基于统计特征的音频中隐藏信息估计方法

谢春辉　程义民　汪云路　陈扬坤
(中国科学技术大学电子科学与技术系　合肥　230027)

摘　要：该文基于信息嵌入不同位平面所引起的幅度统计分布变化，提出了一种音频中隐藏信息长度估计方法。通过该方法，不仅可以判别未知音频中是否包含隐藏信息，而且还能有效识别信息嵌入位平面和较精确地估计隐藏信息长度。该方法已在微机上进行了实验，实验结果表明，该方法识别正确率较高，隐藏信息长度估计较精确，可有效用于音频隐藏分析。

关键词：信息隐藏；隐藏分析；位平面；副峰；嵌入率；统计特征
中图分类号：TP391　　**文献标识码**：A　　**文章编号**：1009-5896(2009)06-1341-04

Estimation of Secret Message in Audio Based on Statistic Characteristics

Xie Chun-hui　Cheng Yi-min　Wang Yun-lu　Chen Yang-kun
(*Dept. of Electronic Science & Technology, University of Science & Technology of China, Hefei* 230027, *China*)

Abstract: Based on the audio amplitude distribution change caused by secret message embedding in different bit-planes, a method of secret message length estimation is presented in this paper. This method can not only discriminate the existence of secret message, but also recognize the embedding bit-plane and estimate the amount of secret message. The proposed method has been implemented on PC. The experimental result shows its good recognition performance and relatively high estimation accuracy of secret message length. The proposed method can be used effectively in audio steganalysis.

Key words: Information hiding; Steganalysis; Bit-plane; Sub-peak; Embedding rate; Statistic characteristics

1　引言

目前，隐藏分析[1-3]已成为信息隐藏技术领域[4,5]的重要研究课题。选择在不同位平面(包括最低位)隐藏信息的方法，具有容量大，容易实现等优点，应用较广。目前，许多隐藏方法和工具，如Steganos，S-Tools，StegoMagic，Stegowav，Hide4PGP等，都可以选择音频作为隐藏载体。近年来，位平面隐藏分析越来越受到关注。

2003 年， Dumitrescu 等 [6] 提 出 的 SPA(Sample Pair Analysis)方法，能较精确地估计隐藏信息长度，但主要针对最低位随机间隔嵌入方法有效。2005年，Yu等[7]提出的方法，在嵌入位平面已知时，可实现某一位平面的隐藏分析。2007年，Ker等[8]将SPA方法扩展到能分析最低两个位平面。

采用上述方法进行隐藏分析时，须预先知道嵌入位平面，且其中一些方法(如SPA)只对随机间隔嵌入方法有效。针对这些问题，本文提出了一种基于统计特征的音频中隐藏信息估计方法，无须预知嵌入位平面，对嵌入任一位平面的隐藏音频，不仅可以判别隐藏信息的存在性，而且还能有效识别信息嵌入位平面和较精确地估计秘密信息长度。该方法对连续嵌入和随机间隔嵌入算法均有效。

2　系统概述

图 1 给出了音频隐藏分析模型，包括训练和测试两个阶段。

图 1(a)表示训练阶段。训练音频进行幅度分布统计后，确定副峰位置并计算其峰值；根据副峰峰值估计信息嵌入率；通过 SVM 训练，可得到判决阈值。

图 1(b)表示测试阶段。对输入的待测音频，采用和训练阶段相同的步骤得到嵌入率估计值后，利用训练阈值进行判决：如为藏密音频，则根据副峰位置，进一步识别秘密信息嵌入位平面。

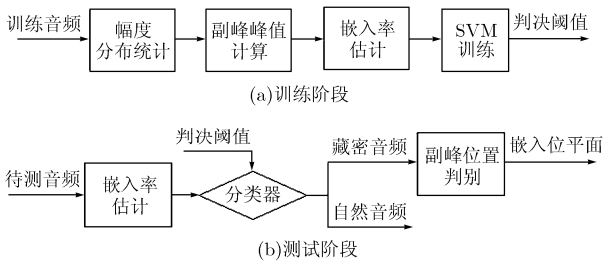


图 1 隐藏分析系统框图

3　隐藏分析方法

3.1 藏密音频幅度分布分析

设自然音频 $\boldsymbol{X}$ 包含 N 个采样点，每个采样点用 q 比特存储，二进制表述为 $b_q b_{q-1}\cdots b_L\cdots b_1$， b_q, b_L 和 b_1 分别表示最高位，第 L 位和最低位[7]。

2008-07-18 收到，2008-10-27 改回

若选择第 L $(1 \le L \le q)$位平面作为嵌入位平面，最大嵌入容量为 N bit，假设嵌入了 N_1 bit 信息，记 $p = N_1 / N$ 为嵌入率，用以描述秘密信息长度。设 H_I 和 $H_I^{'}$ 分别表示自然音频 $\boldsymbol{X}$ 和相应藏密音频 $\boldsymbol{X}'$ 中幅度为 I 的样点集合。$N_I = \| H_I \|$，$N_I^{'} = \| H_I^{'} \|$，$\| \|$表示集合中元素个数。

设 ΔN_I 表示幅度为 I 的样点总数嵌入前后的变化量。则 $b_L = 1$ 时，如图 2(a)所示有

$$\Delta N_I = N_I^{'} - N_I = \left[\frac{p}{2} N_{I-2^{L-1}} + \left(1 - \frac{p}{2}\right) N_I\right] - N_I = \frac{p}{2}(N_{I-2^{L-1}} - N_I) \tag{1}$$

相对变化量可表示为

$$\frac{\Delta N_I}{N_I} = \frac{p}{2}\left(\frac{N_{I-2^{L-1}}}{N_I} - 1\right) \tag{2}$$

同理，当 $b_L = 0$ 时，根据图 2(b)有

$$\frac{\Delta N_I}{N_I} = \frac{p}{2}\left(\frac{N_{I+2^{L-1}}}{N_I} - 1\right) \tag{3}$$

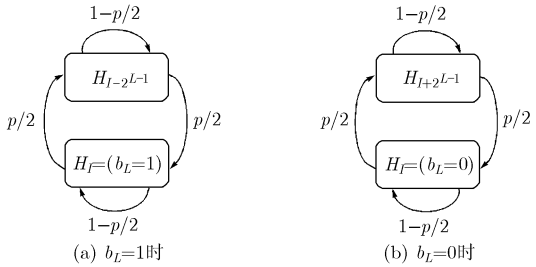


图 2 位平面隐藏对幅度分布的影响

由式(2)，式(3)可见，幅度分布曲线在某点的相对变化量不仅取决于嵌入率 p 和嵌入位平面 L，而且与音频幅度分布特性具有一定的相关性。

经研究表明，大部分音频信号幅度分布符合超高斯分布[9]。信号峭度(kurtosis)[10]为

$$K = \frac{E(X-\mu)^4}{[E(X-\mu)^2]^2} - 3 \tag{4}$$

其中 $\mu = E(X)$，表示期望值。对于 q bit 音频，一般地有 $\mu = 2^{q-1}$，且嵌入率不大时，嵌入前后期望值保持不变。

(1)当 $I = \mu + 2^{L-1}$，即 $b_L = 1$ 时，根据式(2)有

$$\frac{\Delta N_{\mu+2^{L-1}}}{N_{\mu+2^{L-1}}} = \frac{p}{2}\left(\frac{N_\mu}{N_{\mu+2^{L-1}}} - 1\right) \tag{5}$$

对于超高斯分布的音频信号，峭度 K 值较大，幅度分布曲线中心附近比较陡峭，存在 $N_\mu \gg N_{\mu+2^{L-1}}$，使得 $\dfrac{\Delta N_{\mu+2^{L-1}}}{N_{\mu+2^{L-1}}} \gg 0$，该处有较大增量，可能产生副峰，且嵌入率越高，副峰峰值越大。

(2)当 $I = \mu$，即 $b_L = 0$ 时，根据式(3)有

$$\frac{\Delta N_\mu}{N_\mu} = \frac{p}{2}\left(\frac{N_{\mu+2^{L-1}}}{N_\mu} - 1\right) \tag{6}$$

同理，存在 $\dfrac{\Delta N_\mu}{N_\mu} \ll 0$，该处有较明显减量；

(3)当 $I \ne \mu, \mu + 2^{L-1}$ 时，则 $\dfrac{\Delta N_I}{N_I} \approx 0$，该处变化不明显。

3.2 秘密信息长度估计

根据上面的讨论，藏密音频幅度分布曲线可能会在 $I = \mu + 2^{L-1}$ 处形成副峰，副峰出现位置与嵌入位平面相关，副峰峰值与嵌入率相关。

根据式(2)和式(3)有

$$\Delta N_{\mu+2^{L-1}} = \frac{p}{2}(N_\mu - N_{\mu+2^{L-1}}) \tag{7}$$

$$\Delta N_\mu = \frac{p}{2}(N_{\mu+2^{L-1}} - N_\mu) \tag{8}$$

显然 $\Delta N_{\mu+2^{L-1}} = -\Delta N_\mu$。

记 $\hat{p}$，$\widehat{\Delta}$，$\widehat{N}$ 分别为 p，Δ，N 的估计值。对 $\widehat{N}_{\mu+2^{L-1}}$ 采用如下估计

$$\begin{aligned}\widehat{N}_{\mu+2^{L-1}} &= \left(N_{\mu+2^{L-1}-1} + N_{\mu+2^{L-1}+1}\right)\Big/2 \\ &= \left[\left(N^{'}_{\mu+2^{L-1}-1} + N^{'}_{\mu+2^{L-1}-1}\right) - \left(\Delta N_{\mu+2^{L-1}-1} + \Delta N_{\mu+2^{L-1}+1}\right)\right]\Big/2 \approx \left(N^{'}_{\mu+2^{L-1}-1} + N^{'}_{\mu+2^{L-1}+1}\right)\Big/2\end{aligned} \tag{9}$$

则有

$$\begin{aligned}\widehat{\Delta N}_{\mu+2^{L-1}} &= \left(N^{'}_{\mu+2^{L-1}} - \widehat{N}_{\mu+2^{L-1}}\right) \\ &= N^{'}_{\mu+2^{L-1}} - \left(N^{'}_{\mu+2^{L-1}-1} + N^{'}_{\mu+2^{L-1}+1}\right)\Big/2\end{aligned} \tag{10}$$

而

$$\widehat{N}_\mu = N^{'}_\mu - \Delta N_\mu = N^{'}_\mu + \widehat{\Delta N}_{\mu+2^{L-1}} \tag{11}$$

最后根据式(7)有

$$\hat{p} = 2 \cdot \widehat{\Delta N}_{\mu+2^{L-1}} \Big/ (\widehat{N}_\mu - \widehat{N}_{\mu+2^{L-1}}) \tag{12}$$

式(12)右边所有项均可由藏密音频计算得到，而无须自然音频的任何信息，从而为盲检测提供了可能。

3.3 修正因子

下面来考虑推导式(9)过程中的省略项 $-(\Delta N_{\mu+2^{L-1}-1} + \Delta N_{\mu+2^{L-1}+1})/2$ 对估计结果的影响。

当 $I = \mu + 2^{L-1} + 1$ 时，$b_L = 1$，根据式(2)

$$\Delta N_{\mu+2^{L-1}+1} = \frac{p}{2}(N_{\mu+1} - N_{\mu+2^{L-1}+1}) \tag{13}$$

当 $I = \mu + 2^{L-1} - 1$ 时，$b_L = 0$，根据式(3)

$$\Delta N_{\mu+2^{L-1}-1} = \frac{p}{2}(N_{\mu+2^L-1} - N_{\mu+2^{L-1}-1}) \tag{14}$$

则有

$$\Delta N_{\mu+2^{L-1}-1} + \Delta N_{\mu+2^{L-1}+1} = \frac{p}{2}\left[\left(N_{\mu+2^L-1} + N_{\mu+1}\right) - \left(N_{\mu+2^{L-1}-1} + N_{\mu+2^{L-1}+1}\right)\right] \tag{15}$$

一般来说，N_I 服从凹函数分布，有 $\Delta N_{\mu+2^{L-1}-1} + \Delta N_{\mu+2^{L-1}+1} > 0$ 成立，$\widehat{N}_{\mu+2^{L-1}}$ 偏大于 $N_{\mu+2^{L-1}}$，根据式(10)有 $\widehat{\Delta N}_{\mu+2^{L-1}}$ 偏小于 $\Delta N_{\mu+2^{L-1}}$，导致 $\hat{p}$ 偏小于真实值，因此在式(12)中引入修正因子 k，最终得到

$$\hat{p} = 2k \cdot \widehat{\Delta}N_{\mu+2^{L-1}} \Big/ (\widehat{N}_{\mu} - \widehat{N}_{\mu+2^{L-1}}) \tag{16}$$

修正因子 k 取值一般略大于 1。

4　训练和测试

4.1 训练

(1)输入训练音频，统计幅度分布 $N_I^{'}$，$I \in \{0,1,2,\cdots, 2^q-1\}$；

(2) 根据式 (9)，式 (10) 分别计算得到 $\widehat{N}_{\mu+2^{L-1}}$ 和 $\widehat{\Delta}N_{\mu+2^{L-1}}$，其中 $L \in \{1,2,\cdots,q\}$；

(3)记 $I_L = \{I \mid \Delta_I = \max\{\widehat{\Delta}N_{\mu+2^{L-1}}\}, L = 1,2,\cdots,q\}$，根据式(16)计算嵌入率估计值 $\hat{p} = 2k \cdot \widehat{\Delta}N_{I_L} / (\widehat{N}_{\mu} - \widehat{N}_{I_L})$；

(4)按照前面所述步骤，分别计算自然音频和藏密音频的 $\hat{p}$ 值；

(5)将 $\hat{p}$ 组成输入向量，自然音频标记为-1，藏密音频标记为 1，采用 SVM 训练得到判决阈值 p_T。

4.2 测试

对于某一待测音频，采用和训练阶段相同的步骤(1)-步骤(3)，计算得到 $\hat{p}$ 和 I_L 后，按式(17)进行判决：

$$\text{待测音频} \in \begin{cases} \text{自然音频}, & \hat{p} < p_T \\ \text{藏密音频}, & \text{其它} \end{cases} \tag{17}$$

如为藏密音频，进一步判决嵌入位平面：

$$\hat{L} = \log_2(I_L - \mu) + 1 \tag{18}$$

5　实验结果

为了验证该方法的正确性，在微机上进行了实验。实验用 PC 机 CPU P4 1.3GHz，内存 256M，软件平台为 Windows XP 下的 Matlab 7.0.4。

训练阶段，从音频数据库中随机挑选 50 段自然音频，选择在不同位平面嵌入不等量信息得到 50 段藏密音频，将自然音频与藏密音频分别进行标记，输入 SVM 训练，得到判决阈值 p_T =0.077，修正因子 k 取值为 1.1。

测试阶段，从音频数据库中随机挑选 300 段自然音频(WAV 格式，包括 75 段歌曲，75 段音乐，75 段音效，75 段语音)，选定嵌入率 $p = 0.10, 0.12, 0.15, 0.20, 0.25, 0.30$，$L = 1,2,3,4$，采用 Hide4pgp 等隐藏工具[11]生成 24 个藏密音频集(每个集合包含 300 段音频)。最后将自然音频集和 24 个藏密音频集混合，根据式(17)，式(18)进行判决，结果列于表 1，表 2。

表 1 是待测音频中是否存在隐藏信息的判别结果，嵌入率为 0 表示自然音频；表 2 是藏密音频中信息嵌入位平面识别结果。

由表 1，表 2 可以看出，相同嵌入率水平下，隐藏信息嵌入较低位平面时，副峰出现在直方图中心附近变化较剧烈处，识别正确率相对较低；嵌入较高位平面时，副峰出现在距离直方图中心较远的相对平缓处，识别正确率相对较高。嵌入率达到 0.15 时，针对各位平面的识别正确率均接近甚至超过 95%。

根据 3.2 节的分析，该方法不仅能判别隐藏信息存在性和识别嵌入位平面，并且能估计隐藏信息长度。表 3 给出了嵌入率估计结果($\mu(\hat{p})$ 表示 $\hat{p}$ 的均值，$\bar{\mu}(\hat{p})$ 和 $\bar{\sigma}(\hat{p})$ 分别表示同一嵌入率水平下，各位平面 $\hat{p}$ 均值和标准差的平均值)。

表 1 是否存在隐藏信息的判别结果

		$L=1$		$L=2$		$L=3$		$L=4$		正确判别总数	平均正确判别率(%)
		误判数	正确判别率(%)	误判数	正确判别率(%)	误判数	正确判别率(%)	误判数	正确判别率(%)		
	0									267/300	89.00
	0.10	86/300	71.3	96/300	68	32/300	89.3	32/300	89.7	954/1200	79.50
	0.12	53/300	82.3	47/300	84.3	8/300	97.3	8/300	97.3	1084/1200	90.33
嵌入率	0.15	22/300	93.7	4/300	98.7	0/300	100	0/300	100	1174/1200	97.83
	0.20	6/300	98	0/300	100	0/300	100	0/300	100	1194/1200	99.50
	0.25	5/300	98.3	0/300	100	0/300	100	0/300	100	1195/1200	99.58
	0.30	3/300	99	0/300	100	0/300	100	0/300	100	1197/1200	99.75

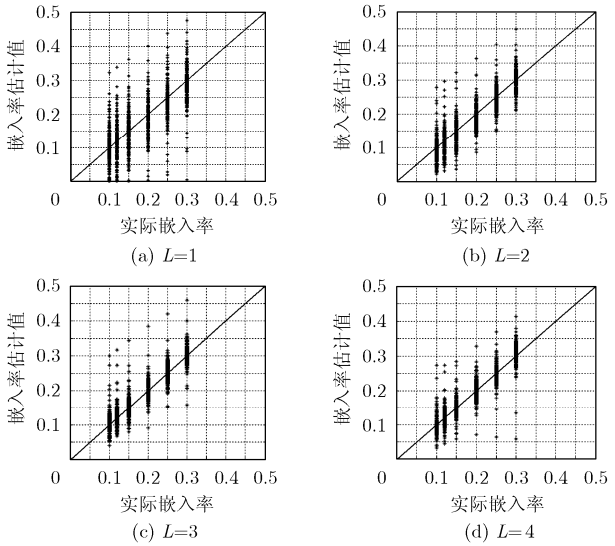
表 2 嵌入位平面识别结果

		$L=1$		$L=2$		$L=3$		$L=4$		正确识别总数	平均正确识别率(%)
		错误识别数	正确识别率(%)	错误识别数	正确识别率(%)	错误识别数	正确识别率(%)	错误识别数	正确识别率(%)		
	0.10	86/300	71.3	103/300	65.7	51/300	83	45/300	85	915/1200	76.25
	0.12	53/300	82.3	51/300	83	25/300	91.7	19/300	93.7	1052/1200	87.67
嵌入率	0.15	22/300	92.7	17/300	97.7	19/300	97	17/300	97.7	1155/1200	96.25
	0.20	6/300	98	1/300	99.7	5/300	98.3	4/300	98.7	1184/1200	98.67
	0.25	3/300	98.3	0/300	100	5/300	98.3	3/300	99	1189/1200	99.08
	0.30	3/300	99	0/300	100	4/300	98.7	2/300	99.3	1191/1200	99.25

表 3 嵌入率估计结果

		$L=1$		$L=2$		$L=3$		$L=4$		平均值	
		$\mu(\hat{p})$	$\sigma(\hat{p})$	$\mu(\hat{p})$	$\sigma(\hat{p})$	$\mu(\hat{p})$	$\sigma(\hat{p})$	$\mu(\hat{p})$	$\sigma(\hat{p})$	$\overline{\mu}(\hat{p})$	$\overline{\sigma}(\hat{p})$
嵌入率	0									0.0349	0.0348
	0.10	0.1083	0.0501	0.0934	0.0335	0.1015	0.0251	0.1014	0.0235	0.1011	0.0330
	0.12	0.1278	0.0559	0.1129	0.0364	0.1204	0.0267	0.1198	0.0253	0.1202	0.0361
	0.15	0.1684	0.0501	0.1429	0.0325	0.1497	0.0240	0.1488	0.0225	0.1524	0.0322
	0.20	0.2095	0.0517	0.1938	0.0304	0.1988	0.0241	0.1984	0.0237	0.2001	0.0324
	0.25	0.2625	0.0522	0.2466	0.0290	0.2508	0.0231	0.2492	0.0259	0.2522	0.0325
	0.30	0.3151	0.0533	0.2997	0.0279	0.3026	0.0231	0.3008	0.0289	0.3045	0.0333

图 3 给出了隐藏信息嵌入不同位平面所得到的藏密音频，通过该方法计算得到的嵌入率估计值 $\hat{p}$ 与实际嵌入率 p 之间的误差分布。图中每个点与对角线之间的纵轴方向距离表示估计误差：$|\hat{p}-p|$。



(a) L=1 (b) L=2 (c) L=3 (d) L=4

图 3 秘密信息长度估计误差分布

6 结束语

本文提出了一种基于统计特征的音频中隐藏信息长度估计方法。通过对未知音频幅度分布曲线的分析，不仅可以判别是否包含隐藏信息，对于藏密音频，还能进一步识别信息嵌入位平面和估计隐藏信息长度。实验结果表明，该方法识别正确率较高，隐藏信息长度估计较精确。该方法不仅可用于音频位平面信息隐藏分析，对于隐藏载体与音频有类似统计特性的隐藏分析也同样适用。

参 考 文 献

[1] Dumitrescu S and Wu Xiaolin. A new framework of LSB steganalysis of digital media [J]. *IEEE Trans. on Signal Processing*, 2005, 53(10): 3936–3947.

[2] Ker A D. Derivation of error distribution in least squares steganalysis [J]. *IEEE Trans. on Information Forensics and Security*, 2007, 2(2): 140–148.

[3] 詹双环, 张鸿宾. 基于小波分解和方差分析的图像信息隐藏盲检测[J]. 电子与信息学报, 2007, 29(6): 1460–1463.
Zhan Shuang-huan and Zhang Hong-bin. Image-based blind steganalysis using wavelet statistics and analysis of variance [J]. *Journal of Electronics & Information Technology*, 2007, 29(6): 1460–1463.

[4] 田源, 程义民, 王以孝. 一种新的数据隐藏方法[J]. 电子学报, 2004, 32(9): 1444–1447.
Tian Yuan, Cheng Yi-min, and Wang Yi-Xiao. A novel method of data hiding [J]. *Acta Electronica Sinica*, 2004, 32(9): 1444–1447.

[5] 程义民, 钱振兴, 王以孝, 等. 基于数位信息的信息隐藏方法[J]. 电子与信息学报, 2005, 27(8): 1304–1309.
Cheng Yi-min, Qian Zhen-xing, and Wang Yi-xiao, *et al*.. A method of information hiding based on the digital-position information [J]. *Journal of Electronics & Information Technology*, 2005, 27(8): 1304–1309.

[6] Dumitrescu S, Wu Xiallin, and Wang Zhe. Detection of LSB steganography via sample pair analysis [J]. *IEEE Trans. on Signal Processing*, 2003, 51(7): 1995–2007.

[7] Yu Xiao-yi, Tan Tie-niu, and Wang Yun-hong. Extended optimization method of LSB steganalysis [C]. IEEE International Conference on Image Processing 2005. Genova, Italy, 2005(II): 1102–1105.

[8] Ker A D. Steganalysis of embedding in two least-significant bits [J]. *IEEE Trans. on Information and Security*, 2007, 2(1): 46–54.

[9] Huang Hesu and Kyriakakis C. Blind dereverberation of audio signals using a modified constant modulus algorithm [C]. 121st Audio Engineering Society. San Francisco, CA, US, 2006: 6974–6977.

[10] 史晓非, 刘人杰, 苗瑞. 一种峭度依赖的参数自适应盲分离算法[J]. 电子与信息学报. 2006, 28(11): 2033–2036.
Shi Xiao-fei, Liu Ren-jie, and Miao Rui. A parameter kurtosis-dependent flexible BSS algorithm [J]. *Journal of Electronics & Information Technology*, 2006, 28(11): 2033–2036.

[11] Johnson M K, Lyu S, and Farid H. Steganalysis of recorded speech [C]. SPIE Symposium on Electronic Imaging. San Jose, CA, US, 2005: 664–672.

谢春辉： 男，1983 年生，博士生，研究方向为多媒体信息处理、隐藏分析等.

程义民： 男，1945 年生，教授，博士生导师，研究领域为信息隐藏、网络多媒体、计算机视觉等.

汪云路： 女，1981 年生，博士生，研究方向为隐藏分析.