抗 PCM A/μ律压缩及其转换的语音密写

胡礼才　　王朔中
(上海大学通信与信息工程学院　上海　200072)

摘　要：该文提出一种在数字语音信号中嵌入隐蔽信息的方法。根据 PCM A/μ 律压缩及其转换特性，采用对某些宿主语音样本预修改和动态选择嵌入比特位的方法，使密写信号不仅没有感知失真，而且经 A/μ 律压缩及其相互转换处理后仍可无误提取嵌入的隐蔽数据，因此含密语音信号不仅可在 Internet 上传输，还可在跨国 PSTN 中传输。实验结果表明了该方法的有效性。

关键词：密写；A/μ 律；量化间隔；A/μ 律转换
中图分类号：TP309　　**文献标识码**：A　　**文章编号**：1009-5896(2007)09-2226-04

Steganography in Speech Immune to PCM A/μ Law Compression and Mutual Conversion

Hu Li-cai　　Wang Shuo-zhong
(*School of Communication and Information Engineering, Shanghai University, Shanghai*　200072 *China*)

Abstract: A steganographic method is proposed for embedding secret information in digital speech signals. Depending on the PCM A/μ law compression/decompression characteristics and mutual conversion between the two laws, appropriate bit planes are chosen in modifying the host speech samples to embed secret data. The induced distortion is highly imperceptible, and the embedded information is robust against A/μ law compression and mutual conversion, therefore suitable for transmission not only through the Internet but also through international PSTN. Experimental results indicate the effectiveness of the proposed method.

Key words: Steganography; A/μ law; Quantization step; Mutual conversion between A/μ laws

1 引言

数字密写就是以图像、音频信号等数字媒体作为掩护，嵌入隐蔽信息，以不引起外界注意的方式通过公共信道进行传递。音频密写是一种以数字音频信号为嵌入对象传递隐蔽信息的有效方法。一种简单而有效的密写方法是最低有效位(LSB)嵌入法[1,2]，如密写工具Steghide[3]。LSB方法的优点是算法简单，听觉隐蔽性好，而且嵌入量大，但这种方法抗噪声和变换的能力很差。

作为改进，Gopalan[4]和Cvejic[5]等人提出了较高有效位嵌入法(第 5 LSB或更高)，这种方法能有效抵抗加性噪声。但随着嵌入层的提高会引起宿主信号和密写信号语图上的差异，并且嵌入层的选择受宿主信号的幅度影响较大(宿主信号幅度小，嵌入层不能选得太高)。另外，它不能抵抗诸如符合ITU建议G711 的PCM A/μ 律压缩，对窄带语音而言，使用起来有很大的局限性。本文在较高有效位嵌入法基础上，根据PCM A/μ 律压缩特性[6]，提出按宿主样本幅度所在的不同量化区间动态选择嵌入位，称为动态有效位嵌入法。该方法除能有效抵抗加性噪声外，更重要的特点是能有效抵抗PCM A/μ 律压缩处理，而且密写信号的语图没有异常特性，故在无宿主信号参照情况下不可觉察。此外，根据A/μ 律压缩算法的转换特性[6]，修改某些量化区间的嵌入层和对某些候选样值幅度进行预修改，进一步可有效抵抗PCM A/μ 律压缩之间的转换，因此密写信号可以压缩为PCM A/μ 律通过Internet和国际间PSTN线路传输。

2 数据嵌入方法

A/μ律压缩是利用语音信号幅度分布的非均匀性[7](小幅度出现的概率大)，将语音信号幅度区间非均匀地划分为 256 个小区间，称为量化区间，量化区间的大小称为量化间隔(见表 1)，小幅度信号量化区间小，大幅度信号量化区间大，处于同一量化区间的样本用一个 8 bit码字表示，译码值用其所在量化区间两端点的平均值表示。可见，语音样本经A/μ 行律压缩/解压后幅度会发生很大变化，因此将隐蔽信息嵌入到某一固定的较高有效位虽然可以使密写信号能抵御一定强度的加性噪声干扰，却不能抵御A/μ 律压缩。

为了适应 A/μ 律压缩，本文根据宿主信号样本所在量化区间的不同，动态选取承载隐蔽信息的有效位。以下分几种情况给出嵌入方法。

2006-03-13 收到，2006-09-04 改回
国家自然科学基金(60372090)资助课题

2.1 抗 A 律嵌入方法

表 1 列出对样本正值进行 A 律压缩的量化区段和编码值范围，对负值的编解码方法与正输入值相同。

一个 16 bit 的数字语音样本

$$S = sa_{15}a_{14}a_{13}a_{12}a_{11}a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1 \tag{1}$$

动态范围为 [–32768, 32767]。由于 A 律压缩标准要求的语音输入范围为[–4096, 4095]，因此要将 S 依 A 律进行压缩，必须先将 S 除以 8，也就是说只对 S 的最高 13 位

$$S' = sa_{15}a_{14}a_{13}a_{12}a_{11}a_{10}a_9a_8a_7a_6a_5a_4 \tag{2}$$

进行编码，然后再将解码结果乘以 8，得到下列译码值：

$$S'' = sa_{15}a_{14}a_{13}a_{12}a_{11}a_{10}a_9a_8a_7a_6a_5a_4000 \tag{3}$$

可见只有将信息嵌入到 S 的 a_4 层以上才有可能从 S'' 中正确提取出来。为了有效地规避语音样值幅度在 A 律编译码后的改变对嵌入信息产生的影响，以便从 A 律译码样值中正确提取嵌入信息，本文根据式(4)在不同的量化区间选用不同的嵌入层：

$$I_n = 4 + \log_2 A_n \tag{4}$$

其中 A_n 为第 n 量化区段的量化间隔， I_n 为第 n 量化区段的嵌入层。

例如当 S' 位于第 1 量化区段时，A_1 =2，I_n =5，数据应嵌入到 S' 的 a_5 。若 S' 位于第 4 量化区段，A_4 =16 则 I_n =8，数据应嵌入到 S' 的 a_8 。

根据式(4)选取的嵌入位可避免译码对嵌入信息的修改，因为译码最大可改变的位 G_n 为：

$$G_n = 4 + \log_2 \frac{A_n}{2} = 3 + \log_2 A_n \tag{5}$$

G_n 总是比 I_n 低一层，说明译码不会影响嵌入的信息。式(4)表示起始嵌入位为 4，当然也可以选得高些，以增强密写的鲁棒性，同时也能正确提取嵌入信息，但会增大嵌入引起的失真。一般来说，所选嵌入隐蔽信息的宿主语音段幅度较大时，嵌入起始位可以选得高些，如 5 或 6。

根据式(4)嵌入隐蔽信息时，对于所有正样本和多数负样本，A 律编码译码后嵌入信息不会改变，但在某些负的量化区间端点上嵌入的信息经 A 律编码译码后会发生改变，这主要是负数的补码表示引起的。

例如，$S' = -176_{10} = 11111010\underline{\mathbf{0}}000_2$(下划线处即为嵌入位置位)，嵌入“1”后成为 $11111010\underline{\mathbf{1}}000_2 = -168_{10}$，译码后为 $11111010\underline{\mathbf{0}}100_2 = -172_{10}$，可以看出嵌入信息发生了改变。为了解决这个问题，本文将原先的宿主信号先进行A律压缩解压后得到的信号作为宿主(对A/μ律压缩而言，一次解压缩与多次解压缩效果一样，不会引起多余失真)，这样根据式(4)嵌入信息后，样本仍然是某个量化区间的中间值，其编译码后值不变，从而保证了嵌入信息的不变。

2.2 抗 A 律嵌入方法

律压缩标准要求的语音输入范围为[–8159, 8159]，因此，要将 S 依律进行压缩，必须先将 S 除以 4，也就是说只对 S 的最高 14 位进行编码。因此，抗　律压缩密写方法中的嵌入层选择关系用下式表示：

$$I_n = 3 + \log_2 A_n \tag{6}$$

式中变量含义同式(4)。

依式(6)嵌入隐蔽信息的含密信号经 μ 律压缩解压后，在某些语音样本上嵌入的信息会发生错误。通过对隐蔽信息进行纠错编码后再嵌入可解决这个问题，但这种方法减少了信息的嵌入量。本文的解决办法是预先修改会使嵌入信息发生错误的语音样本值，然后再嵌入信息。如果用该方法对整个宿主文件的每一个样本都嵌入信息，密写后会缺少某些信号样本取值，从而引起攻击者的怀疑，在实际使用时可以只对宿主文件的部分样本进行嵌入。表 2 列出了需修改样本值和修改值，以及修改和不修改所引起的误差，可以看出，预修改宿主样本再嵌入隐蔽信息所引起的误差比不修改直接嵌入所引起的误差要小。

2.3 抗 A 律及 A/μ 律转换嵌入方法

将抗 A 律压缩密写方法嵌入隐蔽信息后得到的含密信号先进行 A 律压缩，再转换为μ律，然后经μ律解压后按抗 A 律压缩信息提取方法提取信息，嵌入信息在某些样本处出现错误。这里也采用预修改宿主语音样本值方法，解决嵌入信息错误问题。表 3 列出了需修改样本值和修改值及修改和不修改所引起的误差(仅列出了嵌入 0 时需修改的样本值和修改值，嵌入 1 时需修改的样本值是嵌入 0 时需修改样本值的相反数)。之所以如此修改，主要出于两方面考虑，一是修

表 1 正输入 A 律压缩量化区段和编码值范围

量化区段(n)	样本值范围	量化区间数	量化间隔(A_n)	编码值范围
7	2048 ~ 4095	16	128	11110000 ~ 11111111
6	1024 ~ 2047	16	64	11100000 ~ 11101111
5	512 ~ 1023	16	32	11010000 ~ 11011111
4	256 ~ 511	16	16	11000000 ~ 11001111
3	128 ~ 255	16	8	10110000 ~ 10111111
2	64 ~ 127	16	4	10100000 ~ 10101111
1	0 ~ 63	32	2	10000000 ~ 10011111

表 2 抗 μ 律嵌入中需修改样本值和修改值及修改和不修改所引起的误差(各值都为除以 4 之值)

嵌入 0				嵌入 1			
原样本值	修改后值	直接嵌入误差	修改后嵌入误差	原样本值	修改后值	直接嵌入误差	修改后嵌入误差
1023	975	80	48	1983	1919	96	64
−30	−28	3	2	−1023	−975	80	48
−1983	−1919	96	64				

表 3 抗 A 律及 A/μ 律转换嵌入中需修改的样本及修改值(各值都为除以 8 之值)

原样本值	修改后值	直接嵌入误差	修改后嵌入误差	原样本值	修改后值	直接嵌入误差	修改后嵌入误差
−49	−39	0	11	−114	−102	1	13
−51	−41	2	9	−118	−106	3	9
−53	−43	0	11	−122	−110	1	13
−55	−45	2	9	−126	−132	3	13
−57	−47	0	11	−244	−228	3	11
−59	−66	2	10	−252	−236	5	17
−61	−70	0	9	−976	−912	63	17
−63	−74	2	14	−1008	−944	31	49

改不要引起含密信号样本幅度统计直方图的剧烈变化，二是修改造成的误差较小。

2.4 抗律 μ 及 μ /A 律转换嵌入方法

将抗 μ 律压缩密写方法嵌入隐蔽信息后得到的含密信号进行 μ 律压缩，再转换为 A 律，然后经 A 律解压后按抗律压缩信息提取方法提取信息，嵌入信息仅出现在正/负样本第一量化区间内的样本上。可分两个步骤纠正这些错误。第 1 步，将第一量化区间内样本的嵌入层修改为 5(原为 4)。第 2 步，对经上述处理后仍然出现嵌入信息错误的样本进行预修改，表 4 列出了需修改样本值和修改值(错误在嵌入信息“0”和“1”时都存在)。

表 4 抗 μ 律及 μ /A 律转换嵌入中需修改的样本及修改值(各值都为除以 4 之值)

原样本值	修改后值	原样本值	修改后值
−4	−2	−20	−18
−8	−6	−24	−22
−12	−10	−28	−26
−16	−14		

3 实验结果

利用动态比特位嵌入法对 12 段语音进行了嵌入试验，其中男女声各占 6 段，每隔 3 个样本嵌入 1 比特信息。结果表明含密信号能抵抗 A/μ 律压缩及其转换，即经 A/μ 律压缩及转换后数据，解压后仍能正确无误地恢复全部嵌入信息。实验用的语音信号选自北京大学录制的语料库 PKU-SRSC，采样频率为 8kHz，16bit 线性码，每段语音长度为 9600 个样本，持续时间 1.2s。嵌入的信息为另一段语音，该语音经过混合激励线性预测(MELP)压缩编码，码率为 2.4kbps，嵌入信息长度为 56 帧，每帧为 54bit，共 3024bit，1.26s。图 1 显示隐蔽信息嵌入前后载体语音信号的波形变化，其中图 1(a)为宿主语音信号波形，图 1(b)为按抗 A 律压缩密写方法嵌入隐蔽信息后的含密信号波形，图 1(c)为两者之差。

密写信号经主观听觉测试无失真，与宿主信号相比在波形和语图上都不可区分。图 2 为宿主语音信号语图，图 3 为按抗A律压缩密写方法嵌入隐蔽信息的含密信号语图。表 5 列出了 4 种嵌入方法在 12 段宿主语音中嵌入隐蔽信息后的信噪比。信噪比根据式(7)计算[8]，其中$x(n)$为宿主信号，$y(n)$为含密信号。

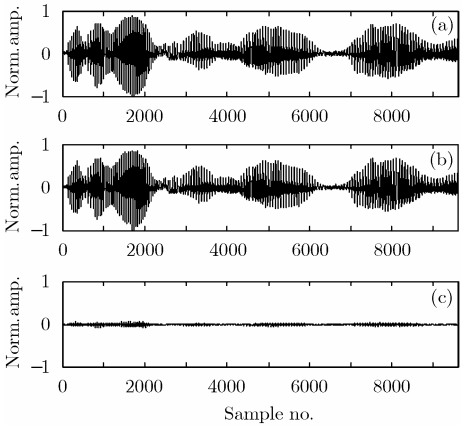

图 1 信息嵌入前后的语音信号波形和嵌入产生的误差

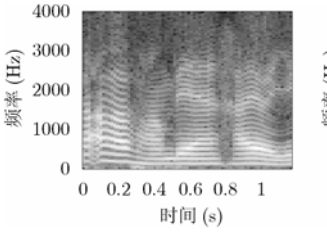

图 2 宿主语音信号语图

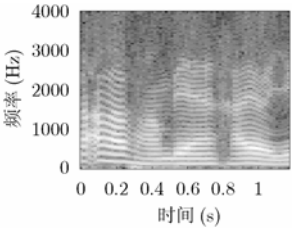

图 3 抗 A 律含密信号语图

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_n x^2(n)}{\sum_n \left[x(n)-y(n)\right]^2} \tag{7}$$

表 5　4 种密写方法的嵌入信噪比(dB)

算 法	最大	最小	平均
抗 A 律	34.84	34.08	34.47
抗 μ 律	34.49	34.07	34.31
抗 A-μ 律	34.29	33.45	33.75
抗 μ -A 律	34.49	34.06	34.31

4　结束语

本文提出的基于窄带语音信号的密写方法具有良好的隐蔽性，较大的嵌入信息量和优良的抗 PCM A/μ 律压缩及其转换性能，因而含隐蔽信息的语音信号不仅能通过因特网传输，还能在 A/μ 律压缩及转换后在 PSTN 网上进行跨国传输。如要提高在 PSTN 网上传输的抗干扰性能，可适当提高嵌入起始位，并采用纠错编码技术。前者会增大密写引起的失真，后者会降低有效的信息传输量，在实际中应在这几个指标之间进行合理的协调以满足应用需要。

参 考 文 献

[1] Wang H and Wang S Z. Cyber warfare: Steganography vs. steganalysis. *Communications of the ACM*, 2004,47(10): 76–82.

[2] Lie W and Chang L. Data hiding in images with adaptive numbers of least significant bits based on the human visual system. Proc. IEEE Int. Conf. Images Processing, Kobe, Japan, Oct. 1999: 286–290.

[3] Steghide Manual, http://steghide.sourceforge.net/documentation/manpage.php.

[4] Gopalan K. Audio steganography using bit modification. ICASSP Hong Kong 2003, II: 421–424.

[5] Cvejic N and Seppänen T. Increasing robustness of LSB audio steganography using a novel embedding method. Proceeding of the International Conference on Information Technology: Coding and Computing (ITCC'04), Las Vegas, Nevada, 2004: 533–537.

[6] Pulse Code Modulation of Voice Frequencies, ITU-T Recommendation G.711.

[7] 杨行峻，迟惠生等编著. 语音信号数字处理. 北京：电子工业出版社，第一版，1995: 169–170.

[8] Bassia P, Pitas I, and Nikolaidis N. Robust audio watermarking in the time domain. *IEEE Transaction on Multimedia*, 2001,3(2): 232–241.

胡礼才：　男，1965 年生，副研究员，研究方向为信道编码、语音编码、信息隐藏.

王朔中：　男，1943 年生，教授，博士生导师，研究方向为信号处理、图像处理、信息隐藏.