

一种基于 DWT 的非自适应数字水印算法及其 MATLAB 实现

An Unadaptive Digital Watermarking Algorithm Based on DWT and Its Realization on MATLAB

黄福莹, 李 丽

HUANG Fu-ying, LI li

(广西大学计算机与电子信息学院, 广西南宁 530004)

(School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi, 530004, China)

摘要:介绍一种基于离散小波变换(DWT)的非自适应数字水印算法,并应用 MATLAB 数学软件实现算法,考察算法的有效性和抗攻击能力。该算法对原始图像进行3级小波变换,将水印信息嵌入到原始图像第3级小波变换的对角线系数分量,通过调整嵌入强度因子来获得满意的水印嵌入效果。该算法对 JPEG 压缩、椒盐噪声和剪切等攻击有很好的鲁棒性。

关键词:数字水印 算法 图像 离散小波变换 MATLAB

中图分类号: TN911.73 **文献标识码:** A **文章编号:** 1002-7378(2009)04-0254-03

Abstract: The paper introduces an unadaptive digital watermarking algorithm based on DWT (discrete wavelet transform) that is realized it on matlab. Both the validity and the attack resistance of the algorithm are tested in the paper. In the algorithm, the original image is divided by three-level wavelet transform and the information of watermarking is embedded into the diagonal coefficient of the third-level wavelet transform. The satisfactory result of the watermarking embedding is gained through adjusting a intensity coefficient. The algorithm is robust enough to some image attacking processing as JPEG compressing, salt & pepper noise and cutting.

Key words: digital watermarking, algorithm, image, DWT, MATLAB

随着计算机网络通信技术特别是互联网的快速发展,数字产品的使用与传播越来越普及。但是,数字产品很容易遭受非法复制、伪造、篡改、侵权等攻击。传统的基于密码学的安全措施对非法攻击显得力不从心,出于对数字产品本身及其所有者利益的保护,数字水印技术孕育而生。数字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,以起到保护版权、秘密通信、数据文件的真伪鉴别和产品标志等作用^[1]。数字水印技术自1993年被提出以来,由于其不容易被接收者或使用者的察觉,不影响原产品的视觉质量,能够抵御外界干扰,具有很好的可靠性(或称鲁棒性),而且不容易被

复制、篡改和伪造,不容易被非法检测和解码消除,所以在信息安全和经济上占有重要地位,已经在数字产品版权保护、票据防伪、数据隐藏、隐藏通信等方面得到广泛应用。数字水印的分类方法有很多种,按数字水印的隐藏位置划分,可以将其分为空间域数字水印和变换域数字水印。空间域数字水印是直接信号空间上叠加水印信息。变换域水印是在离散余弦变换(DCT)域、离散傅里叶变换(DFT)域和离散小波变换(DWT)域上隐藏水印。从总体性能来看,变换域的方法比空间域的方法有更多的优越性。又由于DWT具有很好的时域-频域特性,在DWT域嵌入数字水印具有明显的优点,所以基于DWT域的数字水印研究是近几年数字水印算法研究领域的热点^[2]。

本文研究基于DWT的非自适应数字水印算法

收稿日期:2009-10-10

作者简介:黄福莹(1972-),男,讲师,主要从事数字图像处理研究。

及其 MATLAB 数学软件实现。所谓非自适应嵌入方式,是指水印的嵌入强度始终为常数而不随系数的不同而改变的嵌入方式。它一般分为非自适应加性嵌入方式和非自适应乘性嵌入方式两种,本文采用的是非自适应加性嵌入方式。

1 基于 DWT 的非自适应数字水印算法

基于 DWT 的非自适应数字水印算法的基本思想是:对原始图像进行3级小波变换,然后将水印信息嵌入到原始图像第3级小波变换的对角线系数分量,通过调整嵌入强度因子,以获得满意的水印嵌入效果。

1.1 小波分解

原始图像经过一级小波分解后可以分成4个频率带,分别是低频 LL1、水平 HL1、垂直 LH1 及对角线 HH1。以此类推,第二小波分解对低频频率带 LL1进一步分解成4个频率带。我们将原始图像进行3级小波分解,得到 HL_i, LH_i, HH_i 和 LL_3 共10个频率带,其中 $i=1,2,3, LL_3$ 为低频频率带(详见图1)。

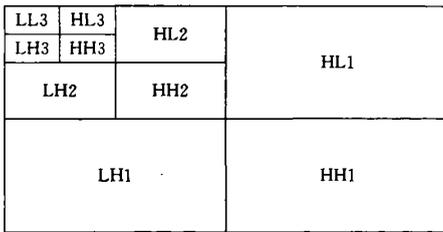


图1 原始图像小波分解

1.2 水印嵌入过程

将水印信息嵌入到第3级小波分解后的对角频率带分量,即图1中的 HH_3 。水印嵌入步骤如下。

(1)选择原始图像第3级小波分解后的对角线分量 HH_3 作为数字水印嵌入位置。水印嵌入算法的公式为

$$I' = I + \alpha\omega, \tag{1}$$

其中, I 和 I' 是水印嵌入前后小波变换系数; α 是水印嵌入的强度因子,用来控制水印嵌入的强度; ω 是嵌入的数字水印。

(2)对已嵌入水印信息的图像进行反小波变换,得到含水印图像。

水印嵌入模型如图2所示。

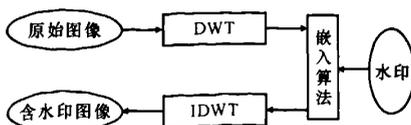


图2 水印嵌入模型

1.3 水印提取过程

水印的提取是水印嵌入的逆过程,水印提取步骤如下。

(1)分别将原始图像和含水印图像进行3级小波分解,选择它们的第3级小波分解的对角线系数。

(2)根据水印提取算法,提取出水印 ω' , 提取算法的公式为

$$\omega' = (I' - I)/\alpha. \tag{2}$$

水印提取模型如图3所示。

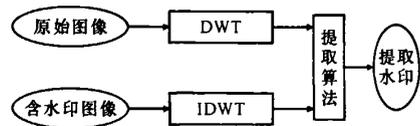


图3 水印提取模型

2 算法的 MATLAB 实现

我们选取一幅 256×256 的灰度图像作为原始图像,一幅 32×32 的二值图像作为水印图像,水印嵌入强度因子 $\alpha = 0.06$ 。应用 MATLAB 数学软件实现基于 DWT 的非自适应数字水印算法,考察该算法的有效性和抗攻击能力。

2.1 水印嵌入和提取

由图4的结果可知,实验得出含水印图像同原始图像之间的 $PSNR = 43.1867\text{dB}$, 已是远大于人的肉眼可以容忍图像的 20dB $PSNR$ 值^[3], 它们之间凭肉眼看不出任何失真。采用相关系数 NC 评价水印的质量,按照“ NC 值大于等于 0.750 即为有效水印”的判定原则^[3], 从未经任何处理的含水印图像提取出的水印与原始水印相似度很高 ($NC = 1.0046$), 主观上感觉不到差异。



图4 水印嵌入和提取结果

(a)原始图像;(b)含水印图像($PSNR = 43.1867\text{dB}$);(c)原始水印图像;(d)提取出的水印图像($NC = 1.0046$)。

2.2 水印攻击检测

对已嵌入水印的图像进行 JPEG 压缩、添加椒盐噪声、剪切和中值滤波攻击,提取出来的水印如图 5 所示。

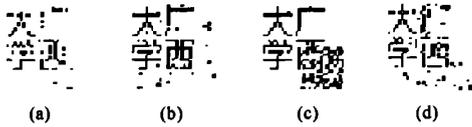


图5 含水印图像受攻击后提取的水印

(a)QF=50%的 JPEG 压缩提取的水印(NC=0.8288); (b)添加椒盐噪声提取的水印(NC=0.9965); (c)裁剪右下角1/4提取的水印(NC=0.7755); (d)3×3中值滤波提取的水印(NC=0.4998)。

从图5可以看出,本文所实现的数字水印算法对 JPEG 压缩、椒盐噪声和裁剪有很好的鲁棒性,对中值滤波的鲁棒性一般,但是肉眼也能识别出来。

3 结束语

本文实现了一种基于 DWT 的非自适应数字水

印算法,算法的 MATLAB 实现结果表明,该算法简单易懂,有一定的抗攻击能力。但是在实验中发现该算法对高斯噪声的鲁棒性很差,添加高斯噪声后水印基本检测不出来,还需要对该算法做进一步完善。

参考文献:

[1] 杨义先. 数字水印基础教程[M]. 北京:人民邮电出版社,2007.

[2] 董长虹,赖志国,余啸海. Matlab 图像处理与应用[M]. 北京:国防工业出版社,2004.

[3] 金聪. 数字水印理论与技术[M]. 北京:清华大学出版社,2008.

[4] 黄达人,刘九芬,黄继武. 小波变换域图像水印嵌入对策和算法[J]. 软件学报,2002,13(7):1290-1295.

(责任编辑:韦廷宗 邓大玉)

(上接第250页)

表3 算法的召回率与准确率试验结果

| 实验编号 | 相似句子数(个) | 召回率 | 准确率 |
|------|----------|------|------|
| 1 | 55 | 0.92 | 0.87 |
| 2 | 16 | 0.94 | 0.84 |
| 3 | 18 | 0.88 | 0.82 |
| 4 | 46 | 0.92 | 0.85 |

3.3 阈值对算法效果的影响

使用3.2中的第2组数据,改变阈值的大小来观察对算法召回率和准确率的影响结果如图1所示。从图1的实验结果表明,阈值在0.5~0.6的时候召回率和准确率达到一个较好的平衡,这和吴平博等[6]的试验结果基本一致。

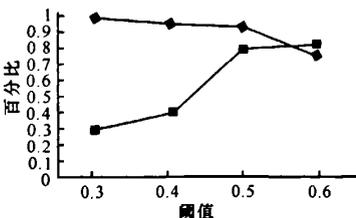


图1 阈值对算法效果的影响

—●—:召回率;—■—:准确率。

4 结束语

在语料库收集过程中进行相似句子的去重是非

常必要的,但是,目前国内尚缺少这类研究。本文尝试利用整体相似因子和局部相似因子计算句子的相似度,并借鉴 KMP 算法的匹配跳跃思想,提出中文字符串匹配类 KMP 算法,对于完全重复的和极其相似的句子去重有很好的效果。由于本文的相似句子去重算法没有考虑到语义上的相似,因此还需要进一步改进算法,以提高去重的准确率。

参考文献:

[1] 黄俊红,范云. 双语平行语料库对齐技术述评[J]. 外语电教化,2007(118):21-25.

[2] 陈桂林,王永成. 字串去重的快速算法研究[J]. 情报学报,2000,19(3):254-258.

[3] Dekang Lin, Patrick Pantel. DIRT-Discovery of inference rules from text [J]. Journal of Natural Language Engineering, Fall-Winter, 2001, 1: 1-6.

[4] 王莹莹. 中文短语相似度计算方法研究及应用[D]. 长沙:长沙理工大学硕士学位论文,2008:24-26.

[5] 严蔚敏,吴伟民. 数据结构(C语言版)[M]. 北京:清华大学出版社,1997:80-81.

[6] 吴平博,陈群秀,马亮. 基于特征申的大规模中文网页快速去重算法研究[J]. 中文信息学报,2003,17(2):28-35.

(责任编辑:邓大玉)