基于 F-范数的小波自适应量化的语音伪装*

陈 亮[1], 刘 实[2], 李云剑[1]
(1. 解放军理工大学通信工程学院, 江苏 南京 210007;
2. 北京理工大学信息技术学院, 北京 100081)

摘 要: 提出一种小波域自适应量化的语音伪装算法。首先对语音信号进行小波变换, 由混沌序列控制嵌入位置, 并采用固定阶距量化实现秘密信息的隐藏。在此基础上通过 F-范数自适应控制量化阶距的变化, 实现了基于 F-范数的语音伪装算法。仿真实验表明算法安全性高、透明性好, 在多种干扰情况下均能盲提取出秘密信息。

关键词: F-范数; 小波; 自适应量化
中图分类号: TP309 文献标识码: A 文章编号: 0529-6579 (2004) S2-0046-04

信息伪装是在 20 世纪 90 年代中后期产生的一门新兴研究领域。信息伪装通过将数字秘密信息隐藏于图像、语音、视频等公开媒体之中传输, 不影响公开媒体的主观特性, 从而达到隐蔽通信的目的。

本文从量化角度出发研究语音信号的伪装算法并提出一种基于 F-范数的小波自适应量化的语音伪装算法。此算法充分利用小波变换的多分辨特性, 将极低速率编码形成的保密语音信号用量化方法隐藏在公开语音信号小波变换域的低频系数中, 量化阶距自适应于公开语音信号被量化小波系数的 F-范数, 并且嵌入后人耳不易察觉到公开语音的微小变化, 系统透明性和鲁棒性高。经噪声、滤波、压缩编码后, 在接收端能恢复出隐藏的保密语音信号。此算法还运用混沌序列对隐藏过程进行了加密, 系统安全性好。

1 秘密信息在语音中的嵌入和提取

1.1 小波变换与小波基的选取

小波变换是当前应用数学和工程学科中一个迅速发展的新领域, 属时频分析的一种。小波在时域和频域都具有表征信号局部特征的能力, 且在低频部分具有较高的频率分辨率和较低的时间分辨率, 在高频部分具有较高的时间分辨率和较低的频率分辨率。本文信息伪装主要利用了小波分析与耳蜗特性一致的多分辨分析思想, 采用 Daubechies 紧支撑小波对语音进行多分辨分析, 该小波函数在有限长 FIR 滤波器中具有最大正则性, 小波的波形比较光滑, 时频局部化特性也比较好。由 Mallat 重构公式, 利用高频系数和低频系数进行语音信号重构。由于小波变换具有很好的能量压缩能力, 能量大部分集中于低频分量, 所以在低频段嵌入信息后人耳不易察觉到信号的微小变化, 能够增强系统的透明性; 同时低频分量的感知容量较大, 允许系数有较大的量化误差, 这可以保证系统的鲁棒性。

1.2 小波域量化的语音伪装方案

1.2.1 秘密信息的嵌入 将原始语音 S 分帧进行 L 级 Mallat 塔式小波分解, 帧长为 Fm, 得到小波系数 C:

$$C = \{ c_{A_L}, c_{D_L}, c_{D_{L-1}}, c_{D_{L-2}}, \cdots, c_{D_1} \}$$

其中, c_{A_L} 为第 L 级小波低频系数; c_{D_m} ($1 \leqslant m \leqslant L$) 为第 m 级小波高频系数。将第 L 级小波低频系数 c_{A_L} 进行量化后嵌入秘密信息。为了增加系统的嵌入容量, 减少计算量, 采用 Daubechies-K 的小波基进行一级小波分解后再量化的方法。

利用伪装系统中分配的用户码获得一密钥 K_2, 利用 K_2 来控制混沌随机数产生器。假设一帧嵌入 N bit ($N \leqslant Fm/2$), 则混沌随机数产生器产生 N 个 [1, $Fm/2$] 之间的随机整数 (若 1 帧中有重复的整数, 则删除第 2 次出现的整数, 并重新产生 1 个随机数), 随机数代表了低频系数的位置。在随机数确定的嵌入位置上, 利用比特 w_i ($w_i \in \{0, 1\}$) 对相应的小波系数进行标量量化, 从而实现信息的嵌入。

假设混沌随机数确定的 N 个系数为: $C_1 = [c_1, c_2, \cdots, c_N]$, 标量量化器的量化台阶为 Δ, 由量化原理可以得到嵌入和提取算法。首先,

* 收稿日期: 2004-09-11
作者简介: 陈亮 (1974 年生), 男, 博士; E-mail: liang_ch@sina.com

根据量化台阶 Δ 对小波系数 c_i 进行取模和余数的运算，假设：

$$m_i = \left[\frac{c_i}{\Delta}\right] \tag{1}$$

$$r_i = c_i - m_i\Delta \tag{2}$$

其中，m_i 代表 c_i 的模，r_i 代表 c_i 的余数。然后，根据比特 w_i 和 m_i 的奇偶性质确定 c_i 的量化值。根据 w_i 和待量化的小波系数 c_i 所在的区间对 c_i 进行量化处理。量化过程可以用下面的数学公式表示：

$$\tilde{c}_i = \begin{cases} m_i\Delta + \dfrac{1}{2}\Delta & m_i \oplus w_i = 0 \\ (m_i - 1)\Delta + \dfrac{1}{2}\Delta & m_i \oplus w_i = 1,\ r_i \leqslant \dfrac{\Delta}{2} \\ (m_i + 1)\Delta + \dfrac{1}{2}\Delta & m_i \oplus w_i = 1,\ r_i > \dfrac{\Delta}{2} \end{cases} \tag{3}$$

$\oplus$表示模二加，可以看出，由量化所引起的误差为：

$$\varepsilon = |c_i - \tilde{c}_i| \leqslant \Delta \tag{4}$$

也即量化所造成的最大量化误差 $\varepsilon_{\max} = \Delta$。

将 $W = \{w_i\}_{i=1}^{N}$ 按照式(3)方法嵌入到小波系数 C_1 中，得到 $\tilde{C}_1 = \{\tilde{c}_i\}_{i=1}^{N}$，然后根据密钥 K_2 将 $\tilde{c}_i$ 恢复到各自对应的位置，由此得到嵌入信息后的小波系数 C。对 C 进行小波重构，即产生嵌入秘密信息后的中间语音 S。

1.2.2　*秘密信息的提取*　秘密信息 W 是由 C_1 惟一确定的。根据 C_1 中各分量位于量化区间的位置即可确定二值秘密信息比特。首先，将中间语音 S' 分帧进行小波分解，得到 C'。然后，根据密钥 K_2 控制混沌随机数产生器同步产生 N 个[1, $Frm/2$] 之间的随机整数，以此确定秘密信息嵌入到低频系数的位置。假设混沌随机数所确定的 N 个系数为：

$$C'_1 = [\tilde{c}'_1, \tilde{c}'_2, \cdots, \tilde{c}'_N]$$

然后，根据量化台阶 Δ 对 $\tilde{c}'_1$ 取模：

$$m'_i = \left[\frac{\tilde{c}'_1}{\Delta}\right] \tag{5}$$

即可得该系数所嵌入的秘密信息比特为：

$$w'_i = m'_i \oplus 0 \tag{6}$$

以此方法依次计算，便可得到秘密信息序列

$$W' = \{w'_i\}_{i=1}^{N}$$

2　采用 F-范数的算法改进

采用量化算法进行语音伪装时，量化阶距 Δ 对于系统的鲁棒性与透明性有直接联系。因为语音信号为非平稳信号，具有时变特性，即有的语音段(如浊音段)小波系数的能量很高，有的语音段(如清音段或无声段)的小波系数能量非常小，因此采用固定阶距的算法难以适应不同情况下的变化，不能保证系统在具有鲁棒性的同时还保持较高的透明性。仿真实验显示，采用固定阶距 Δ 量化方式将信息隐藏在语音中后，虽然在接收端能恢复隐藏的秘密信息，但中间语音在清音段或无声段失真较明显，能被人耳感知，从而影响了保密语音的安全性。

采用自适应量化阶距的方法，可以较好地解决该问题。自适应量化阶距可以跟踪信号特性的变化，自适应地进行调整，从而在保证系统鲁棒性的条件下使中间语音的音质又有了较大幅度的提升。具体算法如下：

一帧语音中，由混沌随机数确定的 N 个系数构成的矢量为：

$$C_1 = [c_1, c_2, \cdots, c_N]$$

量化阶距 Δ' 可根据本帧 C_1 的 F-范数确定：

$$\Delta' = k \| C_1 \|_F \tag{7}$$

其中，k 与鲁棒性要求以及 C_1 的维数有关，可以根据系统环境的不同要求来设定。F-范数可以反映出系数的能量变化情况：在能量较高的语音段，$\| C_1 \|_F$ 较高，从而抬高量化阶距，由于能量高的语音段的感知容量较高，因而量化阶距的升高并不会带来音质的下降；能量较低的语音段的感知容量较小，但 $\| C_1 \|_F$ 也相应减小，从而降低了量化阶距，量化误差随之减小，保证了中间语音的音质。

由于算法在接收端为盲检测，没有原始语音参与秘密信息的提取，故不能获得发送端精确的 Δ'，只能采用估计的方法。因为嵌入秘密信息时，量化使得嵌入秘密信息后的信号相当于叠加了噪声，这种噪声可以认为与信号是不相关的，其均值为零，方差为 σ^2，所以有下式近似成立：

$$\| C_1 \|_F \approx \| C_1 \|_F \approx \| C'_1 \|_F \tag{8}$$

上式表明，接收端通过计算中间语音的 F-范数，即可估计量化阶距：

$$\Delta'_{est} = k \| C'_1 \|_F \tag{9}$$

图 1 比较了发送端与接收端量化阶距的幅度变化情况以及对应的误差，从图中可见自适应量化阶距可以及时跟踪语音的变化，从而有效提高系统的透明性。

3　仿真及性能分析

为了消除实验室环境和条件等多种主客观因素的影响，这里采用归一化相关系数 ρ 来定量评估提取的水印 W' 和原始 W 的相似性，即：

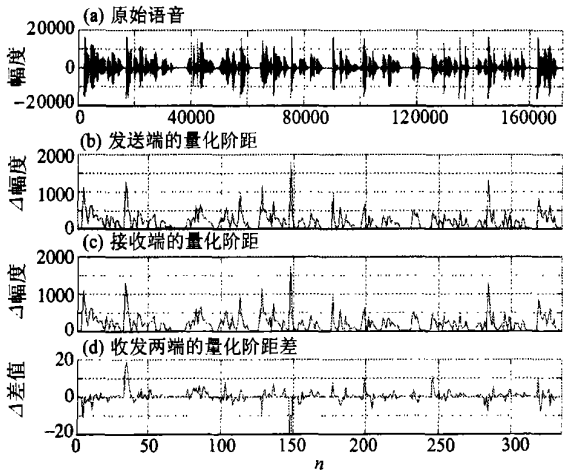

图 1 发送端与接收端的量化阶距以及对应的误差

Fig. 1 Quantization steps and error between transmitter and receiver

$$\rho(W, W') = \frac{W^t \circ W'}{\sqrt{(W^t \circ W)(W'^t \circ W')}} \quad (10)$$

在实验中，采用多段 8 kHz 采样，16 bit 线性量化的语音信号作为语音载体，每帧语音 512 点并采用 Daubechies-1 小波基进行一级小波分解，然后在低频系数中嵌入 64 bit（最多可嵌入 256 bit），量化阶距 $\Delta = 300$。在无攻击情况下，本算法的归一化相关系数 $\rho = 1$，隐藏的极低速率语音可以完全无失真地恢复出来，并且中间语音 S 的平均分段信噪比为 31.986 2 dB。图 2 画出了一段原始语音与中间语音的波形比较图。可见，从波形上比较，原始语音与中间语音是非常相近的。

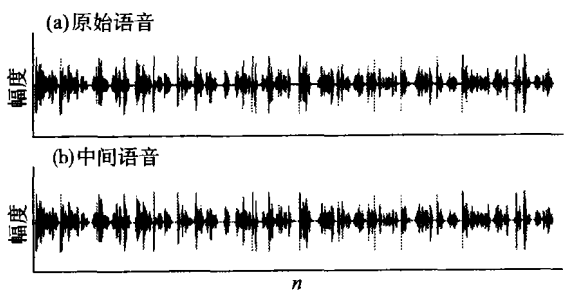

图 2 原始语音与嵌入水印后的中间语音的波形比较图

Fig. 2 The Comparion between original speech and middle speech with watermarking embedded

实验中也比较了在有噪声攻击的情况下算法的性能。采用 NATO-RSG-10 标准噪声库，通过对中间语音加入不同类型和不同信噪比的噪声，比较在不同噪声攻击情况下接收端恢复水印的性能。所有数据为 19.98 kHz 采样，16 bit 量化。这里将全部数据经过低通滤波和降采样率，得到 8 kHz 采样的噪声。实验中主要使用了下面使用了 13 种背景噪声：

(1) Babble Noise——大约有 100 个人在讲话的餐厅中的嘈杂声；

(2) F-16——美军 F16 双座战机巡航时座舱内的噪声（F16）；战机在 300～600 英尺上空以 500 节的速度巡航；

(3) Factory1——汽车工厂中金属切割的噪声；

(4) Factory2——汽车工厂焊接车间的噪声；

(5) Leopard——Leopard 战车以 70 km 的时速行驶时的车内噪声；

(6) White Gaussian Noise——高斯白噪声；

(7) Buccaneer jet——Buccaneer 战机以 450 节巡航时的座舱噪声；

(8) Destroyer operations room——驱逐舰设备控制室噪声；

(9) Destroyer engine room——驱逐舰轮机舱噪声；

(10) HF Channel——解调后的高频无线信道噪声；

(11) Pink Noise——有色噪声；

(12) Vehicle interior——Volvo 轿车雨天以 120 km/h 在沥青路面行驶时的车内噪声；

(13) Tank noise——M109 坦克以 30 km/h 速度行驶时的车内噪声。

图 3 反映了抗噪声攻击性能，对于不同类型的噪声，当信噪比在 22 dB 左右时均可以 100% 恢复出水印，也即提取隐藏信息后经过极低速率语音解码后可以完全恢复出保密语音，因而本文算法具有一定的抗噪声攻击性能。

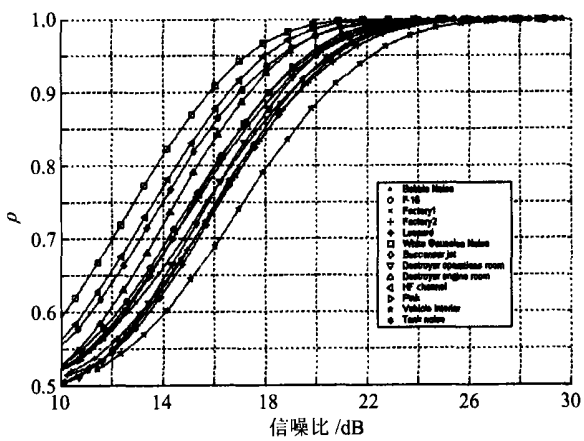

图 3 小波域量化语音伪装方案的抗噪声性能

Fig. 3 anti-noise performance of speech camouflage scheme

需指出的是，抗噪性能与量化阶距直接相关，若量化阶距较小，则算法对噪声攻击比较敏感，而随着量化阶距的不断增大，算法的鲁棒性将逐渐提

高，但是中间语音的音质也随之下降。因此需要在鲁棒性与透明性之间作一折衷。

在无攻击情况下，2 种算法的归一化相关系数 $\rho=1$，但自适应量化阶距算法的中间信噪比为 39.5246，较固定阶距算法提高了 7.5 dB 左右，中间语音的音质有了较大幅度的提高。在信道中有低通滤波、中值滤波以及压缩攻击的情况下，自适应量化阶距算法优于固定阶距算法（表 1—3）。

表 1　低通滤波攻击时的算法性能比较

Tab.1　Performance comparison attacked by LPF

低通滤波/Hz	固定阶距	自适应阶距
≤3500	1	1
≤3000	0.9954	1
≤2500	0.9630	0.9713

表 2　中值滤波攻击时的算法性能比较

Tab.2　Performance comparison attacked by MF

中值滤波	固定阶距	自适应阶距
$N=3$	0.9348	0.9509
$N=5$	0.7681	0.7942
$N=7$	0.7453	0.7662

表 3　压缩编码攻击时的算法性能比较

Tab.3　Performance comparison attacked by speech coding

编码算法	固定阶距	自适应阶距
40 kb/s ADPCM	0.9927	0.9953
32 kb/s ADPCM	0.9745	0.9893
24 kb/s ADPCM	0.9076	0.9435
13 kb/s GSM	0.6224	0.6934

5　结　论

本文提出了在小波域自适应的语音伪装算法，采用混沌序列控制嵌入位置，用嵌入信息控制小波系数的量化，并自适应控制秘密信息的嵌入强度，能实现在接收端不需要原始载体就能提取出嵌入的极低速率编码形成的保密语音，并且获得了比固定嵌入强度方法更好的性能。此算法较好地解决了以语音作载体时存在的信息隐藏难题，为实时实现语音伪装提供了可行方案。

参考文献：

[1]　杨义先，纽心忻，任金强. 信息安全新技术[M]. 北京：北京邮电大学出版社，2002.

[2]　COX I J, KILIAN J, LEIGHTON T. Secure spread spectrum watermark for multimedia[J]. IEEE Trans Image Processing, 1997, 6(12): 1673—1687.

An Adaptive Quantization Speech Camouflage Algorithm Based on F-norm in Wavelet Domain

CHEN Liang[1], *LIU Shi*[2], *LI Yun jian*[1]

(1. Institute of Communications Engineering, PLA University of Sciences & Technology, Nanjing 210007, China;
2. School of Information Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: This paper presents a speech camouflage algorithm based on adaptive quantization in wavelet domain. First speech is segmented to frames and performed discrete wavelet transform. Then secret information is hidden by constant step quantization and the quantized wavelet coefficients are determined by the chaotic sequence. Furthermore, the speech camouflage algorithm is performed based on frobenius norm which adaptively controls the quantization steps. Simulations show that the algorithm is secure, imperceptible. Despite of distortions, the secret information can be extracted without the original speech.

Key words: frobenius norm; wavelet; adaptive quantization