

数字平均法的误差分析

苏 勇, 沈绍宾, 苏 航, 林维正

(同济大学声学研究所, 上海 200092)

摘 要: 文章针对以前提出的声时测量中的随机数字平均技术, 给出了其理论依据, 详细讨论了误差的大小, 并利用概率统计理论进行了论证。同时, 文中还提出均匀延时数字平均技术。在实际应用中, 对该类仪器的测量误差的估计和测量次数的选取以及这类仪器的设计有着重要的指导作用。

关键词: 数字平均; 时间间隔平均; 概率统计; 误差分析

中图分类号: TB52

文献标识码: A

Error analysis of digital averaging

SU Yong, SHEN Shao-bin, SU Hang, LIN Wei-zheng

(Institute of Acoustics, Tongji University, Shanghai 200092, China)

Abstract: This paper presents a theoretical basis of the random digital averaging technique in time interval measurement, and discusses the magnitude of errors based on the statistical theory. A uniform delay digital averaging technique is proposed, which has practical significance in estimating errors, determining number of measurements required, and designing relevant instruments.

Key words: digital averaging; time-interval averaging; statistics; error analysis

1 引 言

Lacy 和 Daniel^[1]提出了提高声时测量精度的数字平均法(digital averaging method)。Aindow 和 Chivers^[2]对时间间隔平均法(time-interval averaging method)与脉冲回鸣法(sing around method)进行了比较, 并讨论了其精度。在国内, 同济大学声学所朱士明等人^[3]首次将其用到了实际测量中, 并提出了过零检测数字平均法, 给出了声时测量平均值的标准差为:

$$STD(\bar{t}) = \frac{T}{\sqrt{6N}}$$

式中 T 为时标脉冲的周期, N 为测量次数, 惠普公司^[4]研制的一些时间脉冲计数器也采用了时间间隔平均的技术, 其给出的声时测量标准差为:

$$STD(\bar{t}) = \frac{T}{\sqrt{N}}$$

式中 T 为时标脉冲的周期, N 为测量次数, 他们都没有对误差在概率以及统计规律的基础上进行详细的讨论。本文从概率和统计理论出发, 对声时测量的误差进行了详尽的推导, 得出了与前者都不同的误差结论。

2 原 理

假设信号 t 稳定不变, 对于其进行单次量化测量, 可以得到量化结果 t_0 。测量的误差与测量设备的量化步长同一量级, 如图 1 所示。

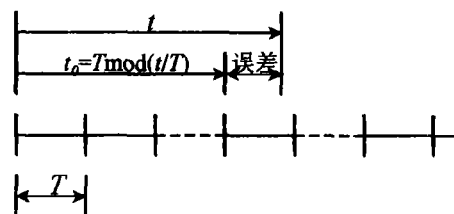


图 1 单次量化测量

对于 t 的测量进行多次并不能减小其测量误差。如果在待测量 t 加上一个单边变量 t_s , t_s 以 T' (T' 为 t_s 的递增步长, $T' \ll T$) 为步长从 0 到 Δt 变化, 则对于 $t + t_s$ 进行 $\frac{\Delta t}{T'}$ 次量化得 t_k , $k = 0, 1, \dots, n-1$, 共 n (n 为量化次数, 量化步长为 T') 个量化结果, 相应量化值 t_i , $i = 0, 1, \dots, N-1$ (N 为可能得到的不同量化值的个数), t_i 的分布为 n_i , $i = 0, 1, \dots, N-1$, n_i 为得到测量值为 t_i 的次数, 则满足关系式 $\sum_{i=0}^{N-1} n_i = n$ 。如图 2 所示, 则可求得 t 的测量值为:

收稿日期: 2001-11-20; 修回日期: 2002-07-31

作者简介: 苏勇(1971-), 男, 江苏人, 博士研究生, 研究方向: 超声检测。

声学技术

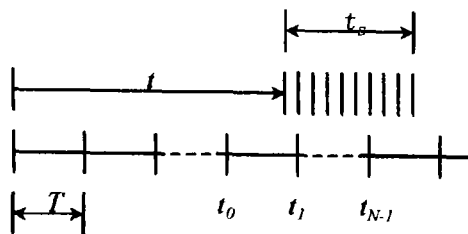


图2 单边多次均匀量化测量

$$t = t_0 + T - n_0 T' = t_0 + T - n_0 \frac{\Delta t}{n} \quad (1)$$

其偏差的量级与 T' 相同。

当 $\Delta t = T$, 即为正单边, 宽度为 T 时, 有:

$$\begin{aligned} t &= t_0 + T - \frac{n_0}{n} T' \\ &= t_0 + T - \frac{n_0}{n_0 + n_1} T \\ &= t_0 + \frac{n_1}{n_0 + n_1} (t_1 - t_0) \\ t &= \frac{t_0 n_0 + t_1 n_1}{n_0 + n_1} \end{aligned} \quad (2)$$

式(2)即为待测量 t 不变时的平均公式。

因此只需要知道 t_0, n_0, t_1, n_1 以及 T , 就可以求出 t , 当 $T' \rightarrow 0$ 时, t 的量化值就越接近真实值。

$$\text{同理, 当 } \Delta t = -T \text{ 时, } t = \frac{t_0 n_0 + t_1 n_1}{n_0 + n_1} + T$$

在实际测量中, 可以在 t 上加一个均匀分布的白噪音 w , 每次量化测量时, w 在 $0 \sim T$ 之间随机变化。当测量次数 n 足够大时, 即可以认为 $t + t_s$ 以 T 的幅度由 t 变化到 $t + T$ 。

下面对测量误差进行统计分析: 由式(2)可知, t 的测量误差主要来自 n_0/n 和 n_1/n , 可用以下概率模型 $(0, 1)$ 表示:

设随机变量 x 为 $(0, 1)$ 分布, 其分布的概率为 $P\{x=1\} = p, P\{x=0\} = q, p+q=1$, 其数学期望为: $E(x) = 1 \cdot p + 0 \cdot q = p$, 而 $E(x^2) = 1^2 \cdot p + 0^2 \cdot q = p$ 。

方差 $D(x) = E(x^2) - [E(x)]^2 = p - p^2 = pq$, 均方差 $\sigma(x) = \sqrt{pq}$

由概率论中心极限分布定律可知: 随机变量 $X_0, X_1, \dots, X_{M-1}, \dots$ 相互独立, 服从同一分布, 因而有相同的数学期望与方差: $E(x_k) = \mu, D(x_k) = \sigma^2 \neq 0, (k=0, 1, \dots, M-1)$, $(\sum_{k=0}^{M-1} x_k)/M$ 值当 M 很大时, 近似服从正态分布 $(\mu, \frac{\sigma}{\sqrt{M}})$ 。

将本问题放入讨论, 则 $\mu = p$ 的无偏估计为

$$(\sum_{k=0}^{M-1} x_k/M) = \frac{n_0}{M}, \text{ 由 } \frac{n_0}{\sqrt{M}} = \frac{\sqrt{pq}}{\sqrt{M}} \text{ 得无偏估计为样}$$

$$\text{本标准差 } s = \sqrt{\frac{1}{M-1} \sum_{k=0}^{M-1} (x_k - \mu)^2} \approx \sqrt{\frac{n}{M} \left(1 - \frac{n}{M}\right)^2},$$

$$\text{其中均方差 } \frac{\sigma}{\sqrt{M}} = \frac{\sqrt{pq}}{\sqrt{M}} = \frac{\sqrt{p(1-p)}}{\sqrt{M}} \leq \frac{1}{2\sqrt{M}}$$

因而, 在本问题中均方差应为:

$$\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{pq}T}{\sqrt{n}} = \frac{\sqrt{p(1-p)}}{\sqrt{n}} T \leq \frac{1}{2\sqrt{n}} T$$

式中 n 为测量次数, p 为待测量量化后的余量占量化精度的比例, 即 $\left[t - T \bmod \left(\frac{t}{T}\right)\right]/T$ 。

当待测量为量化精度整数倍加量化精度的一半时, 均方差最大, 为 $\frac{1}{2\sqrt{n}} T$, 而当待测量是量化精度的整数倍时, 均方差最小, 为 0。

在很多情况下, 由于环境的影响, 以及自身的变化, 待测量不可避免发生变化。因而必须对其变化范围以及影响做出估计, 如下所述:

设相互独立的待测量 t 的分布函数为 $f_t(t)$, 白噪音 w 的分布函数为 $f_w(w)$, 由概率统计理论可知随机变量 $z = t + w$ 的分布函数为:

$$f(z) = f_t(t) * f_w(w) = \int_{-\infty}^{\infty} f_t(z-w) f_w(w) dw$$

但事实上我们很难知道待测量 t 的分布函数 $f_t(t)$ 。故不能再公式(1)对待测量 t 作出估计。然而, 我们注意到当 t 为定值、噪音为单边 $(0 \sim T)$ 时, 得到公式(2)。

以 t 的概率分布函数 $f_t(t)$ 作随机试验, 进行 N 次多次平均, 每次测量 $n_i = n_{i0} + n_{i1}$ 次, 将得到一系列结果 $t_i = \frac{n_{i0}t_{i0} + n_{i1}t_{i1}}{n_{i0} + n_{i1}}, i=0, 1, \dots, N-1$, 可得数学期望值:

$$\begin{aligned} E(t) &= \frac{\sum_{i=0}^{N-1} t_i (n_{i0} + n_{i1})}{\sum_{i=0}^{N-1} (n_{i0} + n_{i1})} = \frac{\sum_{i=0}^{N-1} (n_{i0}t_{i0} + n_{i1}t_{i1})}{\sum_{i=0}^{N-1} (n_{i0} + n_{i1})} \\ &= \frac{\sum_{i=0}^{N-1} n_i t_i}{\sum_{i=0}^{N-1} n_i} \end{aligned}$$

上式即为待测量 t 变化时的平均公式。

上述过程可证明如下:

$$\begin{aligned} E(z) &= \int_{-\infty}^{\infty} z [f_t(t) * f_w(w)] dz \\ &= \int_{-\infty}^{\infty} z \left[\int_{-\infty}^{\infty} f_t(t) f_w(z-t) dt \right] dz \end{aligned}$$

$$= \int_{-\infty}^{\infty} f_t(t) \left[\int_{-\infty}^{\infty} z f_w(z-t) dz \right] dx$$

$$= \int_{-\infty}^{\infty} t f_t(t) dx = E(t)$$

$f_w(w)$ 为噪声 w 所对应的 $(0,1)$ 分布的概率密度函数。

由上式可知,多次平均的结果为变量 t 的数学期望, $z = t + w$ 的单次测量方差为:

$$D(z) = D(t) + D(w)$$

$$= D(t) + T^2 \int_{t-t_n^-}^{t-t_n^+} p(1-p) f_t(p) dp$$

$$\leq D(t) + \frac{T^2}{4}$$

(t_n^-, t_n^+) 为 t 的噪声变化范围(如 t 的变化超出 1 个量化宽度,需要分段积分), n 次测量所得均方差为:

$$\sigma(z) \leq \frac{1}{\sqrt{n}} \sqrt{D(t) + \frac{T^2}{4}} \leq \frac{1}{\sqrt{n}} \sqrt{D(t)} + \frac{T}{2\sqrt{n}}$$

对于声时测量的理想系统,如果触发时刻相对于计时脉冲完全随机,则结束计时时刻等同于声时 t 加上一个 $(0 \sim T)$ 的白噪声,其分析如上所述。

3 计算机仿真与分析

选取两组结果:

待测量:505;待测量噪声:3;量化精度:10;测量次数 $M \times N$:表 1 为 10000×100 ,表 2 为 6400×100 。

由表中可见待测值与测量值的偏差极小。所得均方差为重复多次测量的均方差,与理论值很接近,其中的误差是由于信号本身的噪声所引起的。待测

表 1 次数为 10000×100 的仿真结果

结果值	误差百分比/%	理论方差	实际方差
505.05	0.0091	0.05	0.049
504.91	0.017	0.05	0.050
504.99	0.00099	0.05	0.048
504.99	0.0014	0.05	0.052
504.97	0.0059	0.05	0.055
504.99	0.0022	0.05	0.054
504.95	0.0095	0.05	0.053
504.96	0.0087	0.05	0.056
504.96	0.0079	0.05	0.048
504.99	0.0012	0.05	0.054

表 2 次数为 6400×100 的仿真结果

结果值	误差百分比/%	理论方差	实际方差
505.93	0.013	0.065	0.062
505.06	0.011	0.065	0.061
505.05	0.011	0.065	0.064
504.84	0.031	0.065	0.064
504.93	0.017	0.065	0.065
504.91	0.011	0.065	0.062
504.94	0.016	0.065	0.065
504.89	0.020	0.065	0.063
505.10	0.020	0.065	0.072
505.03	0.0053	0.065	0.058

量噪声对测量结果影响并不显著。

在实际应用中,待测量自身的噪声分布可能不是均匀分布的白噪声,比如可能为高斯分布的噪声,也可以用以上的形式计算。对于声时测量系统,加入的随机噪声可以通过随机触发得到,即相对于计时脉冲,随机触发信号,就可认为加入的噪声是随机噪声。

4 结 论

(1) 从理论上,我们给出了误差最大为 $\frac{T}{2\sqrt{n}} + \frac{\sqrt{D(t)}}{\sqrt{n}}$, $D(t)$ 为待测信号的方差,如在很短时间内, $D(t)$ 可以忽略,量化误差最大为 $\frac{T}{2\sqrt{n}}$ 。

(2) 当 n 很大时,测量时间会延长,则待测量的变化不可忽略,随着平均次数增大所带来的精度会被 t 的变化所抵消,所以盲目增大测量次数是不明智的^[1~4]。

(3) 实际测量中触发时刻的精度对于计时精度非常重要,对于谐波发射,可以用过零检测技术^[1~4]。

参考文献:

- [1] Lewis L Lacy, Daniel A C. Measurements of ultrasonic velocities using a digital averaging technique[J]. J. Acoust. Soc. Am., 1972, 52(1):189-195.
- [2] Aindow J D, Chivers R C. The optimization of electronic precision in ultrasonic velocity measurements: A comparison of the time interval averaging and sing around methods [J]. J. Acoust. Soc. Am., 1983, 73(5):1833-1837.
- [3] 朱士明. 提高声时测量精度的“过零检测数字平均法”[J]. 声学技术, 1990, 9(3):36-40.
- [4] Hewlett-Packard Company. Fundamentals of time interval measurements, Application Note 200-3[Z], 1997.