

对 G. 723. 1 高码率编码语音中进行压缩域信息隐藏的性能分析

游 鹏 季晓勇 卢 鹏

(南京大学电子科学与工程系 南京 210093)

摘 要 近年来,基于混沌理论的保密通信和数据保密得到广泛、深入的研究,提出了许多基于混沌理论的混沌加密算法,但这些算法缺乏可靠的安全性和鲁棒性。本文在对文献[1]提出的一种在语音混沌保密通信过程中进行信息隐藏方案的分析中,通过对 G. 723. 1 高速率编码语音中适合进行压缩域信息隐藏的码位性能分析和仿真实验,进一步证明了文献[1]方案的有效性和安全性等特点。

关键词 G. 723. 1, 高码率编码, 信息隐藏

Performance Analysis of Using G. 723. 1's High-rate Codes to Hiding Information in Packed Communication System

YOU Peng JI Xiao-yong LU Peng

(Department of Electronic Science and Engineering, Nanjing Univ, Nanjing 210093, China)

Abstract In recent years, a growing number of chaos encryption algorithms based on chaos theory have been proposed, many of them fundamentally flawed by a lack of robustness and security. In this paper, we propose a novel feedback for the features of G. 723. 1's high-rate codes to achieve an approach that hides the information in chaos-based secure communication system by sending information in the less significant bytes of the coded audio data. The result of experiment proves that the system mentioned in literature[1] is effective and secure.

Keywords G. 723. 1, High-rate codes, Information hiding

1 引言

文献[1]提出了一种在语音混沌保密通信过程中进行信息隐藏的方案。本文尝试在实验的基础上对 G. 723. 1 高码率编码语音中适合进行压缩域信息隐藏的码位,进行进一步的性能分析。

G. 723. 1 协议是双速率语音编码建议,高速率语音编码 6.3 kbps。对于 6.3 kbps 语音编码,每帧编码结果包含 189bit,添加 3bit 数据,实际每帧包含 192bit,具体分为数据类型 2 bit、线谱对参数索引 24 bit、闭环基音周期 18 bit、增益 48 bit、脉冲平移 4 bit、保留 1 bit、脉冲位置 73 bit、脉冲符号 22 bit。

2 对各码位进行嵌入数据仿真实验

2.1 2bit 的数据类型位

用来描述本帧的数据类型。2bit 能够表示四种状态,其中 00 代表 6.3 kbps 高码率编码数据帧。双方约定码率的情况下,解码器可以忽略实际数值按照约定码率解码,不会对语音质量造成不良影响,因此可以用于隐藏数据,但因隐蔽性较差而不常采用。

2.2 24bit 的线谱对参数索引位

24bit 的线谱对参数索引又可细分为未量化线谱对矢量、量化线谱对矢量、线谱对残差矢量。采用容量为 256 的码本对上述 3 个矢量进行矢量量化,尝试逐渐增加子矢量码字的修改位数,嵌入保密数据对参数修改尽量选择与原始参数距离最远误差最大。分析实验结果得知,嵌入信息量极限大约

15bit 左右。在保证一定语音质量的前提下,为增加线谱对参数索引所能隐藏的信息量,采取将数据平均嵌入三个子矢量码字的最后几位,分散误差以避免对某个子矢量的过度修改。

表 1 三个子矢量码字各隐藏 1-4bit 数据的语音效果

隐藏数据	SNR	SSNR	SD
3bit	8.586	5.049	0.567
6bit	7.517	4.089	0.626
9bit	7.330	3.972	0.727
12bit	6.158	2.958	0.863

具体仿真结果参考表 1,根据上述分析,嵌入 6bit 数据方案和嵌入 9bit 数据方案相比语音质量并没有明显提高,嵌入 12bit 数据方案缺乏隐蔽性,因此,在对语音质量要求比较高的情况下,应选取在三个子矢量码字中各嵌入 1bit 数据共 3bit 的方案;在对隐藏信息量要求比较高的情况下,一般选取在三个子矢量码字中各嵌入 3bit 数据共 9bit 的方案。

图 1 中表示 G. 723. 1 标准解码语音的波形与隐藏 3bit 保密数据时解码语音后的波形差值。分别用时域和频域波形显示。

初步观察结论,嵌入数据前后解码语音差值信号,波形在整个时域和频域上都分布得比较平均,接近白噪声特性难以被察觉。

2.3 18bit 的闭环基音周期位

18bit 的闭环基音周期参数可以继续细分为各 7bit 绝对值的第一、第三子帧闭环基音周期和各 2bit 绝对值的第二、第四子帧闭环基音周期。尝试单独修改每一子帧闭环基音周

期最低有效位,比较嵌入不同位置对语音质量的影响情况。实验数据显示,尽管奇数子帧基音周期和偶数子帧基音周期形式不同分别用 7bit 绝对值和 2bit 相对值表示,但嵌入 1bit 保密数据后语音受影响程度基本相同。

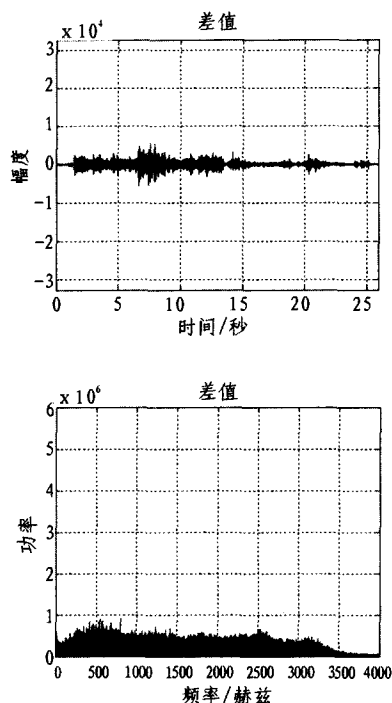


图 1 线谱对参数索引嵌入 3bit 数据的时域和频域波形图

增加嵌入数据信息量,如果认为前半帧和后半帧类似不加区分,当每帧嵌入 2bit 保密数据时,会出现六种不同情况。

表 2 闭环基音周期参数隐藏 2bit 数据语音效果

子帧闭环基音周期低位修改				SNR	SSNR	SD
一	二	三	四			
2	0	0	0	6.994	5.114	0.400
0	2	0	0	8.314	5.809	0.361
1	1	0	0	8.466	5.866	0.369
0	0	1	1	8.462	5.770	0.387
1	0	1	0	8.371	5.670	0.381
0	1	0	1	8.512	5.807	0.376

分析表 2 结果,闭环基音周期隐藏保密数据能力不佳,即使少量嵌入也会造成较强的干扰。计算闭环基音周期是 G. 723.1 参数编码过程中最后一个步骤,如果结果由于信息隐藏发生变化,此后的波形编码过程很难予以纠正。所以保密数据不适于嵌入闭环基音周期。

2.4 48bit 的增益位

对 48bit 的增益参数而言,首先,其中的码本标志不能用来隐藏任何数据。因为如果码本选择错误,语音充满尖锐噪声完全不能接受,信噪比和分段信噪比甚至变成负数。其次经过尝试,在增益组合的每一子帧每一参数修改 1bit 测试隐藏数据能力。利用基音增益隐藏保密数据,解码语音信噪比偏低,听觉感受是语音断续影响理解。而利用脉冲增益隐藏 1bit 保密数据虽然可以取得良好效果,但由于隐藏信息量不足,脉冲增益缺乏实用价值,因此,尽管增益占用编码结果的 25% 空间,但对重建语音的重要性使得不应该发生任何程度的变化,难以用来进行信息隐藏。

2.5 4bit 的脉冲平移位

4bit 的脉冲平移参数可以细分四个子帧的脉冲平移参数。

每一子帧包含 60 个采样点,描述一个采样点位置至少需要利用 6bit 数据。为了降低码率,残差信号的波形编码过程中脉冲要么全在奇数位置采样点上要么全在偶数位置采样点上,这样采样点位置只有 30 种可能,仅需 5bit 即可描述,每个采样点位置信息能够节省 1bit,付出的代价是增加脉冲平移参数。每一子帧的脉冲平移参数通过 1bit 表示脉冲究竟是在奇数位置采样点上还是在偶数位置采样点上。

由于每一子帧的脉冲平移参数只有 1bit,不存在高低有效位之分,而每帧的脉冲平移参数总共 4bit,可以在进行排列组合后测试所有能嵌入保密数据的方法对语音质量的影响。实验数据表明,当隐藏相同数量的保密数据时,语音质量各项指标基本相同,每一子帧的脉冲平移对重建语音的价值近似相等。表 3 反映的是分别隐藏 1-4bit 的几种代表情况。

表 3 脉冲平移参数隐藏数据的语音效果

子帧脉冲平移修改				SNR	SSNR	SD
一	二	三	四			
✓				10.550	9.402	0.223
✓	✓			8.087	6.408	0.272
✓	✓	✓		5.967	4.524	0.302
✓	✓	✓	✓	4.618	3.306	0.335

从表 3 可以看出,随着隐藏信息量逐渐增加,语音受到更多影响导致各项指标降低。收听解码效果,语音亲切自然,难以察觉失真,但能感到和隐藏信息量成正比的噪声。

与嵌入参数编码结果的方法相比,嵌入脉冲平移的方法解码后语音频谱距离更小,说明经过信息隐藏的语音在频谱上畸变较小,失真现象微弱符合人类听觉习惯,然而信噪比和分段信噪比略低,说明降低语音质量的主要原因是信息隐藏造成的大量噪声。

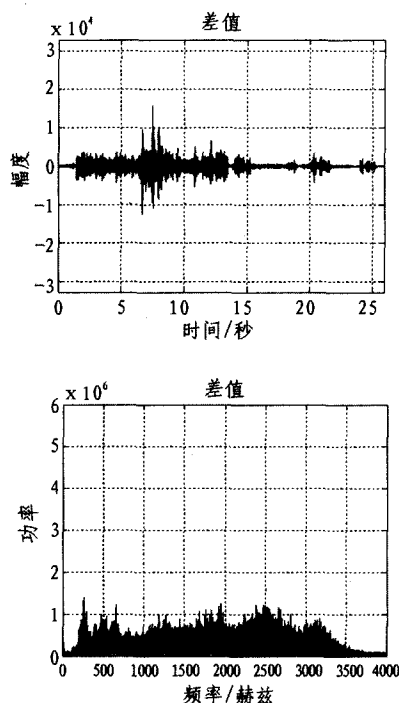


图 2 脉冲平移嵌入 4bit 数据的时域与频域波形图

图 2 表示子帧脉冲平移参数码字各隐藏 1bit 保密数据时解码语音的波形与 G. 723.1 标准解码语音之间差值的时域

波形和频域波形。

利用脉冲平移进行信息隐藏可以有效保证语音自然还原,但会引入噪声影响通信效果,这种方法的最主要局限性在于参数自身容量仅有 4bit 导致隐藏信息量较低,只能适用于低速情况。

2.6 1bit 的保留位

这是用于凑齐 1byte 的方便操作,没有实际意义,编码时默认值为 0,解码时直接跳过不做处理。所以无须对编解码器做任何修改,可以嵌入 1bit 保密数据。与数据类型码位类似,虽不干扰语音,但隐蔽性略逊。

2.7 73bit 的脉冲位置

73bit 的脉冲位置参数可以继续细分为每一子帧脉冲位置最高有效位组合 13bit 和第一、第三子帧脉冲位置各 16bit,第二、第四子帧脉冲位置各 14bit。

将脉冲位置按权重赋值,每一子帧计算的第一个脉冲位置权重最高赋值最大,第二个脉冲位置其次,最后计算的脉冲位置权重最低赋值最小。在奇数子帧中 6 个脉冲位置值相加得到 20bit 位置参数,在偶数子帧中 5 个脉冲位置值相加得到 18bit 位置参数。如果直接记录位置,每个脉冲存在 30 种可能性需要 5bit 编码,奇数子帧共要 30bit,偶数子帧共要 20bit,权重相加方法各节约 10bit 和 2bit。对每一子帧脉冲位置参数的最高四位有效位再编码,形成 13bit 索引。

分别对构成脉冲位置参数的 5 个部分的最低有效位取反,分析研究是否适合隐藏数据。

表 4 脉冲位置参数隐藏 1bit 数据的语音效果

隐藏 1bit 信息位置	SNR	SSNR	SD
子帧脉冲位置组合	3.642	4.262	0.349
第一子帧脉冲	11.818	12.730	0.137
第二子帧脉冲	12.916	12.433	0.145
第三子帧脉冲	13.202	12.757	0.150
第四子帧脉冲	11.861	12.054	0.220

从表 4 可以看出,保密数据嵌入子帧脉冲位置组合时,对语音质量影响较大,保密数据嵌入每一子帧脉冲位置时,对语音质量影响较小。原因是子帧脉冲位置组合包括每一子帧脉冲位置最高有效位,即使发生略微变化对还原效果的影响也超过每一子帧脉冲位置其他较低有效位。主观感觉嵌入子帧脉冲位置组合后语音产生失真,但在可以接受的范围内,嵌入每一子帧脉冲位置时,语音几乎没有发生任何畸变,和未嵌入保密数据的语音对比收听难发现差异,隐蔽性好,适宜进行信息隐藏。

采用不同方案增加嵌入每帧脉冲位置的信息量,选择保密数据嵌入脉冲位置最合适的手段和数量。

表 5 脉冲位置参数隐藏 2-10bit 数据的语音效果

第一子帧脉冲位置修改	SNR	SSNR	SD
2bit	9.104	11.376	0.149
3bit	9.157	10.543	0.160
4bit	9.163	9.887	0.171
5bit	8.938	9.146	0.183
6bit	7.446	8.512	0.197
7bit	7.432	7.999	0.197
8bit	6.676	7.392	0.211
9bit	6.052	6.838	0.224
10bit	4.896	6.341	0.231

观察表 5 发现脉冲位置隐藏保密数据的两个特点,一是前八组嵌入 2bit、后九组嵌入 4bit 表现有同一趋势,相同信息量单独嵌入某一子帧的测试结果优于分散嵌入两个子帧,分散嵌入两个子帧测试结果优于平均嵌入四个子帧,说明在脉冲位置隐藏信息集中嵌入方法效果最佳,二是使用相同方法隐藏相同信息量,如果处理涉及第四子帧,频谱距离将明显增加,前三子帧对语音影响基本相似。主观感受与客观数据相似,造成清晰度降低的主要原因是噪声增加。

奇数子帧的脉冲位置用 20bit 数值描述,偶数子帧的脉冲位置用 18bit 数值描述,少量数据嵌入某一子帧影响不大。由于预测编码经常需要上一帧第四子帧,如果用来隐藏数据,会同时干扰本帧和下一帧。根据上述讨论,可行的方法是选择前三子帧中的某一帧实施信息隐藏。

继续增加隐藏信息量,以嵌入第一子帧为例。

甚至当信息量高达 1byte,表 5 显示语音质量依旧能够接受,只是听起来环境比较嘈杂稍影响清晰度,有时必须集中精力收听,直至嵌入 10bit 之后语音出现明显失真,不便继续增加信息量。

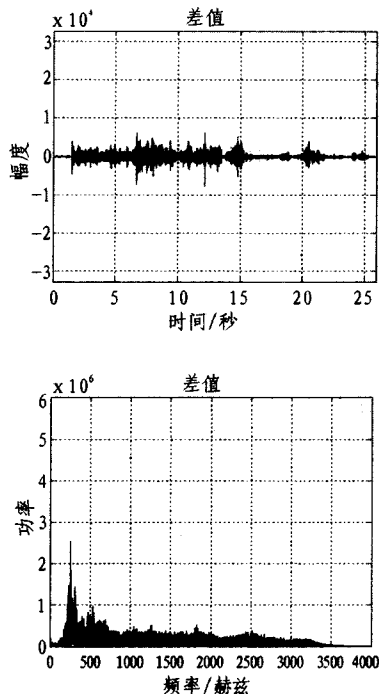


图 3 脉冲位置嵌入 4bit 数据的时域与频域波形图

图 3 表示第一子帧脉冲位置参数隐藏 4bit 保密数据时解码语音与 G. 723. 1 标准解码语音之间差值的时域波形和频域波形。

因为主要变更的是低权重脉冲信号的位置,所以即使大规模处理,信号语音质量降低也相当有限。时域波形体现隐藏效果和嵌入线谱对参数索引时很类似,频域差值的平均功率比其它方法更低,但存在其他方法没有出现的功率峰值。

嵌入脉冲位置是进行信息隐藏的理想途径之一,除了能兼顾语音质量和隐藏信息量以外,由于数据隐藏的位置相对集中,嵌入和提取算法实现起来更加方便,而频域功率峰值可能会引起攻击者注意降低系统的安全性。

2.8 22bit 的脉冲符号位

22bit 的脉冲符号参数可以继续细分为四个子帧脉冲符号。奇数子帧包含 6 个脉冲,需要 6bit 描述脉冲符号,偶数子

帧包含 5 个脉冲,需要 5bit 描述脉冲符号。

在实验中分别修改第一子帧和第二子帧中所有脉冲的极性,讨论语音质量受到影响的程度。

表 6 脉冲符号参数隐藏 1bit 数据的语音效果

第一子帧修改脉冲符号	SNR	SSNR	SD	第二子帧修改脉冲符号	SNR	SSNR	SD
第一位	5.239	7.952	0.200	第一位	7.141	7.360	0.214
第二位	6.008	7.888	0.210	第二位	6.194	7.366	0.221
第三位	8.334	8.474	0.207	第三位	7.758	7.824	0.213
第四位	8.538	8.686	0.196	第四位	9.350	8.463	0.186
第五位	9.865	9.199	0.190	第五位	10.287	9.009	0.170
第六位	10.073	9.403	0.184	第六位	10.073	9.403	0.184

实验结果如表 6 所示,可以发现数据存在一定的规律性。在同一子帧中,保密数据嵌入的符号位置越高,语音还原质量越好,而嵌入奇数子帧或偶数子帧同一位置的语音之间差别并不显著。主观感觉与之一致,在低位嵌入的语音无论噪声还是失真都比在高位嵌入的语音更严重。

出现这种现象的原因是,用来模拟残差信号的脉冲是依次单独进行计算的,首先寻找和原始残差信号相关性最高的第一个脉冲,并把相关信息保存在最低的位置,其次计算原始残差信号与第一个脉冲的区别,再寻找与差值相关性最高的第二个脉冲,并把相关信息保存在稍高的位置,然后计算原始残差信号与第一个脉冲、第二个脉冲的区别,如此反复直至得出算法规定的奇数子帧的 6 个脉冲、偶数子帧的 5 个脉冲为止。与残差信号相关性较低的脉冲信息放在高位,影响语音效果较小可以尝试进行信息隐藏,具体来说奇数子帧的第五位和第六位与偶数子帧的第四位和第五位。

按照从最高位开始、最多嵌入 2bit 的指导思想,每帧最多能隐藏 8bit 保密数据。比较利用不同子帧的脉冲符号高位隐藏不同信息量的保密数据对语音质量的影响。

表 7 脉冲符号参数隐藏 2-4bit 数据语音效果

子帧脉冲符号高位修改				SNR	SSNR	SD
一	二	三	四			
2	0	0	0	7.083	6.574	0.230
1	1	0	0	6.691	5.863	0.255
2	2	0	0	3.916	2.637	0.327
1	1	1	1	3.348	2.526	0.362

(上接第 179 页)

- [5] 陆锦军,王执铨. 基于混沌特性的网络流量预测[J]. 南京航空航天大学学报,2006,38(2):217-221
- [6] Vapnik V N. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995
- [7] 王晓丹,郑春颖,吴崇明. 一种新的 SVM 对等增量学习算法. 计算机应用,2006,26(10):2440-2443
- [8] 吴少雄,黄恩洲. 基于支持向量机的控制图模式识别. 计算机应用,2007,27(1):61-64

表 7 实验数据表明,无论数据分散隐藏到前半帧和后半帧还是集中隐藏到前半帧或后半帧,解码语音效果都不理想,隐藏信息量略微增加信噪比和分段信噪比迅速减小。收听感觉语音时常发生失真,强烈干扰使细节含混不清,机械感让听觉不舒适。

虽然在 G. 723.1 编码结果中脉冲符号占有超过 10% 的份额,但即使隐藏少量数据也会造成语音发生可感知的变化,令通信过程失去了最基本的保密性,因此不使用脉冲符号进行信息隐藏。

结束语 本文基于信息隐藏及其特点,尝试讨论一种在语音混沌保密通信过程中进行信息隐藏方案中,对 G. 723.1 高码率编码语音中适合进行压缩域信息隐藏的码位进行进一步的性能分析。提出了具体的实验方案。并综合实验结果,6.3kbps 算法中适合信息隐藏的位置有数据类型位、线谱对参数索引、脉冲平移、保留位、脉冲位置,最多分别可以隐藏 2bit,9 bit,4 bit,1 bit,8 bit 保密数据。

嵌入数据类型位、保留位无损语音质量,隐藏信息量虽小却能共同使用略作弥补,因容易被察觉,适用于隐藏非关键数据;嵌入线谱对参数索引、脉冲平移、脉冲位置通过牺牲部分效果提高安全性能,但由于对语音有破坏,不能叠加使用以免累积误差,适用于隐藏关键数据。实验结果进一步证明了文献[1]方案的有效性和安全性。

参考文献

- [1] 何招友,季晓勇. 基于混沌理论的语音保密通信技术. 信息安全与通信保密,2006(7):68-70
- [2] 倪皖荪,华一满. 混沌通讯. 物理学进展,1996,16(3-4):645-655
- [3] 代明,张宇,华一满,倪皖荪. 超混沌迭代同步在数字语音保密通讯中的应用. 南京大学学报(自然科学),1999,35:110-113
- [4] 陆佩芸,许留荣,季晓勇. 基于 DSP 的嵌入式语音保密通信模块. 微处理机,2003(10)
- [5] ITU-T Recommendation G. 723.1: DUAL RATE SPEECH CODER FOR MULTIMEDIA COMMUNICATIONS TRANSMITTING AT 5.3 AND 6.3 kbit/s
- [6] 张雄伟,曹铁勇. DSP 芯片的原理与开发应用. 北京:电子工业出版社,2000
- [7] 韩纪庆,张磊,郑铁然. 语音信号处理. 北京:清华大学出版社,2004
- [8] 蔡莲红,黄德智,蔡锐. 现代语音技术基础与应用. 北京:清华大学出版社,2003
- [9] 苏晓生. 掌握 MATLAB 6.0 及其工程应用. 北京:科学出版社,2002
- [9] 董辉,付鹤林,冷伍明. 基于支持向量机的时间序列回归与预测[J]. 系统仿真学报,2006,18(7):1785-1788
- [10] 刘涵,刘丁,李琦. 基于支持向量机的混沌时间序列非线性预测[J]. 系统工程理论与实践,2005,9:94-99
- [11] 周红刚,杨春德. 基于免疫算法与支持向量机的异常检测方法[J]. 计算机应用,2006,26(9):2145-2147
- [12] 杨虞微,左洪福,陈果. 支持向量机时间序列预测模型的参数影响分析与自适应优化[J]. 航空动力学报,2006,21(4):767-772