

基于方差及线性回归的语音隐藏盲检测方法

汪云路, 程义民, 田 源, 谢于明

(中国科学技术大学电子科学与技术系, 合肥 230026)

摘要: 描述了一种基于方差及线性回归的语音隐藏信息盲检测方法。该方法用语音质量评估参数生成一个多元线性回归模型, 可对可疑的语音进行隐藏信息盲检测, 语音质量评估参数种类和数量, 可用方差分析法预先选择。该方法经计算机模拟, 对常用的变换域信息隐藏方法有较好的检测效果。

关键词: 语音信息隐藏; 盲检测; 语音质量评估参数; 方差分析; 线性回归模型

中图分类号: TB556

文献标识码: A

文章编号: 1000-3630(2008)-01-0071-05

Blind audio steganalysis based on variance and multivariate regression analysis

WANG Yun-lu, CHENG Yi-min, TIAN Yuan, XIE Yu-ming

(University of Science and Technology of China, Hefei 230026, China)

Abstract: An audio steganalysis method based on variance and multivariate regression analysis, which can be exploited for audio detection, is studied. With the aid of audio quality measures a multivariate regression scheme is proposed to detect probable stego-audio. Audio quality measures are chosen based on ANOVA (analysis of variance) technique. Simulation results indicate that this method generally works well for transform domain steganographic.

Key words: audio message hiding; blind detecting; audio quality measures; ANOVA; multivariate regression scheme

1 引 言

随着信息技术和计算机技术的发展, 信息隐藏作为信息安全的一个重要方面, 已经获得了广泛的应用。伴随信息隐藏技术的不断提出, 隐藏检测方法也受到人们的关注。在现代语音通信网络发展的同时, 语音信息隐藏与分析也越来越为人所重视。

2005年, Oktay Altun 等人采用 n 维形态学变换和边缘失真, 测度未隐藏信息的音频和隐藏信息的音频之间的统计差异, 从而检测隐藏信息^[1]; 同年 Xuemin Ru 等人提出了一种针对 wav 格式的语音隐藏信息侦测算法。该方法先对语音信号进行 4 阶 1 维小波分解, 通过对小波带系数值的线性预测, 提

取出有效的统计特征, 再利用支持向量机 (SVM, support vector machine) 判断是否隐藏有信息^[2]; 同年 Sorina Dumitrescu 等人提出了一种基于高阶统计量的检测模型。该方法通过建立视频载体的高阶统计量模型, 检测是否存在隐藏信息, 适用于最低有效位 LSB (least significant bits) 隐藏算法。该方法同样也适用于图像和音频^[3]。

本文给出了一种基于方差及线性回归的语音隐藏盲检测方法, 利用语音质量评估参数和方差估计法, 选择能反映藏有信息的数字语音和未藏信息的数字语音在统计上差异的参数, 建立质量评估参数的线性回归模型, 并使用该模型进行语音信息隐藏的盲检测。

2 系统概述

图 1 为基于方差分析 (ANOVA) 及线性回归模

收稿日期: 2006-12-05; 修回日期: 2007-04-24

作者简介: 汪云路 (1981-), 女, 安徽人, 博士研究生, 研究方向: 信息隐藏分析、网络多媒体等。

通讯作者: 汪云路, E-mail: cvlab@mail.ustc.edu.cn

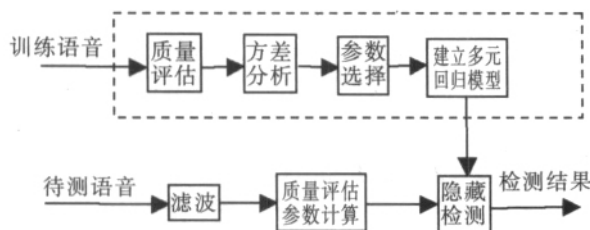


图1 语音信息隐藏盲检测方法框图

Fig.1 Schematic description of blind audio steganalysis based on ANOVA

型的语音信息隐藏盲检测方法框图。它主要包括两个部分,多元回归模型的建模以及语音信息隐藏的盲检测。

框图的上半部分虚线内为多元线性回归模型的建模部分,训练语音包括原始语音子集和藏有信息的语音子集。经中值滤波,由滤波前后的语音集计算各个质量评估参数值。对每一质量评估参数进行方差分析,选择那些能较好反映原始语音及藏有信息的语音统计上差异的参数,建立多元回归模型。这些种类的质量评估参数及相应模型,可用于语音信息隐藏盲检测。

框图的下半部分为语音信息隐藏检测部分。对待测语音先进行中值滤波处理,再由滤波前后语音计算预先选择的那些种类的质量评估参数及相应评估值,进而应用预先训练获得的多元线性回归模型,最终获得语音中是否藏有信息的检测结果。

3 语音质量评估参数的选择

经过信道传输后接收到的语音信息通常具有稳定且满足系统要求的语音质量,因此,选择语音质量评估参数进行信息隐藏盲检测,有一定的稳定性和实用性。本文用方差分析方法,经过实验主要选用以下五种常用且有效的质量评估参数,来建立多元回归盲检测模型。

(1) SNR 信噪比

$$SNR = \frac{\sum_{i=1}^N (X_i)^2}{\sum_{i=1}^N (X_i - Y_i)^2} \quad (1)$$

式中: X 和 Y 分别为输入和输出语音(即滤波前语音和滤波后语音,以下同), N 为语音总采样点数。

(2) Mel 谱距离^[4]

定义 M 个等三角形带宽响应的滤波器, 则对总长 L 帧的输入和输出信号 X 及 Y 的 Mel 谱距

离为:

$$MelD = \frac{1}{L} \sum_{l=1}^L \sum_{i=1}^M \left[\sum_{k=1}^M X_k^l \cos \left[i \left(k - \frac{1}{2} \right) \frac{\pi}{M} \right] - \sum_{k=1}^M Y_k^l \cos \left[i \left(k - \frac{1}{2} \right) \frac{\pi}{M} \right] \right]^2 \quad (2)$$

$i=1, 2, \dots, M$

式中: X_k^l 、 Y_k^l ($k=1, 2, \dots, M$) 分别为输入与输出语音信号第 l 帧第 k 个滤波器的对数谱能量输出, i 为 Mel 倒谱系数的个数。

(3) LPC 倒谱距离^[5]

对总长 L 帧的输入和输出语音信号 X 及 Y , 则 LPC 倒谱距离^[5]根据第 l 帧输入和输出信号的第 k 阶 LPC 系数, 计算出各自的倒谱系数 $C_x(l, k)$ 和 $C_y(l, k)$, 再求出它们之间的倒谱距离(P 为预测器的阶数):

$$LPCD = \frac{1}{L} \sum_{l=1}^L \sqrt{[C_x(l, 0) - C_y(l, 0)]^2 + 2 \sum_{k=1}^{P-1} [C_x(l, k) - C_y(l, k)]^2} \quad (3)$$

(4) LPC 倒谱联合线性频谱动态特性^[6]

用倒谱系数对时间的线性回归分析来表示频谱动态特性, 则:

$$A_x(l, k) = \frac{\sum_{j=-3}^3 C_x(l+j, k)}{\sum_{j=-3}^3 j^2} \quad (4)$$

$A_x(l, k)$ 和 $A_y(l, k)$ 分别为第 l 帧的输入和输出语音信号 X 及 Y 第 k 阶倒谱系数的线性回归系数, 则 LPC 倒谱联合线性频谱动态特性可用下式表示:

$$LPCD \pm \frac{1}{L} \sum_{l=1}^L \sqrt{\sum_{k=1}^P [C_x(l, k) - C_y(l, k)]^2 + 3.84 \sum_{k=1}^P [A_x(l, k) - A_y(l, k)]^2} \quad (5)$$

4.84

(5) 联合 LPC 倒谱系数、倒谱回归系数与能量^[7]

考虑到信号失真时其能量的变化, 则联合 LPC 倒谱系数、倒谱回归系数和能量的对数值同时作为语音信号失真的测度, 可得到^[7]:

$$LPCD' \pm \frac{1}{L} \sum_{l=1}^L \sqrt{\sum_{k=1}^P [C_x(l, k) - C_y(l, k)]^2 + 10 \sum_{k=1}^P [A_x(l, k) - A_y(l, k)]^2 + 0.34 [C_x(l, 0) - C_y(l, 0)]^2} \quad (6)$$

11.34

以上初选的参数还需要进一步通过方差分析法, 判断该质量评估参数在统计上是否能将隐藏信息的语音与未藏信息语音的特征值区分开。

4 方差分析

方差分析法^[9]常用于两个或两个以上样本均数差别的显著性检验, 通过分析研究不同来源的变异对总变异的贡献大小, 从而可确定不同来源对研究结果的影响。

若 A 是包括原始语音和 $k-1$ 种信息隐藏方法的语音样本集合, 共有 $M=n*k$ 个语音样本, 可表示为 $A=\{A_1, A_2, \dots, A_k\}$, 其中 A_1 是包含 n 个原始语音样本的子集, $A_i (i=2, \dots, k)$ 是第 i 种信息隐藏方法得到的语音样本子集。 A_i 经过中值滤波后, 组成一个新的子集 A_i' ; 通过 A_i 和 A_i' 中相对应元素可计算出第 j 个语音质量评估参数的子集 $X_{ij}=\{X_{i1j}, X_{i2j}, \dots, X_{imj}\}$, 其中 $m=1, 2, \dots, n$ 为语音样本个数, $j=1, 2, \dots, p$ 为语音评估参数个数; 故 A 集合中所有语音样本第 j 个质量评估参数集合可表示为 $X_j=\{X_{1j}, X_{2j}, \dots, X_{kj}\}$ 。

由方差分析法知, 对第 j 个质量评估参数, 可计算相应 F_j 值:

$$F_j = MS_{\text{组间}} / MS_{\text{组内}} \quad (7)$$

$$\text{其中, } MS_{\text{组间}} = \frac{\sum_{i=1}^k n(\bar{X}_{ij} - \bar{X}_j)^2}{k-1} \quad (8)$$

$$MS_{\text{组内}} = \frac{\sum_{i=1}^k \sum_{m=1}^n (X_{imj} - \bar{X}_{ij})^2}{N-k} \quad (9)$$

其中, $\bar{X}_{ij}$ 为子集 A_i 第 j 个质量评估参数的平均值, $\bar{X}_j$ 为 A 集合中第 j 个质量评估参数的总平均值。 n 为子集 A_i 中元素个数, 为方便, 这里每一个子集 $A_i (i=1, \dots, k)$ 中元素个数取同为 n ; 实际各子集中元素个数不一定相同。 $MS_{\text{组间}}$ 为组间平均平方和, 反映原始语音和用不同隐藏方法嵌入信息的语音, 经中值滤波后质量评估参数间的差异。 $MS_{\text{组内}}$ 为组内平均平方和, 反映用同一种隐藏方法嵌入信息的语音, 经过中值滤波后质量评估参数间的差异。

这样 F 值大小便反映出各质量评估参数对信息隐藏前后的语音质量参数变化的大小。

选出 q 个能反映隐藏前后语音质量差异较大的质量评估参数, 记为 $\{X_{ij}|j=1, 2, \dots, q\}$, 可用于线性回归模型建模及隐藏分析。

5 多元线性回归模型

5.1 建模

若训练语音集为 $B=\{b_1, b_2, \dots, b_N\}$, b_i 为第 i 个训练语音样本。 b_i 经过中值滤波后, 计算经优选出的 q 个质量评估参数值, 记做 $Z=\{z_{ij}|i=1, 2, \dots, N, j=1, 2, \dots, q\}$, 式中 N 为语音总样本数, 包括原始语音和用不同隐藏方法嵌入信息的语音; q 为经优选出的质量评估参数的个数。用线性回归模型^[9], 可将 z_{ij} 与随机误差 ε_i 表示为:

$$\begin{cases} y_1 = \beta_1 z_{11} + \dots + \beta_j z_{1j} + \dots + \beta_q z_{1q} + \varepsilon_1 \\ y_2 = \beta_1 z_{21} + \dots + \beta_j z_{2j} + \dots + \beta_q z_{2q} + \varepsilon_2 \\ \vdots \\ y_i = \beta_1 z_{i1} + \dots + \beta_j z_{ij} + \dots + \beta_q z_{iq} + \varepsilon_i \\ \vdots \\ y_N = \beta_1 z_{N1} + \dots + \beta_j z_{Nj} + \dots + \beta_q z_{Nq} + \varepsilon_N \end{cases} \quad (10)$$

$$\text{简化为矩阵形式, 即: } Y = Z\beta + \varepsilon \quad (11)$$

其中, $Y=[y_1, y_2, \dots, y_N]^T$, 若 $z_{11}, z_{12}, \dots, z_{iq}$ 为藏有信息的语音样本的 q 个质量评估参数值, 则 $y_i=1$, 否则 $y_i=-1, i=1, 2, \dots, N$ 。 $\beta=[\beta_1, \beta_2, \dots, \beta_N]^T$, $\varepsilon=[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N]^T$ 。有许多优化方法, 可给出一个令 ε 最小的 β 的最优估值 $\hat{\beta}$, 限于篇幅这里仅给出用最小二乘法的结果。

由多元回归的最小二乘法可知^[10], 线性回归模型中系数 β 的最小二乘估计, 即最大似然矩阵 $\hat{\beta}$ 为:

$$\hat{\beta} = (Z^T Z)^{-1} (Z^T Y) \quad (12)$$

5.2 检测

对测试语音集 $C=\{c_1, c_2, \dots, c_M\}$, c_i 为待测语音样本。 c_i 经过中值滤波后, 分别计算相应的 q 个质量评估参数值, 得到 $R_i=\{r_{ij}|j=1, 2, \dots, q\}$, 代入下式:

$$\hat{y}_i = \hat{\beta}_1 r_{i1} + \hat{\beta}_2 r_{i2} + \dots + \hat{\beta}_q r_{iq} \quad (13)$$

若 $\hat{y}_i \leq 0$, 表示 c_i 无隐藏信息; 若 $\hat{y}_i > 0$, 则 c_i 有隐藏信息。

6 实验结果

6.1 质量评估参数选择

从语音数据库中随机选取 100 段格式为 wav (wave, 波形格式数据) 的原始语音样本, 分别对它们用 FFT、DCT 和 DWT 域信息隐藏方法嵌入等量的数据, 共有 $k=4$ 组数据, 每组 100 个样本, 共计 $N=400$ 个样本, 表 1 给出 F 值较好的四个参数, Mel 谱

距离参数、LPC 倒谱距离参数、LPC 倒谱联合线性频谱动态特性、联合了 LPC 倒谱系数、倒谱回归系数与能量以及 SNR 信噪比等参数方差分析的实验结果。

表 1 部分语音评估参数的方差分析结果
Table 1 ANOVA tests for several audio quality measures

语音质量评估参数	方差分析法 F 值
SNR信噪比	0.326692
Mel谱距离	3.40335
LPC倒谱距离	3.45961
LPC倒谱联合线性频谱动态特性	3.21185
LPC倒谱、倒谱回归系数与能量三者融合	3.35765

由表 2 可知,除了 SNR 信噪比之外,其他四种语音质量评估参数,均大于 $F_{0.05}(3, 396)$,而小于 $F_{0.01}(3, 396)$ 。根据方差分析的定义,表明对于这几个质量评估参数,有无隐秘信息的嵌入,在统计学上有较好的可分性。

6.2 信息隐藏盲检测实验

为验证本文提出的方法的正确性,我们进行基于统计学的语音隐藏盲检测实验。训练集同上,图 2 给出其中一段样本经中值滤波后,分别用 Mel 谱距离参数、LPC 倒谱距离参数、LPC 倒谱联合线性频谱动态特性及 LPC 倒谱系数、倒谱回归系数、能量三者联合这四种质量评估参数进行语音质量评估的实验结果。

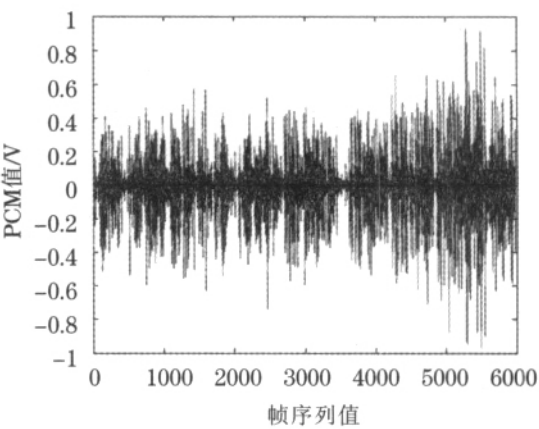
用训练集质量评估参数和式(10)~(12)建立多元回归模型,并求出似然矩阵;随机选取 20 段人声 wav 语音作为测试集,测试集与训练集交集为 ,用 FFT、DCT、DWT 域信息隐藏方法嵌入等量数据;再分别选择语音数据库中的 80 段人声语音与之混合,分别进行中值滤波后,用上述语音质量评估参数计算参数值,用式(13)进行检测得最终检测结果。

表 2 隐藏信息盲检测率表
Table 2 Performance of the steganalyzer for all the experiments

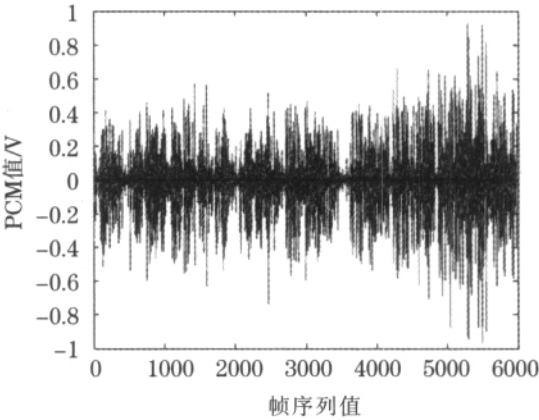
	误检率	漏检率	正确检测率
FFT	17%	6%	77%
DCT	17%	7%	76%
DWT	18%	4%	78%

7 结 论

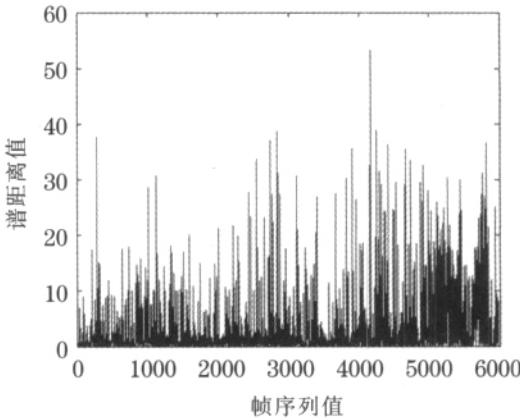
本文提出了一种基于统计学的信息隐藏分析方法,通过使用常用语音质量评估参数对信息隐藏前



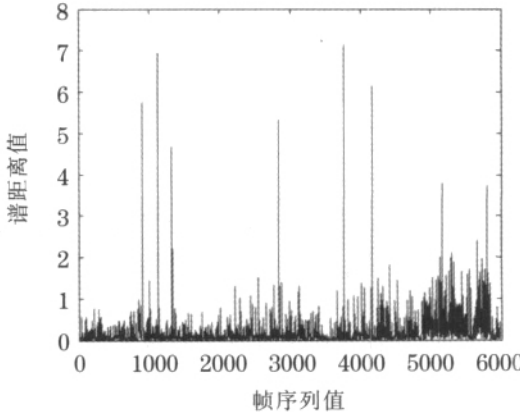
(a)原始语音



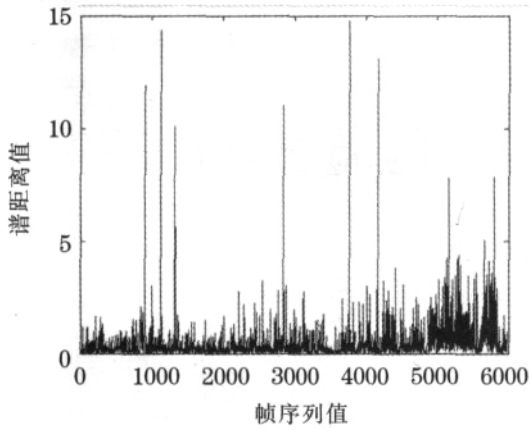
(b)中值滤波后语音



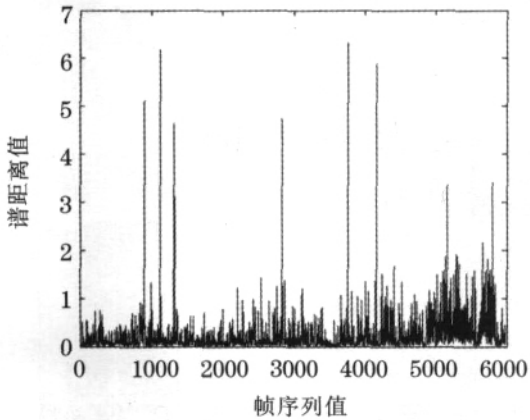
(c)图(a)(b)的 Mel 谱距离



(d)图(a)(b)的 LPC 倒谱距离



(e)图(a)(b)的 LPC 倒谱联合线性频谱动态特性



(f)图(a)(b)的 LPC 倒谱系数、倒谱回归系数
与能量三者联合

图 2 实验语音的四种质量评估参数(谱距)

Fig.2 Four quality measures (spectrum distances)
of testing audio

后的语音进行质量评估, 并对评估的数据进行方差分析, 选用其中的有效质量评估参数来建立多元线性回归模型, 使用此模型实现语音信息隐藏的盲检测。实验表明, 该方法运算简便, 检测效果较好。

参 考 文 献

[1] Altun Oktay, Morphological steganalysis of audio signals

and the principle of diminishing marginal distortions[C]. Acoustics, Speech, and Signal Processing, 2005, 2: 21-24.

[2] RU Xuemin, ZHANG Hongjuan, HUANG Xiao. Steganalysis of audio: attacking the Steghide[J]. Machine Learning and Cybernetics, 2005, 7: 3937-3942.

[3] S Dumitrescu, WU Xiaolin. On global optimality of gradient descent algorithms for fixed-rate scalar multiple description quantizer design[C]. Data Compression Conference, 2005: 388-397.

[4] 陈国, 胡修林, 张蕴玉, 等. 语音质量客观评价方法研究进展[J]. 电子学报, 2001, 29(4): 548-552.

CHEN Guo, HU Xiulin, ZHANG Yunyu. Research advance on objective measures of speech quality[J]. Acta Electronica Sinica, 2001, 29(4): 548-552.

[5] 黄惠明, 王瑛, 赵思伟, 等. 语音系统客观音质评价研究[J]. 电子学报, 2000, 28(4): 112-114.

HUANG Huiming, WANG Ying, ZHAO Siwei. Study of objective quality evaluation for the speech systems[J]. Acta Electronica Sinica, 2000, 28(4): 112-114.

[6] 黄惠明, 师聘, 线性预测编码倒谱系数距离(LPC-CD)在语音信号质量客观测度中的应用[J]. 计算机与网络, 1999, 25(16): 25-26.

HUANG Huiming, SHI Cheng. Technology of objective measuring speech quality-linear predictor code cepstrum-distance[J]. Computer and Networks, 1999, 25(16): 25-26.

[7] 吴淑珍, LouisC.W.Pols. 汉语通信系统客观音质评价方法[J]. 声学学报(中文版), 1998, 23(2): 170-174.

WU Shuzhen, LouisC.W.Pols. A measure for objective speech quality evaluation of China communication channels [J]. Acta Acustica(Chinese), 1998, 23(2): 170-174.

[8] 辛益军. 方差分析与实验设计, 北京: 中国财政经济出版社[M]. 2001, 49-65.

XIN Yijun. Analysis of variance and experiments design[M]. Beijing: China Financial and Economic Publishing House, 2001, 49-65.

[9] 徐志敏, 刘希玉, 多元回归分析及其在辅助决策中的应用[J]. 计算机工程, 2004, 30(21): 129-131.

XU Zhimin, LIU Xiyu. Multivariate regression analysis and the application in assistant decision-making[J]. Computer Engineering, 2004, 30(21): 129-131.

[10] 刘晓莉. 试验设计中多元回归分析方法的研究[J]. 数理统计与管理, 2001, 20(4): 14-16.

LIU Xiaoli. The research for method of regression analysis on experiment design[J]. Application of Statistics and Management, 2001, 20(4): 14-16.