

用于版权保护与内容认证的半脆弱音频水印算法

王向阳^{1,2} 祁薇¹

摘要 提出了一种可同时用于版权保护与内容认证的半脆弱数字音频盲水印算法。该算法具有以下特点: 1) 结合音频自身特性, 不仅能够自适应划分音频数据段, 而且能够智能调节水印嵌入强度, 增强了数字音频水印的鲁棒性; 2) 能够同时进行数字音频的版权保护与内容认证, 并可大致确定篡改发生区域。仿真实验表明, 文中算法不仅具有较好的不可感知性, 而且对常规处理具有较好的鲁棒性, 同时能够对替换等恶意篡改作出报警并确定被篡改位置。

关键词 半脆弱水印, 数字音频, 自适应量化, 篡改检测

中图分类号 TP391

A Semi-fragile Audio Watermarking for Copyright Protection and Content Authentication

WANG Xiang-Yang^{1,2} QI Wei¹

Abstract Semi-fragile watermarking methods aim at detecting unacceptable malicious manipulations, while allowing for acceptable regular manipulations such as lossy compression. In this paper, a semi-fragile audio watermarking algorithm for copyright protection and content authentication is proposed. The features of the proposed algorithm are as follows: 1) The original audio signal can be adaptively divided, and the watermark depth can be intelligently adjusted according to human auditory masking and local audio character; 2) The copyright protection and content authentication for digital audio can be utilized simultaneously. Experimental results show that our semi-fragile audio watermarking scheme is not only inaudible and robust against various common images processing, but also fragile to malicious modification. Especially, it can detect effectively the tampered regions.

Key words Semi-fragile watermarking, digital audio, adaptive quantization, tamper detection

1 引言

作为数字音频作品知识产权保护的一种有效手段, 数字音频水印技术已成为国际学术界研究的一个热点。半脆弱音频水印技术结合了鲁棒音频水印

技术与脆弱音频水印技术的优点^[1]。与鲁棒音频水印相似, 半脆弱音频水印可以在一定程度上容忍施加于含水印 (利用半脆弱水印技术) 数字音频上的常见信号处理操作 (如叠加噪声、低通滤波、MP3 压缩等)。与脆弱音频水印相似, 半脆弱音频水印能够对恶意篡改做出报警反应且具有定位篡改区域的能力。显然, 半脆弱音频水印技术在网络时代具有更重要的应用价值^[2]。

近年来, 数字音频水印技术研究取得了很大进展, 并陆续提出了一系列优秀的数字音频水印嵌入方案^[3~5]。赵等^[6]给出了一种基于固定量化步长的混合域数字音频水印嵌入算法。全等^[7]以改进的心理声学模型为基础, 提出了小波包变换域脆弱型音频水印算法。

以上分析表明, 尽管数字音频水印技术研究已日益受到重视, 但可同时用于版权保护与内容认证的半脆弱音频水印算法研究尚未见报道。

本文以二值图像作为水印信号, 提出了一种可同时用于版权保护与内容认证的半脆弱数字音频盲水印算法。该算法不仅具有较好的不可感知性, 而且对常规处理具有较好的鲁棒性, 同时能够对替换等恶意篡改作出报警并确定被篡改位置。

收稿日期 2006-5-29 收修改稿日期 2006-12-25
Received May 29, 2006; in revised form December 25, 2006
视觉与听觉信息处理国家重点实验室 (北京大学) 开放基金 (0503), 信息安全国家重点实验室 (中国科学院软件研究所) 开放基金 (03-06), 大连市科技基金 (2006J23JH020), “图像处理与图像通信” 江苏省重点实验室 (南京邮电大学) 开放基金 (ZK205014), 江苏省计算机信息处理技术重点实验室 (苏州大学) 开放课题基金 (KJS0602) 资助
Supported by Open Foundation of State Key Laboratory of Vision and Auditory Information Processing (Peking University) (0503), Open Foundation of State Key Laboratory of Information Security of P. R. China (03-06), Natural Science Foundation of Dalian City of P. R. China (2006J23JH020), Open Foundation of Key Laboratory of Image Processing and Image Communication (Nanjing University of Posts and Communications) (ZK205014), Open Foundation of Jiangsu Province Key Laboratory for Computer Information Processing Technology (Soochow University) (KJS0602)
1. 辽宁师范大学计算机与信息学院 大连 116029 2. 北京大学视觉与听觉信息处理国家重点实验室 北京 100871
1. School of Computer and Information Technique, Liaoning Normal University, Dalian 116029 2. State Key Laboratory of Vision and Auditory Information Processing, Peking University, Beijing 100871
DOI: 10.1360/aas-007-0936

2 水印信号的嵌入

假设原始数字音频信号为 $A = \{a(i), 0 \leq i < Length\}$. 二值水印图像为 $W = \{w(i, j), 0 \leq i < M, 0 \leq j < N\}$. 其中, $w(i, j) \in \{0, 1\}$ 代表二值水印图像的第 i 行、第 j 列像素值. 半脆弱音频水印嵌入方案 (关键步骤) 可描述如下 (参见图 1).

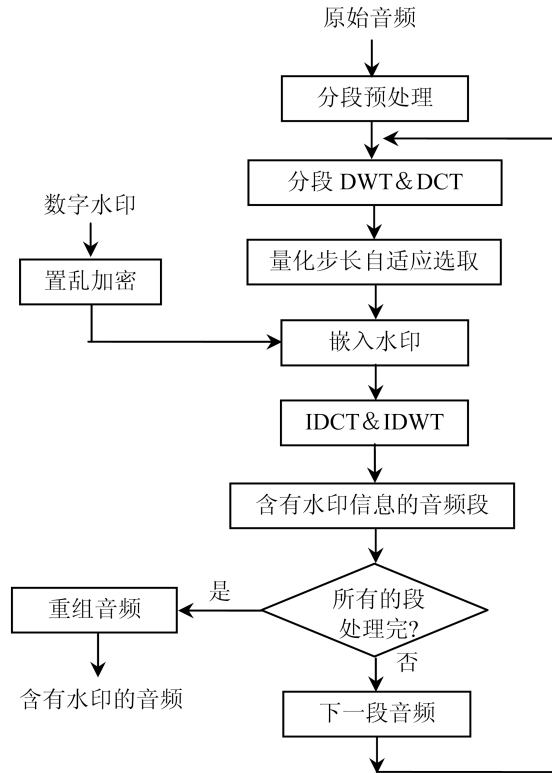


图 1 数字水印嵌入原理

Fig. 1 The principle of watermark embedding

2.1 预处理

为了消除二值水印图像的像素空间相关性, 本文引入 Arnold 变换^[3], 将二值水印图像 W 置乱加密为安全水印矩阵 W_1 .

为了有效克服由 MP3 压缩等攻击所引起的解压缩的数字音频文件长度与原始数字音频文件长度不等的问题, 我们采用对 M 个音频数据段的长度做自适应调整.

2.2 水印信号的嵌入

为了提高数字水印的嵌入与检测速度, 本算法将采用逐行嵌入水印信息的策略, 即每个音频数据段将嵌入二值水印图像的一行水印信息.

1) 音频数据段的离散小波变换 (DWT). 对音频数据段 $A(i)$ 进行 H 级小波变换, 并得到小波系数 $A(i)^H, D(i)^H, D(i)^{H-1}, \dots, D(i)^1$. 其中,

$A(i)^H$ 为音频段 $A(i)$ 小波变换的第 H 级近似分量; $D(i)^1, \dots, D(i)^{H-1}, D(i)^H$ 分别为音频段 $A(i)$ 小波变换的第 $1 \sim H$ 级的细节分量.

2) 近似分量的离散余弦变换 (DCT). 考虑到小波变换域的近似分量能够有效抵御各种干扰, 这里选择 H 级小波变换的近似分量 $A(i)^H$ 进行 DCT, 得到 $A(i)^{HC}$ 来完成水印的嵌入.

3) 水印信号的嵌入. 本文采用量化策略, 将水印信息嵌入到 DCT 域的低中频部分, 即量化调制前 N 个 DCT 系数 (N 为二值水印图像的宽度). 量化规则如下

$$a(i, j)^{HC} = \begin{cases} IQ(a(i, j)^{HC}) \times S(i, j) + S(i, j)/2, & \text{如果 } Q(a(i, j)^{HC}) = w_1(i, j) \\ IQ(a(i, j)^{HC}) \times S(i, j) - S(i, j)/2, & \text{如果 } Q(a(i, j)^{HC}) \neq w_1(i, j) \end{cases} \quad (1)$$

其中, $0 \leq i < M, 0 \leq j < N$, $S(i, j)$ 表示第 i 个音频数据段内第 j 个 DCT 系数的量化步长, $IQ(a(i, j)^{HC}) = \lfloor a(i, j)^{HC}/S(i, j) \rfloor$ ($\lfloor \cdot \rfloor$ 为向下取整函数), $Q(a(i, j)^{HC}) = \text{mod}(IQ(a(i, j)^{HC}), 2)$ (mod 是取模运算).

4) 量化步长的选取. 为了克服均匀量化方案存在的不足, 本文给出了量化步长的自适应确定方法, 其主要基于如下事实: a) 根据听觉时域掩蔽特性, 高能量音频段所发生的较大畸变难以被察觉, 故可有较大的量化步长值; b) 由小波分解理论, 若原始音频段能量较大, 则其近似分量 $A(i)^H$ 和细节分量 $D(i)^H$ 的能量也较大, 且 $A(i)^H$ 与 $D(i)^H$ 具有相同的数量级; c) 根据听觉频域掩蔽特性, 大幅度频率分量的改变不易被感知, 故可有较大的量化步长.

本文选取的第 i 个音频数据段 $A(i)$ 内第 j 个系数 $a(i, j)^{HC}$ 的量化步长 $S(i, j)$ 为

$$S(i, j) = 2^{\lfloor \log_2(|a(i, j)^{HC}| \cdot S_0(i)) \rfloor} \quad (2)$$

这里, $0 \leq i < M, 0 \leq j < N$, 并且

$$S_0(i) = \frac{\sum_{j=0}^{L/2^H-1} |d(i, j)^H|}{L/2^H}, \quad (0 \leq i < M) \quad (3)$$

其中 $d(i, j)^H$ 表示细节分量 $D(i)^H$ 内小波系数.

2.3 逆 DCT, 逆 DWT

对嵌入水印后的音频数据段 $A'(i)^{HC}$ ($0 \leq i < M$) 做逆 DCT, 以 $A'(i)^H$ 代替 $A(i)^H$ 并进行 H 级逆 DWT, 最终得到整个含水印音频信号 A' .

3 水印信号的抽取

本文讨论的半脆弱音频水印算法属于盲水印算法, 即提取数字水印信号时不需要原始音频载体的参与. 数字水印的提取过程如下:

1) 按照嵌入过程相同的操作, 对待检测的数字音频信号 A^* 做分段处理, 并作调整.

2) 对音频数据段 $A^*(i)$ ($0 \leq i < M$) 进行 H 级 DWT, 得到小波系数 $A^*(i)^H$.

3) 对音频数据段小波域的近似分量 $A^*(i)^H$ 做 DCT, 得到 $A^*(i)^{HC}$.

4) 按照第 2.2 节步骤 4), 计算量化步长 $S(i, j)^*$, 并提取水印信息, 提取公式为

$$w'_1(i, j) = [a^*(i, j)^{HC} / S(i, j)^*] \bmod 2 \quad (0 \leq i < M, 0 \leq j < N) \quad (4)$$

5) 对所提取出的水印信息 W'_1 进行逆置乱加密, 便可得到二值水印图像 W^* .

4 版权信息的判断与恶意篡改的认证

设 W_1 是置乱原始水印 W 而得到的二值水印矩阵, W'_1 为从接收到含水印音频信号 A' 中提取出的二值水印矩阵 (未经过逆置乱加密), 则版权信息判断与恶意篡改认证的步骤为:

1) 定义篡改矩阵 \hat{W}_1 (篡改矩阵 \hat{W}_1 中元素为 1 的区域意味着该区域可能被篡改)

$$\hat{W}_1 = W'_1 \oplus W_1 \quad (\oplus \text{ 为异或运算}) \quad (5)$$

2) 对篡改矩阵 \hat{W}_1 进行去噪. 实验结果表明, 可以通过篡改矩阵 \hat{W}_1 确定篡改发生位置, 但对待检测音频实施 MP3 压缩等常规处理操作后, 由 \hat{W}_1 所确定出的篡改发生区域呈均匀分布或随机分布 (即算法会在数字音频未遭受恶意攻击的情况下产生报警). 为此, 本文采用四邻域估计方法对篡改矩阵 \hat{W}_1 进行了去噪, 以消除恶意攻击误报 (包括虚警和漏警) 影响.

3) 计算音频数据段的篡改比率. 根据音频数据段与篡改矩阵的对应关系 (如图 2 所示), 我们可以将音频数据段 $A(i)$ 的篡改比率定义为 $r(i)$

$$r(i) = \frac{\sum_{j=0}^{N-1} \hat{w}_1(i, j)}{N} \quad (6)$$

其中, $\hat{w}_1(i, j) \in \hat{W}_1$ (\hat{W}_1 已经过去噪处理), N 为水印图像的列数.

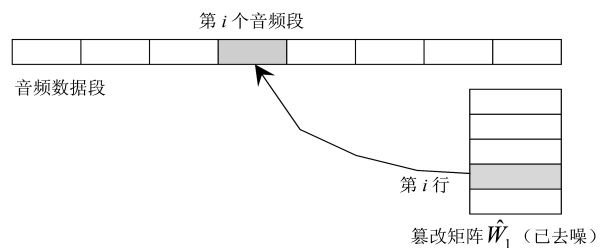


图 2 音频数据段的篡改检测

Fig. 2 Audio segment tamper detection

4) 适当选取检测阈值 T , 进行音频段的版权信息判断与恶意篡改认证. 规则如表 1 所示.

表 1 判断认证规则

Table 1 Judgement authentication regulation

条件	结果
$r(i) = 0$	音频数据段未遭受任何攻击
$r(i) \leq T$	音频数据段已遭受常规处理攻击
$r(i) > T$	音频数据段已遭受恶意篡改攻击

此外, 为了全面刻画数字音频的篡改程度, 可进一步定义被篡改音频段数 R_{num} 和整个数字音频的篡改程度 R_{total} , 即

$$R_{num} = \sum_{i=0}^{m-1} q(i) \quad (7)$$

$$q(i) = \begin{cases} 1, & \text{如果 } r(i) > T \\ 0, & \text{如果 } r(i) \leq T \end{cases}$$

$$R_{total} = R_{num} / M$$

其中, M 为二值水印图像的高度, $r(i)$ 为音频数据段 $A(i)$ 的篡改比率.

5 仿真实验

为了验证本文算法的高效性, 以下给出了抗攻击能力测试及篡改定位能力的实验结果, 并与相关算法进行了对照. 实验中, 选用的原始音频载体为采样频率为 44.1 kHz, 分辨率为 16 比特, 长度为 9.75 秒的单声道数字音频信号. 数字水印采用了 64×64 的二值图像. 小波变换采用了常见的 Daubechies-1 小波基. 小波变换级数为 $H = 3$. 检测阈值选取 $T = 0.1$. 本文还利用被篡改音频段数 R_{num} 和数字音频的篡改程度 R_{total} 对整个数字音频的篡改情况进行描述.

5.1 抗攻击能力测试

为了检测算法的鲁棒性, 仿真实验分别对本文算法及文献 [6] 算法的含水印音频信号进行了一系列常规攻击. 表 2 给出了几种算法的抗常规攻击能力对照结果.

文献 [7] 等所提出的完整性认证方案, 虽然可以进行数字音频的篡改定位, 但它们均无法抵抗常规攻击, 即均无法有效区分常规攻击与恶意攻击. 为了检测本文算法抵抗常规处理与恶意篡改的联合攻击能力, 仿真实验对篡改后的含水印音频信号进行一系列攻击, 表 3 给出了本文算法的抗联合攻击能

力对照结果. 不难看出, 本文方法仍具有较好的鲁棒性.

5.2 内容认证与篡改定位能力测试

为了检测本文算法的内容认证与篡改定位能力, 我们首先对含水印音频信号进行了恶意篡改, 然后进一步对其进行常规信号处理攻击. 表 4 (见下页) 给出了四种情况下的篡改矩阵 \hat{W}_1 、篡改音频段数 R_{num} 和整个数字音频的篡改程度 R_{total} . 其中, 四种情况分别为含水印音频、经过常规处理的含水印音频、经过篡改的含水印音频、经过常规处理与篡改的含水印音频. 图 3 给出了篡改定位结果 (经过篡改的含水印音频).

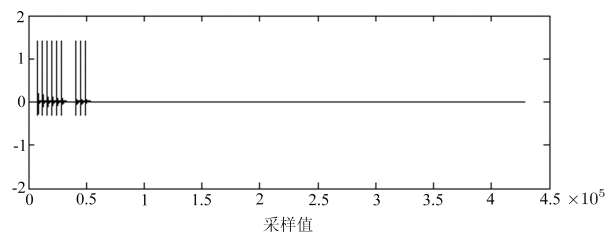


图 3 篡改后的含水印音频信号的篡改定位结果

Fig. 3 Tamper region locating of the tampered digital watermarking audio

表 2 数字水印对常规攻击的抵抗能力

Table 2 The watermark detection results for various attacks

攻击	峰值信噪比 (PSNR)		归一化相关系数 (NC)	
	文献 [6] 算法	本文算法	文献 [6] 算法	本文算法
未攻击	51.4854	50.6497	1.0000	1.0000
重采样	30.4207	30.4134	0.8356	0.9190
重量化	45.3476	45.2952	0.9726	0.9991
加白噪声	45.6669	44.7293	0.9186	0.9991
MP3-64 kb	50.3612	49.6058	0.3707	0.9944
MP3-56 kb	49.523	48.9140	0.3726	0.9872
低通滤波	36.6745	36.6447	0.3771	1.0000







表 3 数字水印对联合攻击的抵抗能力

Table 3 The watermark detection results for complex attacks

本文算法	水印图像	篡改	篡改+重新采样	篡改+重新量化	篡改+加白噪声	篡改+MP3-64kb	篡改+MP3-56kb	篡改+低通滤波
		NC	0.8469	0.7803	0.8466	0.8444	0.8481	0.8492
PSNR	30.705	27.6016	30.5945	30.5807	30.7803	30.7932	30.1195	

表 4 内容认证与篡改定位结果

Table 4 The results of content authentication and tampered region locating

	含水印音频		常规攻击		恶意篡改		联合攻击	
	未攻击	重新采样	MP3-56 kb	篡改	篡改+ 重新采样	篡改+ MP3-56 kb		
篡改矩阵 (已去噪)								
R_{num}	0	0	0	12	12	12		
R_{total}	0	0	0	0.1875	0.1875	0.1875		

实验数据表明, 本文算法不仅对恶意篡改具有脆弱性, 而且可以通过篡改矩阵 \hat{W}_1 大致确定发生篡改的位置 (见图 3), 同时还能给出整个数字音频的篡改程度 (如表 4 中的 0.1875)。

6 结论

本文以听觉掩蔽特性与数字音频局部相关特性为基础, 提出了一种可同时用于版权保护与内容认证的半脆弱数字音频盲水印算法。实验结果表明, 本文算法不仅具有较好的不可感知性, 而且对常规处理具有较好的鲁棒性, 同时能够对替换等恶意篡改作出报警并确定被篡改位置。此外, 该方案还具有计算简单、容易实现等优点, 这大大扩展了其应用范围。

References

- Li Wei, Yuan Yi-Qun, Li Xiao-Qiang. Overview of digital audio watermarking. *Journal on Communications*, 2005, **26**(2): 100~111
(李伟, 袁一群, 李晓强. 数字音频水印技术综述. 通信学报, 2005, **26**(2): 100~111)
- Ekici O, Sankur B, Akcay M. Comparative evaluation of semifragile watermarking algorithm. *Journal of Electronic Imaging*, 2004, **13**(1): 209~216
- Seok J, Hong J, Kim J. A novel audio watermarking algorithm for copyright protection of digital audio. *ETRI Journal*, 2002, **24**(3): 181~189
- Wu Shao-Quan, Huang Ji-Wu, Huang Da-Ren. DWT-based audio watermarking with self-synchronization. *Chinese Journal of Computers*, 2004, **27**(3): 365~370
(吴绍权, 黄继武, 黄达人. 基于小波变换的自同步音频水印算法. 计算机学报, 2004, **27**(3): 365~370)

- Lee H S, Lee W S. Audio watermarking through modification of tonal maskers. *ETRI Journal*, 2005, **27**(5): 608~661
- Zhao Chun-Hui, Li Fu-Chang. A blind audio watermark algorithm based on DWT and DCT. *Journal of Electronics and Information Technology*, 2003, **25**(11): 1570~1574
(赵春晖, 李福昌. 一种 DWT 与 DCT 结合的盲音频水印算法. 电子与信息学报, 2003, **25**(11): 1570~1574)
- Quan Xiao-Mei, Zhang Hong-Bin. Fragile audio watermarking algorithm for telltale tamper proofing and authentication. *Journal of Electronics and Information Technology*, 2005, **27**(8): 1187~1192
(全笑梅, 张鸿宾. 用于篡改检测及认证的脆弱音频水印算法. 电子与信息学报, 2005, **27**(8): 1187~1192)



王向阳 辽宁师范大学计算机系教授。主要研究方向为信息隐藏与数字水印、多媒体数据压缩技术。本文通信作者。

E-mail: wxy37@126.com

(WANG Xiang-Yang Professor at Department of Computer Science, Liaoning Normal University. His research interest covers information hiding, digital watermarking, and multimedia data compressing technology. Corresponding author of this paper.)



祁薇 辽宁师范大学硕士研究生。主要研究方向为数字水印技术。

E-mail: pinpin053@163.com

(QI Wei Master student at Liaoning Normal University. Her research interest covers digital watermarking.)