{其他情况} \end{cases}$$

(6) 检查水印信息嵌入量是否等于 length ,如果是,就结束,否则转(2);

● 提取过程:

(1) 读入含水印文件,得到原始图像的数据矩阵 $H_1(i, j)$, $i = 1:256, j = 1:256$;

(2) 计算出每个像素值出现的次数,即图像中每个像素出现的频率,用数列 $H_1(s_k)$ 表示:

$$H_1(s_k) = n_k \quad k = 0, \dots, 255 \quad (5)$$

其中 s_k 为第 k 个灰度级的灰度值, n_k 为图像中灰度值为 s_k 的像素个数。

(3) 计算出 $H_1(s_k)$ 中元素之间的距离,得到数列 $D_1(s_k)$,即灰度级为 $k+1$ 的像素个数与灰度级为 k 的像素个数之间的差,用下式表示:

$$D_1(s_k) = |n_{k+1} - n_k| \quad k = 0, \dots, L-2 \quad (6)$$

(4) 在 $D_1(s_k)$ 中,找出最大的值 D_{\max} ,即用下式表示:

$$D_{\max} = \max(D_1(s_k), k = 0, \dots, L-2) \quad (7)$$

(5) 水印嵌入到灰度级为 L 的像素值中,具体嵌入办法如下:

$$w(k) = \begin{cases} 1 & \text{if } |H(f(i, j) + 1) - H(f(i, j))| = D_{\max} \\ 0 & \text{if } |H(f(i, j)) - H(f(i, j) - 1)| = D_{\max} \end{cases}$$

(6) 检查水印信息提取量是否完等于 length ,如果是,转(7),否则转(2);

(7) 将得到的一维水印信息化成二维的矩阵 W ,然后根据密钥 key 将 W 重构(见图2)。

3 实验结果

这里采用 lena 图像、peppers 图像、baboon 图像和 fishingboat 图像四副灰度图像作为原始载体对象,采用 tongji 二值图像作为水印图像。图3中是四副原始图像及其不同的直方图,从图3可以看出,不同的图像有不一样的像素值分布,即有不同形状的直方图。这也是采用这四副图像作为实验对象的原因,它们都有比较明显的纹理,直方图都有较大的不同,更好地体现算法的效果。

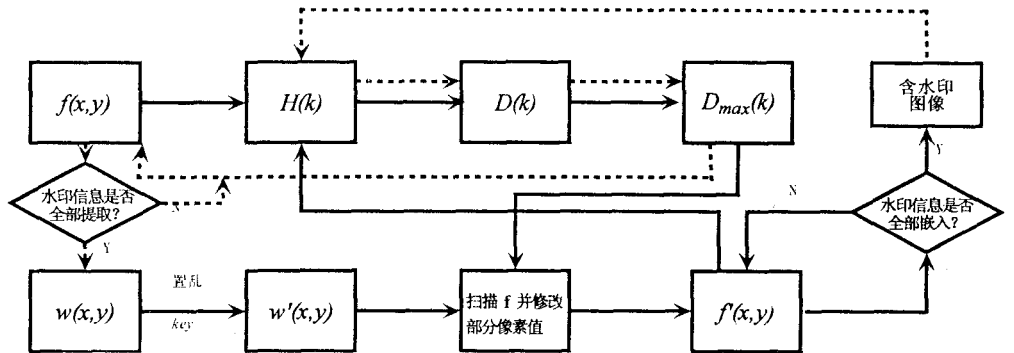


图2 水印的嵌入和提取过程

(实线为嵌入过程,虚线为提取过程)

图4是利用上述算法嵌入水印后的图像以及其对应的直方图,与图3中的四副原始图像对照,人眼看不出其差别,足以证明此算法的水印隐蔽性较好。在把图3和图4中对应直方图进行比较,不难看出直方图的布局发生了一些变化,其变化的原因在于水印的嵌入改变了部分像素值的大小。另外,可以观察出直方图的改变的地方在于相邻像素频率差距最大的地方有了改变,改变的程度取决于嵌入水印的大小和像素频率的大小。

图5是对 lena 图像进行不同角度的旋转后提取出来的水印图像。从提取出来的水印可以看出,尽管有的水印噪音很多,但水印信息仍然很清晰地显示出来。可见此算法具有较强的抗几何攻击性,即具有较好的鲁棒性。再有,从表1的结果得知,嵌入水印后的四副



图3 四副原始图像以及其对应的直方图

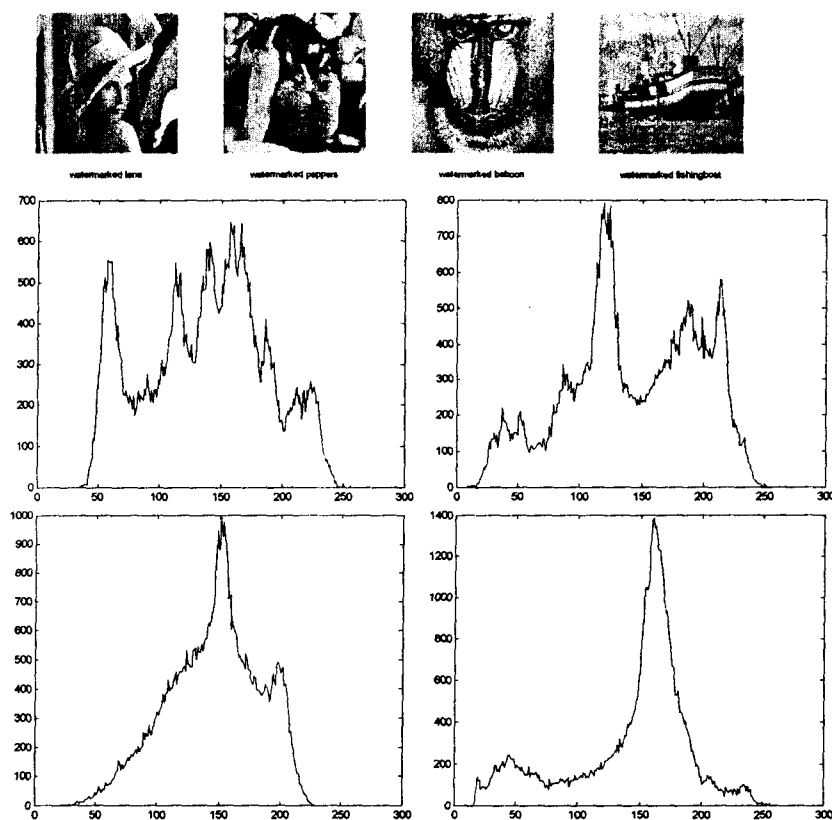


图4 嵌入水印以后的四副图像以及其对应的直方图



图5 对嵌入水印后 lena 图像进行不同角度旋转后提取的水印

图像的峰值信噪比都在 40dB 以上, 超出了人眼所能接受的底限 35dB, 因此, 再一次充分证明了本算法具有较强的不可见性。

表1 嵌入水印后四副图像的 PSNR

Image	Lena	baboon	peppers	fishingboat
PSNR	43.6622	42.8132	44.4131	44.4479

4 结束语

文中提出了一种基于直方图的灰度图像的水印算法, 避免了变换域水印算法引起的误差, 而且这种算法简单易行, 具有广泛的实用性。其实, 直方图作为灰度图像的一种比较重要的特征, 充分体现了图像像素值的布局, 也给水印的嵌入提供了最好的依据。所以该算法就是基于像素出现频率的不同提出的, 将水印嵌入到频率差别较大的像素中去。在这里, 嵌入的过程不是利用在水印系统中经常使用的加法规则和乘法规则, 而是根据自己制定的判断规则去改变原始像素值, 而且改变的大小在 $[-1, 1]$ 范围内, 一般不会影响图像的可视质量, 不可见性较好。

参考文献:

- [1] 王万良, 管秋, 杨旭东. 小波域数字水印改进算法及其性能分析[J]. 计算机辅助设计与图形学报, 2004, 16(9): 1235-1238.
- [2] 牛夏牧, 陆哲明, 孙圣和. 基于多分辨率分解的数字水印技术[J]. 电子学报, 2000, 28(8): 1-4.
- [3] 孙中伟, 冯登国. DCT 变换域嵌入图像水印的检测算法[J]. 软件学报, 2005, 16(10): 1798-1804.
- [4] 于帅珍. 一种基于小波域的彩色图

(下转第 184 页)

DM)与传统的朴素贝叶斯(Naive Bayes)分类法和支撑向量机(SVM)方法在邮件正确率与召回率方面进行比较。

表 1 很直观地显示出基于判别模型的垃圾邮件分类效果要优于 Naive Bayes 与 SVM 分类方法(该实验采用的阈值是 0.9)。

表 1 正确率和召回率对比试验结果

分类方法	SP	SR
Naive Bayes	97.32%	95.08%
SVM	98.21%	96.54%
DM	99.06%	97.40%

第二个对比实验,采用文中提出的方法与 K 近邻算法(KNN)在邮件过滤的精确率方面进行比较,从图 2 可以看出基于判别模型的方法在过滤精度方面要优于 KNN。从算法的复杂度来看,KNN 没有训练过程,原理比较简单,它的过滤速度慢,因此不适用于过滤速度要求较高的场合。

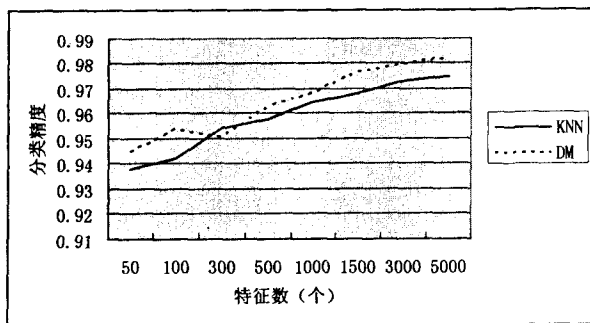


图 2 在不同特征个数下精确率的比较

4 结束语

基于判别方法的垃圾邮件过滤在现代研究中引起比较少的关注,文中提出了一个简单的方法,在某些情况下,它的分类效果要优于其他传统方法。实验很清楚地表明,基于判别模型的方法相对于传统的方法,在垃圾邮件的过滤方面,可以有效地提高正确率和准确率。

参考文献:

- [1] 中国互联网协会反垃圾邮件中心. 年度反垃圾邮件报告 [EB/OL]. 2007-04-05. <http://www.anti-spam.cn/>.
- [2] 潘文锋. 基于内容的垃圾邮件过滤研究[D]. 北京: 中国科学院研究生院, 2004.
- [3] Hulten G, Goodman J. Tutorial on junk email filtering [R/OL]. In ICML 2004: <http://www.research.microsoft.com/~joshuago/tutorialOnJunkMailFilteringjune4.pdf>.
- [4] 张文良, 黄亚楼, 倪维健. 基于差分贡献的垃圾邮件过滤特征选择方法[J]. 计算机工程, 2007, 33(8): 80-82.
- [5] Goodman J, Yih Wen-tau. Online discriminative spam filter training [C]//The Third Conference on Email and Anti-spam (CEAS). California: [s. n.], 2006.
- [6] 马莉, 柴乔林. 基于 Postfix 的垃圾邮件过滤技术的实现[J]. 计算机工程与设计, 2005, 26(4): 999-1001.
- [7] Androutsopoulos I, Paliouras G, Karkalexis V. Learning to filter spam E-mail: A comparison of a naive bayesian and a memory-based approach [C]//The Fourth Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD). France: [s. n.], 2000: 1-13.
- [8] 程卫华, 尤晋元. 基于内容过滤的反垃圾邮件系统的设计与实现[J]. 安徽大学学报, 2007, 31(3): 30-33.

(上接第 180 页)

像自适应水印[J]. 计算机仿真, 2006, 23(7): 132-134.

- [5] Honsigner C W. Lossless recovery of an original image containing embedded data. US patent: 6278 791B1[P]. 2001.
- [6] Jessica F, Goljan M, Du Rui. Invertible Authentication [C]//Proc. SPIE Photonics West, vol. 3971, Security and Watermarking of Multimedia Contents III. San Jose, California: [s. n.], 2001: 197-208.
- [7] Coltue D, Bolon P. Watermarking by Histogram Specification [C]//Proceeding of SPIE. Bellingham: Society of Photo-Optical Instrumentation Engineers, 1999: 252-263.
- [8] Coltue D, Chassery J M, Bolon P. Image Authentication by Exact Histogram Specification [C]//2001 IEEE Fourth Workshop on Multimedia Signal Proceeding. CA: Institute of Electrical and Electronics Engineers Inc, 2001: 701-710.
- [9] Pei S C, Zeng Y C. Hiding Multiple Data in Color Image by Histogram Modification [C]//Proceeding of the 17th International Conference on Pattern Recognition. Piscataway: Institute of Electrical and Electronics Engineers Inc, 2004: 799-802.
- [10] Cheng H, Isnardi M A. Spatial Temporal and Histogram Video Registration for Digital Watermark Detection [C]//Proceeding of 2003 International Conference on Image Processing. Piscataway: Institute of Electrical and Electronics Engineers Computer Society, 2003: 393-396.
- [11] 李妍, 张佑生, 张挺. 一种基于直方图的可逆数字水印算法[J]. 计算机技术与发展, 2006, 16(10): 122-124.