基于奇异值分解的数字图像水印方法

刘瑞祯, 谭铁牛
(模式识别国家重点实验室, 中国科学院自动化研究所, 北京 2728 信箱, 北京 100080)

摘 要: 随着计算机和网络技术的飞速发展, 数字图像、音频和视频产品愈来愈需要一种有效的版权保护方法, 另外通信系统在网络环境下的信息安全问题也日益显露出来. 数字图像水印技术为上述问题提供了一个潜在的解决方案. 所谓水印技术就是将数字、序列号、文字、图像标志等版权信息嵌入到多媒体数据中, 以起到版权保护、秘密通信、数据文件的真伪鉴别和产品标志等作用. 本文提出了一种新的基于奇异值分解的数字水印算法并且对该方法的理论基础给出分析. 实验结果表明这种方法要比目前提出的流行算法鲁棒.

关键词: 数字水印; 奇异值分解; 鲁棒性
中图分类号: TN9111 73 文献标识码: A 文章编号: 0372 2112 (2001) 02 0168 04

SVD Based Digital Watermarking Method

LIU Rui2zhen, TAN Tie2niu
(National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences P. O. Box 2728, Beijing 100080, China)

Abstract: The advent of the Internet and the wide availability of computers, scanners and printers make digital data acquisition, exchange and transmission a simple task. However, making digital data accessible to others through networks also creates opportunities for malicious parties to make salable copies of copyrighted content without permission of the content owner. Digital watermarking is likely to be a potential solution to this problem. Digital watermarking has been proposed as a solution to the problem of copyright protection of multimedia documents in networked environments. It makes possible to embed a watermark (such as identification data, serials number, text or image etc.) to multimedia documents allowing copyright protection, secret communication, document authentication and so on. In this paper, we present a new digital image watermarking method based on SVD (Singular Value Decomposition), and then give some theoretical analysis about the algorithm. Extensive experimental results show that this method is much more robust than other methods presented before.

Key words: digital watermark; singular value decomposition; robustness

1 引言

因特网的出现以及计算机、扫描仪和打印机的广泛使用使得数字化多媒体数据的获取、交换和传输变得异常简单. 但是通过网络传播数据也使有恶意的个人或团体在没有得到数据文件所有者许可的情况下能肆意地复制和传播有版权保护的文档. 数字水印技术则为这种问题提供了一个潜在的解决方案[1~5].

数字水印就是向欲被保护的多媒体数据嵌入某种信息(即水印)以保护所有者的权益. 一个有效的水印算法框架至少应满足下面两个特性: 不可觉察性(Imperceptibility): 即原始的数据文档和添加水印后的数据文档对人的感觉器官应是一样的; 鲁棒性(Robustness): 给定一个水印文档, 非授权的个人或团体在使文档可用的情况下无法剔除水印. 鲁棒性问题对数字水印非常重要. 有效的数字水印应该能够承受大量不同的物理和几何失真, 包括有意的(如恶意攻击)或无意的(如图像压缩、滤波、扫描与复印、噪声污染、尺寸变化等等). 显然在经过这些操作后, 鲁棒的水印算法应仍能从水印图像中提取出嵌入的水印或证明水印的存在. 如果不掌握水印的所有有关知识, 数据产品的版权保护标志应该很难被伪造. 若攻击者试图删除水印则将导致多媒体产品的彻底破坏.

图像的水印技术根据水印嵌入的方式可以大致分为两类: 空域技术(水印被直接嵌入在图像的亮度值上)[6~9]和变换域技术(将图像做某种数学变换, 然后水印被嵌入于变换系数中)[10~12]. 从目前的情况看, 变换域方法正变得日益普遍. 因为变换域方法通常都具有很好的鲁棒性, 对图像压缩、常用的图像滤波以及噪声均有一定的抵抗力. 绝大多数变换域方法均采用了酉变换, 如离散余弦变换 DCT[10,13]、离散傅立叶变换 DFT[14]和离散小波变换 DWT[15~17]等等. 其中一个著名的方法是 Cox 等人提出的基于离散余弦变换的扩展谱通信(Spread Spectrum Communication)方法[10].

收稿日期: 1999 2082 27; 修回日期: 2000 2112 01
基金项目: 国家自然科学基金(No. 59825105); 中科院百人计划项目

在文献[10]中,作者先计算图像的离散余弦变换,然后将水印叠加到DCT域中幅值最大的前 k 个系数上(不包括直流分量),通常为图像的低频分量. 若DCT系数的前 k 个最大分量表示为$\{d_i\}$, i= 1, , k,水印为服从高斯分布的随机实数序列 W= $\{w_i\}$, i= 1, , , k,那么水印的嵌入算法为 $d_i = d_i(1+aw_i)$,其中常数 a 为尺度因子,控制水印添加的强度或能量. 然后用新的系数做反变换得到水印图像 I. 水印的检测过程是分别计算原始图像 I 和水印图像I 的离散余弦变换,提取嵌入的水印 W^*,再做相关检验 W# W^* / $\sqrt{W^* \# W^*}$ 以确定水印的存在与否. 该方法即使当水印图像经过一些通用的几何变形和信号处理操作而产生比较明显的变形后仍然能够提取出一个可信赖的水印拷贝.

本文提出一种新的基于奇异值分解(SVD)的数字水印方法,并将该方法与Cox的扩展谱方法进行比较. 由于SVD的良好特性,这种水印方法非常鲁棒,实验结果清楚地表明了这一点.

2　基于SVD的图像水印

数值分析中的奇异值分解(SVD)是一种将矩阵对角化的数值算法. 在图像处理中应用 SVD 的主要理论背景是:(1)图像奇异值的稳定性非常好,即当图像被施加小的扰动时,图像的奇异值不会有大的变化;(2)奇异值所表现的是图像的内蕴特性而非视觉特性.

本节将表述用SVD算法来完成水印的嵌入和提取方法.

211　奇异值分解

从线性代数的角度看,一幅灰度图像可以被看成是一个非负矩阵. 若一幅图像用 A 表示定义为 A I $R^{n@n}$(为方便起见,以后均只对方阵进行讨论),其中 R 表示实数域. 则矩阵 A 的奇异值分解定义如下:

$$A = USV^T \tag{1}$$

其中 U I $R^{n@n}$和 V I $R^{n@n}$均为正交阵, S I $R^{n@n}$为对角阵,上标 T 表示矩阵转置.

212　水印的嵌入和检测

SVD方法的基本原理是将水印嵌入到原始图像的奇异值中. 在水印的嵌入过程中,先做 n@n 灰度图像A 的奇异值分解,得到两个正交矩阵 U、V 及一个对角矩阵 S. 尽管假设 A 是方阵,但其他非方阵可以完全用同样的方法来处理. 这个特性是SVD方法的一个优点,因为很多流行的水印算法都不能直接处理长方阵[10, 15, 18].

水印 WI $R^{n@n}$被叠加到矩阵S 上,对新产生的矩阵 S+ aW 进行奇异值分解,得到 U_1、S_1 和 V_1(S+ aW= $U_1 S_1 V_1^T$),其中常数 a> 0 调节水印的叠加强度. 然后将矩阵 U、S_1 和 V^T 相乘,得到处理后的包含水印的图像 A. 即如果矩阵 A 和 W分别表示原始图像和水印,那么通过如下三个步骤得到水印图像 A:

$$\begin{aligned} A &\Rightarrow USV^T \\ S+aW &\Rightarrow U_1 S_1 V_1^T \\ A &\Leftarrow US_1 V^T \end{aligned} \tag{2}$$

在水印的检测过程中,如果给出矩阵 U_1、S、V_1 和可能损坏的水印图像 A^*,那么通过简单的逆过程就就可以提取出可能已经失真的水印 W^*,即:

$$\begin{aligned} A^* &\Rightarrow U^* S_1^* V^{*T} \\ D^* &\Leftarrow U_1 S_1^* V_1^T \\ W^* &\Leftarrow \frac{1}{a}(D^* - S) \end{aligned} \tag{3}$$

注意到三个矩阵 U_1、S 和 V_1 的总的自由度为 n^2,即等于一个 n@n 矩阵的自由度. 与其他一些水印算法要求原始图像来提取水印不同的是,SVD算法需要上面的三个矩阵来提取水印,但没有要求额外的信息量.

W(原始水印)和 W^*(提取的水印)的相似性通过相关检验来衡量. 将 W 和 W^* 看作一维向量,并按标准方法计算两者的相关系数 c(W, W^*). 对二维水印(如公司的标志图像),将其映射为一维向量,或直接计算它们的二维相关系数.

213　误差估计

在图像水印的过程中,有两个重要问题需仔细考虑:

(1)原始图像和水印图像的差别如何量化;

(2)如何确定嵌入的水印的信息量或能量. 这两个问题并不孤立,实际上是如何进行添加水印后的误差估计,而这为目前大多数文献所忽略.

下面我们试图探讨这个问题.

定义1: 设矩阵 AI $R^{n@n}$,其谱范数(也称22范数)定义如下:

$$+ A +_2 = \sqrt{K_{max}} = s_{max} \tag{4}$$

其中 K_{max}和 s_{max}分别表示 A^TA 的最大特征值及A 的最大奇异值.

引理 1: 若 UI $R^{n@n}$和 VI $R^{n@n}$为正交矩阵,且 A I $R^{n@n}$,则

$$+ UAV +_2 = + A +_2 \tag{5}$$

引理2: 设 AI $R^{n@n}$, DA 为矩阵A 的一个扰动,定义 A= A+ DA. 令矩阵 A 和A 按递减排列的第 i 个奇异值分别为 $s_i(A)$和 $s_i(\tilde{A})$,则

$$| s_i(A) - s_i(\tilde{A}) | \le + DA +_2, \quad i = 1, 2, , , n \tag{6}$$

引理2也称为奇异值扰动定理. 根据上面的定义及引理,可以得到如下结果:

定理1: 如果 A、A、W 和$s_i(\#)$的定义如上,则有

$$| s_i(\tilde{A}) - s_i(A) | \le a + W +_2, \quad i = 1, , , n \tag{7}$$

证明: 由式(2)、(4)、(5)和(6)有

$$\begin{aligned} | s_i(\tilde{A}) - s_i(A) | &= | s_i(S_1) - s_i(S) | \\ &= | s_i(S + aW) - s_i(S) | \\ &\le a + W +_2 \end{aligned}$$

从式(7)可以看到用 + W+ $_2$ 可衡量图像 A 和A 之间的误差. 这样通过调节水印的谱范数使得到的水印图像在鲁棒性和可觉察性之间达到一个平衡. 一个最简单的方法是直接调节系数 a 的值. 定理1为我们选择水印,确定水印的嵌入位置以及控制水印的能量提供了理论指导. 这些在实际应用中有重要意义的信息是目前很多水印算法不能提供的.

3 实验结果

用灰度图像仔细研究 SVD 算法对各种图像失真的抵抗能力,并将其与 Cox 的扩展谱方法[10]做比较,以检验新算法的鲁棒性能.结果显示 SVD 方法非常鲁棒.算法用了多种不同的图像和图像类型做了测试,但限于篇幅,这里只给出了使用图像 Lena 的鲁棒性测试结果.该图像大小为 200@200,有 256 个灰度级,像素值介于[0,1]之间.测试结果包含以下五个方面:加噪声、低通滤波、JPEG 压缩、图像裁剪和旋转.

与 Cox 方法类似,本文选取的水印是一个服从高斯分布的伪随机数组成的 2500@1 向量.在用 SVD 方法叠加水印时,水印向量被映射为一个 50@50 的矩阵.而在 Cox 方法中,水印向量则直接叠加到图像 DCT 域上幅值最大的前 2500 个系数上(不包括直流分量).在 Cox 方法中,控制水印嵌入能量的调节系数 a 的值设置为 0.1(Cox 等人所采用的一个典型值[10]).使用 50 个 2500@1 的随机向量作为测试用的水印,其中只有第 10 个为被嵌入到图像中的正确水印.原始图像和水印图像的相似性度量由两者的二维相关系数 e_c 来评价.用 Cox 方法得到的 e_c 值为 0.9957.

图 1 显示用 SVD 方法对图像 Lena 嵌入数字水印的结果.图 1(a)是原始图像,添加水印后的图像显示在图 1(b)中,图 1(c)为放大 64 倍后的绝对差图像,尽管经放大后图 1(a)与图 1(b)的差别显而易见,但人眼却难以区别两幅图之间的不同.为了使 SVD 和 Cox 两种方法创建的水印图像具有可比性,SVD 方法中尺度因子 a 值设置为 0.2,相应的两幅图像之间的相关系数 e_c 的值为 0.9966(即与 Cox 方法的 e_c 值接近).从图 1(c)中我们注意到差图像显示出原始图像的纹理特征,表明图像中信息量多的部分添加的水印信息量也多,信息量少的部分添加的信息量少.

图 1 SVD 方法对图像 Lenna 的数字水印
(a)原始图像,(b)添加水印后的图像,(c)绝对误差图像

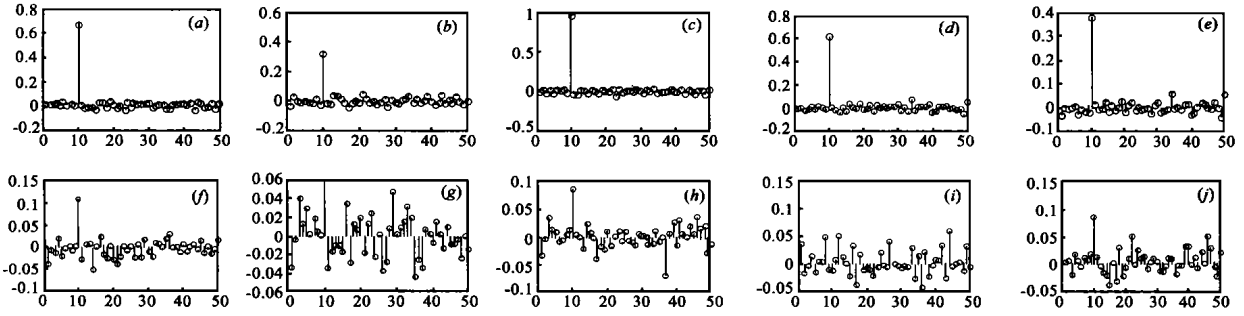

图 2 SVD 和 Cox 方法的鲁棒性测试结果以及比较.(a)~(e)为 SVD 方法分别对叠加高斯噪声、低通滤波、JPEG 压缩、图像旋转以及图像裁剪后的测试结果;(f)~(j)为相应的用 Cox 方法得到的测试结果

表 1 测试方法及其参数

测试方法	参 数
叠加高斯噪声	均值为 0,方差为 0.05
低通滤波	Gaussian 低通滤波器,大小为 16@16,方差为 4
JPEG 压缩	压缩质量系数 5%,压缩比 18.1
图像旋转	旋转 30°,并进行裁剪以保持原始图像大小
图像裁剪	裁剪图像的左半边

图 2 为 SVD 和 Cox 方法的鲁棒性测试结果以及比较.先按照式(2)得到嵌入水印后的图像)图 1(b),然后对图像分别进行叠加高斯噪声、低通滤波、JPEG 压缩、图像旋转以及图像裁剪等操作,得到失真的带水印图像.然后在水印检测过程中,从失真的带水印图像中提取出遭到破坏的水印 W^*,并计算 W^* 和 W(原始水印)之间的相关系数 c(W, W^*),检测结果显示于图 2(a)~(e).图中纵坐标表示相关系数值,横坐标表示 50 个 2500@1 的水印.5 个测试结果清楚表明第 10 个(即真正的水印)的相关值均大于其他的值,这表明可检测到正确的水印.实验的参数设置见表 1.

作为比较,本文也用 Cox 方法得到带水印的图像,然后对带水印的图像进行同样的操作,并计算水印的相关系数.图 2(f)~(j)为 Cox 方法的测试结果.显然用 Cox 方法得到的水印图像在经过低通滤波和图像旋转后,水印完全被破坏了,而对其它三种鲁棒性测试,得到的相关系数值均比较小,这表明水印在经过这些操作后实际上已经有非常大的失真.

在图像的尺寸大小变化和二维可视化水印(如公司的图像标志等)等方面,本文也做了相应的鲁棒性测试.所有这些结果均表明即使水印图像经过比较严重的失真,SVD 方法仍然能够提取出正确的水印或确定水印的存在与否.结果也同时说明新算法要比通用的 Cox 方法鲁棒得多.

4 结论

本文提出了一个新的数字水印算法.水印被叠加到原始图像的 SVD 域上.该算法的数学背景非常清晰,而且水印图像和原始图像之间的误差容易估计.因此一些重要的问题如水印的叠加位置的确定,水印的叠加能量和容量的控制都可以容易地解决.大量的实验数据与 Cox 方法相比较,表明新方法是非常鲁棒的.对静止图像来讲,SVD 方法是一项很有前途的数字水印技术.

致谢:在本文的完成过程中得到很多人的热心帮助.作者

在与王蕴红博士及博士生孙洪赞和冯涛的讨论中受益非浅,在此表示感谢.本文的工作已申请专利(专利号 99107964.7).

参考文献:

[1] B. R. Macq and I. Pitas. Special issue on watermarking [J]. Signal Processing, 1998, 66(3): 281- 282.

[2] M. D. Swanson, M. Kobayashi and A. H. Tewfik. Multimedia data-embedding and watermarking technologies [A]. Proceedings of the IEEE [C]. 1998: 86(6): 1064- 1087.

[3] E. Koch, J. Rindfrey, and J. Zhao. Copyright Protection for Multimedia Data [A]. Proc. of the International Conference on Digital Media and Electronic Publishing [C]. Dec. 1994: 6- 8.

[4] S. H. Low, N. F. Maxemchuk and A. M. Lapone. Document identification for copyright protection using centroid detection [J]. IEEE Trans. on Communications, 1998, 46(3): 372- 383.

[5] J. M. Acken. How watermarking adds value to digital content [J]. Communications of the ACM, 1998, 41(7): 74- 77.

[6] N. Nikolaidis, I. Pitas. Robust image watermarking in the spatial domain [J]. Signal Processing, 1998, 66(3): 385- 403.

[7] R. G. V. Schyndel, A. Z. Tirkel and C. F. Osborne. A Digital Watermark [A]. Proc. of ICIP. 94 [C]. 1994, 2: 86- 90.

[8] N. Nikolaidis and I. Pitas. Copyright protection of images using robust digital signatures [A]. Proc. of ICASSP. 96 [C]. 1996, 4: 2168- 2171.

[9] V. Darmstaedter, J. F. Delaigle, J. J. Quisquater and B. Macq. Low cost spatial watermarking [J]. Computers & Graphics, 1998, 22(4): 417- 424.

[10] I. J. Cox, J. Kilian, F. T. Leighton and T. Shamoon. Secure Spread Spectrum Watermarking for Multimedia [J]. IEEE Trans. on Image Processing, 1997, 6(12): 1673- 1687.

[11] M. D. Swanson, B. Zhu and A. H. Tewfik. Transparent robust image watermarking [A]. Proc. of ICIP. 96 [C]. 1996, 3: 211- 214.

[12] C. T. Hsu and J. L. Wu. Hidden digital watermarks in images [J]. IEEE Trans. on Image Processing, 1999, 8(1): 58- 68.

[13] M. D. Swanson, B. Zhu and A. H. Tewfik. Multiresolution scene-based video watermarking using perceptual models [J]. IEEE Journal on Selected Areas in Communications, 1998, 16(4): 540- 550.

[14] J. O. Ruanaidh, W. Dowling and F. Boland. Phase watermarking of digital images [A]. Proc. of ICIP. 96 [C]. 1996, 3: 239- 242.

[15] D. Kundur and D. Hatzinakos. A robust digital image watermarking method using wavelet-based fusion [A]. Proc. of ICIP. 97 [C]. 1997, 1: 544- 547.

[16] C. T. Hsu and J. L. Wu. Multiresolution watermarking for digital images [J]. IEEE Trans. on Circuits and Systems II-Analog and Digital Signal Processing, 1998, 45(8): 1097- 1101.

[17] G. h. Berbecel, T. Cooklev and A. N. Venetsanopoulos. Multiresolution technique for watermarking digital images [A]. Proc. of ICCE. 97 [C]. Jun. 1997: 354- 355.

[18] A. Piva, M. Barni, F. Bartolini and V. Cappellini. DCT-based watermark recovering without resorting to the uncorrupted original image [A]. Proc. of ICIP. 97 [C]. 1997, 3: 520- 523.

作者简介:

刘瑞祯　1969 年出生,1986 年考入北京航空航天大学,1998 年进入中科院自动化所模式识别室读博士.研究兴趣包括信息隐藏和数字水印,图像处理与编码,计算机视觉等领域.

谭铁牛　1964 年出生,1980 年考入西安交通大学,1989 年获英国伦廓大学帝国理工学院博士学位。现为中科院自动化研究所所长,模式识别国家重点实验室主任、研究员.从事图像处理、计算机视觉和模式识别等相关领域的研究工作,现已在主要的国际学术期刊和国际学术会议上发表论文七十多篇,申请专利 5 项.