

文章编号 : 1007-130X(2006)10-0061-02

一种基于线性神经网络的半易损水印技术

A Linear-Neural-Network-Based Semi-Fragile Watermarking Technique

余艳玮, 卢正鼎

YU Yan-wei, LU Zheng-ding

(华中科技大学计算机科学与技术学院, 湖北 武汉 430074)

(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

摘要: 本文提出了一种基于线性神经网络的半易损水印技术。该算法利用线性神经网络对小波低频系数及其邻域系数的关系进行建模, 然后利用被选的嵌入点处的小波系数与模型输出值之间的大小关系来嵌入二值水印。基于该模型提取水印时不需要原始图像和水印, 只需模型的参数即可。实验表明, 该算法不仅对于恶意篡改有较强的定位能力, 而且可以抵抗 JPEG 压缩。

Abstract: A linear neural-network-based semi-fragile watermarking technique is proposed in the paper. The method establishes a relational model among these coefficients and other coefficients in their vicinities in the low frequent band using linear neural networks. Based on the linear neural network model, watermarks can be extracted only with the parameters of the model without the participation of the original image and watermarks. Experimental results show the proposed method can not only locate the tempered areas exactly but also resist the JPEG compression.

关键词: 半易损性水印; 离散小波变换; 神经网络

Key words: semi-fragile watermarking; discrete wavelet transform; neural network

中图分类号: TP309

文献标识码: A

1 引言

随着计算机和网络技术的飞速发展, 多媒体信息可能会遭受各类无意或有意的篡改攻击, 这使得人们对数字媒体的完整性和内容的真实性产生质疑。如果篡改涉及到国家安全、法庭举证等重要内容时, 可能会造成不良的社会影响或重大政治经济损失。因此, 对数字媒体内容的真实性、完整性实施有效保护已成为一个严峻的现实问题。数字水印中的易损性数字水印为上述问题提供了一个很好的解决方案^[1]。它利用人类知觉系统的冗余, 在不影响数字媒体感官(视或听)质量的前提下, 将与媒体内容相关或不相关的标志信息作为水印直接嵌入媒体内容中。当媒体内容需认证时, 可将水印提出以鉴定其是否真实完整。

近年来, 国内外有关易损性数字水印的论述日益增加。

根据识别篡改的能力, 一般把易损性数字水印分为完全易损性水印和半易损性水印两种^[2]。完全易损性水印技术^[3,4]主要应用于保护那些要求极其精密的数字媒体, 如医学图像等。由于完全易损性水印随着数字媒体内容的改变而破坏, 使得通过认证系统可以识别数字媒体任何像素值的改变, 其中包括正常的图像处理(如压缩等)。半脆弱水印^[5,6]是同时具有易损性水印和鲁棒性水印两种功能的水印技术, 它在容忍一定程度的常见信号处理操作的同时, 还具有与完全易损水印相似的定位篡改区域的能力。

文献[5]中介绍了一种基于矢量量化压缩的半易损水印算法, 虽然可以大致定位可能被篡改的位置, 但不能精确刻画出篡改的形状。文献[6]给出了一种基于小波系数对之间正负关系的半易损水印系统, 通过比较小波系数带 LH_2 和 HL_2 中相对应的系数之间的大小来判定当前小波系数对其中嵌入的水印码值。虽然在水印检测的过程中不

* 收稿日期: 2004-12-08; 修订日期: 2005-06-02

基金项目: 电子信息产业发展基金资助项目(04C26211100937); 科技型中小企业创新基金资助项目(04C26214201284); 国家自然科学基金资助项目(60502024)

作者简介: 余艳玮(1981-), 女, 湖北汉川人, 博士生, 研究方向为数字水印; 卢正鼎, 教授, 博士生导师, 研究方向为分布式异构系统集成、信息安全、数字水印、数据挖掘等。

通讯地址: 430074 湖北省武汉市华中科技大学计算机科学与技术学院; Tel: (027) 87544285; E-mail: yuyanwei198112@163.com

Address: School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P. R. China

需要原始水印就可以给出相应的检测水印结果。但是,在对可能的水印篡改攻击进行定位的时候,却需要比对原始水印和检测水印之间的异同,来给出相应的篡改矩阵,从而定位出可能被篡改的位置和形状。文献[7]提出了一种基于BP神经网络的脆弱水印技术,它利用BP神经网络建立图像的像素点与它们邻域中的其它像素点之间的关系模型,然后通过调整被选像素点与模型输出值之间的大小关系来嵌入水印图案的位信息。由于该技术中水印是直接对像素点进行嵌入的,所以比较脆弱,连最基本的JPEG压缩处理都抵抗不了;而且,由于BP神经网络的训练速度通常很慢,很容易陷入局部极小点等缺点,因此在水印系统中的实用性不是很好。本文提出了一种基于线性神经网络(LNNB)的小波域的半易损水印算法。

2 LNNB 水印算法

线性神经网络是最简单的一种神经网络,它由一个或多个线性神经元构成。线性神经网络采用线性函数作为传递函数,并采用基于最小二乘算法(LMS)的Widrow-Hoff学习规则来调节网络的权值和阈值。此外,采用newlind函数还可以直接根据网络的输入矢量和目标矢量设计出期望的线性网络,来改善线性神经网络的收敛速度和精度。虽然线性神经网络通常用来描述输入、输出间的线性映射关系,但是考虑到线性网络的训练与其他功能更强大的神经网络(如BP网络等)相比计算量要小得多,而且它对非线性映射最佳线性拟合的性能也较好,因此本文采用线性神经网络来最佳线性拟合输入和输出间的非线性映射。通过线性神经网络,对小波低频系数及其邻域系数的关系进行建模,然后利用被选的嵌入点处的小波系数与模型输出值之间的大小关系可以实现二值水印的嵌入。

小波变换(DWT)是变换域水印技术中用到的变换方法之一。与DCT相比,DWT有很多独特的优点,如可以保证在“JPEG 2000”有损压缩下水印不会被去除;图像的局部特征如局部纹理、亮度等,对于图像分析和处理是非常关键的,经过小波变换仍可保留住。考虑到小波分解后的低频部分是能量的集中部分,本文采用将水印嵌入到DWT域中的低频部分。而且,由于小波变换是一种空间-频域分析方法,每个DWT系数反映了图像在局部空域和局部频域上的内容。因此,利用小波变换的这个特点能够检测到图像被篡改的区域。

2.1 水印的嵌入

设原始图像 I 为大小为 $n \times n$ 的灰度图像,水印 W 为大小为 $m \times m$ ($m \leq n/6$) 的二值图像。水印嵌入和提取过程可描述为:

- Step1 基于混沌序列调制二值图像水印 W 。
出于安全的考虑,水印在嵌入前要被调制成杂乱无章的数据。考虑到混沌序列对初始值的极度敏感性,故对水印进行混沌调制。设最后调制得到的水印为 w^* , 则

$$W^* = \{w^*(i,j) | w^*(i,j) = w(i,j) \oplus i, i, j, 1 \dots m\}$$
- Step2 对原始图像 I 进行小波分解。
用 $DWT(I, l)$ 表示对图像 I 进行 l 级小波分解, L 为小波分解后的低频块, $l(i,j)$ 为 L 块的位置 (i,j) 处的低频小波系数, 其中 $1 \leq i, j \leq n/2^l$ 。
- Step3 利用线性神经网络学习低频子带处的小波系数与其邻域的小波系数间的关系。
按行、列依次不重叠地选取小波分解后低频部分 (i,j) 处的系

数值 $t(i,j)$ 及其 3×3 邻域系数 $X(i,j) = \{l(i-1, j-1), (i-1, j), l(i-1, j+1), l(i, j-1), l(i, j), l(i, j+1), l(i+1, j-1), l(i+1, j), l(i+1, j+1)\}$ 形成学习样本集, 每个样本包括输入向量 $X(i,j)$ 和期望输出量 $t(i,j)$ 。通过线性神经网络模型, 可以由 (i,j) 处 3×3 邻域的系数 $X(i,j)$ 映射出位置 (i,j) 处的小波系数值 $y(i,j)$ 。

Step4 嵌入水印。

通过修改 $t(i,j)$ 来嵌入水印, 如果 $t(i,j)$ 所对应的水印位 $w^*(i,j)$ 为 1, 则调整 $t(i,j)$ 使之大于线性神经网络模型的输出 $y(i,j)$; 否则, 调整 $t(i,j)$ 使之小于线性神经网络模型的输出 $y(i,j)$;

$$t(i,j) = \begin{cases} \max(t(i,j), y(i,j) + \epsilon), w^*(i,j) = 1 \\ \min(t(i,j), y(i,j) - \epsilon), w^*(i,j) = 0 \end{cases} \quad (1)$$

Step5 反小波变换。

载体图像的低频部分嵌入水印后, 对载体图像小波系数块进行重组, 然后进行相应尺度的反小波变换得到加入了水印的图像。

在对嵌入了水印的图像进行主观评价的同时, 本文实验采用峰值信噪比 (PSNR) 作为图像质量的客观评价标准。

2.2 水印的提取

水印的提取过程除了不要求神经网络的训练和反小波变换外, 可看作是嵌入的逆过程:

Step1 把待验证的图像 \tilde{I} 经小波变换分解, 得到低频小波系数。

Step2 选取嵌入水印位信息的低频小波系数 $\tilde{t}(i,j)$ 及其 3×3 邻域系数 $\tilde{X}(i,j)$ 。根据线性神经网络模型, 如果 $\tilde{t}(i,j)$ 大于模型的输出 $\tilde{y}(i,j)$, 则水印位 $w(i,j)$ 为 1; 否则, $w(i,j)$ 为 0:

$$w(i,j) = \begin{cases} 1, \tilde{t}(i,j) > \tilde{y}(i,j) \\ 0, \tilde{t}(i,j) \leq \tilde{y}(i,j) \end{cases} \quad (2)$$

Step3 对 w 进行混沌反调制, 得到提取出的水印 \tilde{W} 。

利用篡改评估函数 $TA F$ 来判断检测图像的可信度。

$$TA F(W, \tilde{W}) = \frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m w(i,j) \oplus \tilde{w}(i,j) \quad (3)$$

其中, w 为嵌入的水印, \tilde{w} 为提出出来的水印, \oplus 表示异或运算。

为了判断图像是否被恶意篡改, 给定阈值 θ 。如果 $TA F(W, \tilde{W}) > \theta$, 则表明图像被恶意篡改; 如果 $0 < TA F(W, \tilde{W}) < \theta$, 则认为图像的失真是一些合理的图像处理的结果; 如果 $TA F(W, \tilde{W}) = 0$, 则表示认证图像和水印图像完全一致。阈值 θ 可由用户根据应用需要而设定, 本文中的 θ 取 0.2。

3 实验结果

本文对原图像进行尺度为 1 的小波变换。线性神经网络模型的学习速率参数取最快的稳定学习速率值的 1.8 倍, 再利用它对小波低频系数及其 3×3 邻域中的小波系数的关系进行建模。另外, 式(1)中的 ϵ 取 10。

图 1a 所示为 512 × 512 的灰度 Lena 原图, 图 1b 所示为嵌有水印的图像, 图 1c 所示为 85 × 85 的二值水印图像。从图 1b 可以看到嵌入水印后的水印图像质量无明显下降, 水印的不可感知性非常好。



图 1 水印嵌入结果

3.1 对 JPEG 压缩的抵抗能力

对压缩后的嵌有水印的图像进行检测, 结果如表 1 所示 (下转第 65 页)

的角度设计了一种基于熵的遗传聚类分割算法。通过和其他聚类算法以及标准遗传算法的结果进行比较,证明了算法是可行而且有效的。

参考文献:

[1] Dorin Comaniciu, Peter Meer. Image-Guided Decision Support System for Pathology[J]. Machine Vision and Applications, 1999, 11(4): 213-224.

[2] J Bosworth, S T Acton. The Morphological Lomo Filter for Multiscale Image Processing[A]. Proc of the IEEE Int'l Conf on Image Processing[C]. 1999. 24-28.

[3] J C Rajapakse, J N Gedde. Statistical Approach to Segmentation of Single-Channel Cerebral MR Image[J]. IEEE Trans on Medical Imaging, 1997, 16(2): 176-186.

[4] J H Holland. Adaptation in Natural and Artificial Systems[M]. The University of Michigan Press, 1975.

[5] H M G Velasco, C J G Orellana, F J L Aligué et al. Segmentation of Bovine Livestock Images Using GA and ASM in a Two-Step Approach[J]. International Journal of Pattern Recognition & Artificial Intelligence, 2003, 17(4): 601-616.

[6] A S Abutaleb, M Kamel. A Genetic Algorithm for the Estimation of Ridges in Fingerprints[J]. IEEE Trans on Image Processing, 1999, 8(8): 1134-1139.

[7] Y A Alsultanny, M M Aqel. Pattern Recognition Using Multilayer Neural-Genetic Algorithm[J]. Neuro Computing, 2003, 51: 237-247.

(上接第 5 页)

参考文献:

[1] IBM, Microsoft, BEA. Web Services Atomic Transaction (WS-Atomic Transaction) [EB/OL]. <http://www.ibm.com/developerworks/library/ws-atomtran/>, 2003-09.

[2] S Dalal, S Temel, M Little, et al. Coordinating Business Transactions on the Web[J]. IEEE Internet Computing, 2003, 7(1): 30-39.

[3] 范国钟, 钟华, 黄涛, 等. Web 应用服务器研究综述[J]. 软件学报, 2003, 14(10): 1728-1738.

(上接第 62 页)

示。从实验数据可发现,当质量品质下降到 60 时, TAF 的值仍没有大的变化,可见本文算法对于 JPEG 压缩有很好的抵抗力。

表 1 对 JPEG 压缩的抵抗能力

质量品质	PSNR(dB)	TAF
90	52.765 452	0.073 356
80	49.386 013	0.073 91
70	48.110 818	0.083 183
60	42.661 680	0.098 27
50	38.871 544	0.162 35

3.2 对恶意篡改的定位能力

对图像的恶意篡改主要有剪切、替换等操作。

图 2 是将嵌有水印的图像进行剪切后,用黑块替补处理后的结果。图 3 是嵌有水印图像的左上角小块被替换的结果。可以看出,本文算法对剪切、替换有很好的易损性,

而且对篡改区域有较强的定位能力。



图 2 剪切处理后的结果



图 3 替换处理后的结果

4 结束语

本文提出了一种基于线性神经网络的半易损水印技术。与张军^[7]等提出的水印技术相比,该技术的优点表现在:(1) 线性神经网络的训练速度比较快,对于输入、输出间的最佳线性拟合的性能也很好;(2) 对于恶意篡改比较敏感,而且能较准确地定位篡改区域;(3) 对于 JPEG 压缩有较强的鲁棒性。但是,该技术还无法智能地识别篡改的类型,需要进一步改善。这也正是下一步研究工作要实现的目标。

参考文献:

[1] F P A Petitcolas, R J Anderson, M G Kuhn. Information Hiding: A Survey [J]. Proceedings of the IEEE, 1999, 87(7): 1063-1077.

[2] C Y Lin, S F Chang. Semi-Fragile Watermarking for Authentication JPEG Visual Content [A]. Proc of SPIE 3971 [C]. 2000. 140-151.

[3] F Sattar, B Barkat. A New Time-Frequency Based Fragile Watermarking Scheme for Image Authentication[J]. IEEE Trans on Communications, 2003, 61(2): 992-995.

[4] J Hu, J Huang, D Huang, et al. Image Fragile Watermarking Based on Fusion of Multi-Resolution Tamper Detection [J]. Electronics Letters, 2002, 38(24): 1512-1513.

[5] Z M Lu, C H Liu, D G Xu, et al. Semi-Fragile Image Watermarking Method Based on Index Constrained Vector Quantization[J]. Electronics Letters, 2003, 39(1): 35-36.

[6] 易开祥, 孙鑫, 王道顺, 等. 一种新的基于图像内容的半易损水印技术[J]. 系统系统仿真学报, 2002, 14(9): 1200-1207.

[7] 张军, 王能超. 用于图像认证的基于神经网络的水印技术[J]. 计算机辅助设计与图形学学报, 2003, 3(3): 307-312.