

多普勒计程仪的数据平滑方法

邹 洪, 向大威, 景永刚

(中国科学院声学研究所东海研究站, 上海 200032)

摘要: 多普勒计程仪输出数据中有局外数据,可能会导致组合导航系统中的卡尔曼滤波器发散。文章介绍了 4 种数据平滑方法:动窗平均、动窗加权平均、动窗中值滤波以及动窗稳健平均。仿真结果表明:动窗稳健平均的输出不会被多个局外数据所影响,并且其处理增益接近理论极限。

关键词: 多普勒计程仪(DVL);组合导航系统;局外数据;动窗平均;动窗加权平均;动窗中值滤波;动窗稳健平均
中图分类号: TB556 **文献标识码:** A **文章编号:** 1000-3630(2008)-04-0507-04

Data smoothing methods for DVL

ZOU Hong, XIANG Da-wei, JING Yong-gang

(Shanghai Acoustic Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Shanghai 200032, China)

Abstract: The outliers contained in the DVL output data might cause the divergence of the Kalman filter of the integrated navigation system. Four data smoothing methods (moving window averaging, moving window weighted averaging, moving window median filtering and moving window robust averaging) are described in this paper. Simulation results indicate that the output of the moving window robust averaging is not influenced by a small fraction of outliers and the processing gain is approximately equal to the theoretical limit.

Key words: Doppler velocity log(DVL); integrated navigation system; outlier; moving window averaging; moving window weighted averaging; moving window median filtering; moving window robust averaging

1 引 言

由多普勒计程仪、惯性导航仪和卡尔曼滤波器所组成的组合导航系统是一种比较理想的自主导航系统^[1]。多普勒计程仪的输出不会漂移,但存在着一定的随机误差。惯性导航仪输出中的随机误差很小,但存在着一定的漂移。组合导航系统,依靠卡尔曼滤波器,尽可能多地利用多普勒计程仪和惯性导航仪的优点而摒弃它们的缺点,从而使系统的输出漂移和随机误差都变得非常小。

在正常情况下,多普勒计程仪的输出等于载体速度与随机误差之和。但是在有些情况下多普勒计程仪的输出中会出现局外(outlier)数据。所谓局外数据就是个别的数据与大多数数据的特性不一致,可能是偏离载体速度,也可能是与大多数数据的随机误差统计特性不一

致。例如,在正常情况下相对于水层进行测速时,多普勒计程仪的输出等于载体相对于水层的速度与随机误差之和。当水层中有运动着的鱼或海洋生物时,由于鱼或海洋生物的回波比体积散射的回波要大,因此多普勒计程仪的输出就等于载体相对于鱼或海洋生物的速度与随机误差之和。只要鱼或海洋生物相对于海底的速度与水层相对于海底的速度不一致,就会产生局外数据。局外数据有可能使卡尔曼滤波器产生发散现象。

因此,多普勒计程仪输出数据的平滑方法,不仅应当有一定的处理增益,同时必须能很好地抑制局外数据。下文将介绍四种平滑方法以及它们的性能。这四种平滑方法就是:动窗平均、动窗加权平均、动窗中值滤波和动窗稳健平均。

2 输入数据与输出数据

设有 N 个输入数据,即:

$$v_k = s + e_k, \quad k=1, 2, 3, \dots, N \quad (1)$$

收稿日期:2008-01-30;修回日期:2008-03-21

作者简介:邹洪(1962-),男,江西南昌人,高级工程师,研究方向为声纳工程、导航。

通讯作者:景永刚, E-mail: ygjing@sina.com

式中 s 是载体的速度 (相对于海底的速度或者相对于水层的速度)。由于载体的惯性, 载体速度在短时间内可以视为不变的, 因此我们把它看作一个与 k 无关的量。式(1)中的 e_k 是随机误差。随机误差一般是高斯型的, 且均值 μ_e 为 0, 而标准偏差为 σ_e 。

若输入数据中的第 m 个数据是局外数据, 则该局外数据表示为:

$$v_m = s + d_m + e_m \quad (2)$$

式中 d_m 不等于 0, 是一个相对于 s 的偏移量。

我们要求平滑处理的输出与输入能一一对应。也就是说, 有一个输入 v_k 就应有一个输出:

$$z_k, \quad k=1, 2, 3, \dots, N$$

3 动窗平均的窗宽

若动窗平均窗宽的设定值为 Q , 当数据序号 k 小于窗宽的设定值时, 数据个数就不能满足设定窗宽的要求, 这样就有必要定义一个实际窗宽 q , 且:

$$\begin{cases} q=k, & k < Q \\ q=Q, & k \geq Q \end{cases} \quad (3)$$

4 帧数据与帧矢量

将第 k 个数据窗内的数据称作第 k 帧数据:

$$v_{k-q+1}, v_{k-q+2}, \dots, v_k \quad (4)$$

利用这 q 个数据组成的横矢量称之为第 k 个帧矢量:

$$\mathbf{v}_k = [v_{k-q+1}, v_{k-q+2}, \dots, v_k] \quad (5)$$

5 均值与中值

为了下文的书写方便, 我们先作以下的定义。

第 k 个帧矢量 $\mathbf{v}_k$ 的均值:

$$\text{mean}(\mathbf{v}_k) = \frac{1}{q} \sum_{j=k-q+1}^k v_j \quad (6)$$

为了定义第 k 个帧矢量 $\mathbf{v}_k$ 的中值, 先定义一个新矢量 $\mathbf{t}_k$, 即:

$$\mathbf{t}_k = [t_k(1), t_k(2), t_k(3), \dots, t_k(q)] \quad (7)$$

并且

$$\mathbf{t}_k = \text{sort}(\mathbf{v}_k) \quad (8)$$

式(8)表明矢量 $\mathbf{t}_k$ 的 q 个元素就是矢量 $\mathbf{v}_k$ 的 q 个元素按升序排列的结果。于是, 第 k 个帧矢量 $\mathbf{v}_k$ 的中值:

$$\text{median}(\mathbf{v}_k) = \begin{cases} t_k(\frac{q+1}{2}), & q=\text{奇数} \\ \frac{t_k(\frac{q}{2}) + t_k(\frac{q}{2}+1)}{2}, & q=\text{偶数} \end{cases} \quad (9)$$

6 动窗平均

动窗平均的算法很简单, 即:

$$z_k = \text{mean}(\mathbf{v}_k), \quad k=1, 2, 3, \dots, N \quad (10)$$

7 动窗加权平均

为了得到动窗加权平均的权系数, 首先计算残差矢量 $\mathbf{r}_k$, 且:

$$\mathbf{r}_k = \mathbf{v}_k - \text{mean}(\mathbf{v}_k) \quad (11)$$

然后计算残差矢量的标准偏差 σ_k , 且:

$$\sigma_k = 1.3 \text{mean}(|\mathbf{r}_k|) + 10^{-12} \quad (12)$$

由于随机误差是高斯的, 因此残差也应该是高斯的。在高斯假设下, σ_k 与 $\text{mean}(|\mathbf{r}_k|)$ 相差大约 1.3 倍^[2], 所以在式(12)中出现了一个比例系数 1.3。式(12)中的常数项 10^{-12} 是为了防止出现 $\sigma_k=0$ 而引入的。例如, 当 $k=1$ 时, 也就是帧矢量只有一个元素时, 就会出现 $v_1 = \text{mean}(\mathbf{v}_1)$, 由式(11)可得 $r_1=0$, 如果式(12)中没有常数项, 就会使 $\sigma_1=0$ 。

进而定义一个门限值 th_k , 即:

$$th_k = 4.6\sigma_k \quad (13)$$

再定义一个矢量 $\mathbf{u}_k$, 即:

$$\mathbf{u}_k = [u_k(1), u_k(2), u_k(3), \dots, u_k(q)] \quad (14)$$

并且:

$$u_k(j) = \begin{cases} \frac{|\mathbf{r}_k(j)|}{th_k}, & |\mathbf{r}_k(j)| \leq th_k \\ 1, & |\mathbf{r}_k(j)| > th_k \end{cases}, \quad j=1, 2, \dots, q \quad (15)$$

有了这个矢量, 就可以很方便的求出权系数矢量:

$$\mathbf{w}_k = [w_k(1), w_k(2), w_k(3), \dots, w_k(q)] \quad (16)$$

当我们采用双平方加权时, 则有:

$$w_k(j) = (1 - u_k(j)^2)^2, \quad j=1, 2, 3, \dots, q \quad (17)$$

由式(15)显见, 残差的绝对值 $|\mathbf{r}_k(j)|$ 越小, 则 $u_k(j)$ 也就越小, 再根据式(17), $w_k(j)$ 就越接近于 1。反之, 若残差的绝对值 $|\mathbf{r}_k(j)|$ 越大, 则 $u_k(j)$ 也就越大, 于是 $w_k(j)$ 就越小。当残差的绝对值 $|\mathbf{r}_k(j)| \geq th_k$ 时, 也就是 $|\mathbf{r}_k(j)| \geq 4.6\sigma_k$ 时, $w_k(j)$ 就等于 0。由于 σ_k 是残差的标准偏差, 所以当残差的绝对值 $|\mathbf{r}_k(j)|$ 等于或大于 4.6 倍残差的标准偏差时, 就意味着残差落在 $-4.6\sigma_k$ 到 $4.6\sigma_k$ 范围外。当不存在局外数据时, 残

差落在 $-4.6\sigma_k$ 到 $46\sigma_k$ 范围外的出现概率大约只有 0.000002。双平方加权时, w 与 u 的关系如图 1 所示。

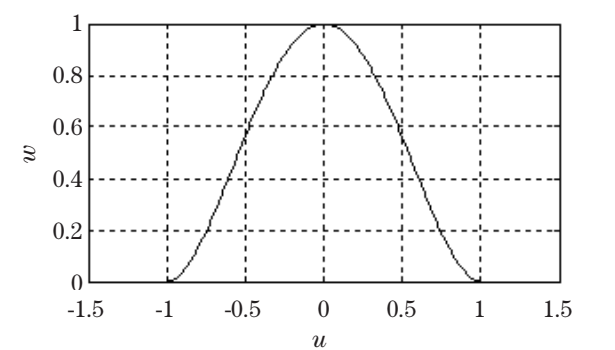


图 1 双平方加权
Fig.1 Bisquare weighting

有了 w_k 以后, 就可以很方便地得到动窗加权平均的输出:

$$z_k = \frac{\mathbf{w}_k \mathbf{v}_k'}{\text{sum}(\mathbf{w}_k)}, \quad k=1, 2, 3, \dots, N \tag{18}$$

式中 $\text{sum}(\mathbf{w}_k)$ 为矢量 $\mathbf{w}_k$ 的诸元素之和。

8 动窗中值滤波

动窗中值滤波的算法也很简单, 即:

$$z_k = \text{median}(\mathbf{v}_k), \quad k=1, 2, 3, \dots, N \tag{19}$$

9 动窗稳健平均

动窗稳健平均的算法和动窗加权平均的方法非常相似, 只是残差矢量 $\mathbf{r}_k$ 的计算方法和残差矢量的标准偏差 σ_k 的计算方法有所不同, 分别为

$$\mathbf{r}_k = \mathbf{v}_k - \text{median}(\mathbf{v}_k) \tag{20}$$

和

$$\sigma_k = 1.3 \text{median}(|\mathbf{r}_k|) + 10^{-12} \tag{21}$$

也就是说, 在计算残差矢量 $\mathbf{r}_k$ 时, 用 $\mathbf{v}_k$ 的中值来代替 $\mathbf{v}_k$ 的平均值, 而在计算残差矢量的标准偏差 σ_k 时, 用 $|\mathbf{r}_k|$ 的中值来代替 $|\mathbf{r}_k|$ 的平均值。一般情况下, 多普勒计程仪输出数据的标准偏差是给定的, 这样就不必再计算式(21), 只要令 σ_k 等于这个标准偏差即可。

动窗稳健平均的其它算法与式(13)、式(14)、式(15)、式(16)、式(17)以及式(18)完全相同。

10 测试数据

为了测量各种平滑方法的处理增益以及抑制局外数据的能力, 可以采用具有下列参数的数据作

为测试数据。

当不存在局外数据时, $N=100000, s=10$, 随机误差是高斯型的, 且 $\mu_e=0$ 和 $\sigma_e=0.1$ 。

当存在单个局外数据时, $N=500, s=10$, 随机误差是高斯型的, 且 $\mu_e=0$ 和 $\sigma_e=0.1$ 。在 $m=150$ 时出现单个局外数据, 而 $d_m=-2$ 。

当存在多个局外数据时, $N=500, s=10$, 随机误差是高斯型的, 且 $\mu_e=0$ 和 $\sigma_e=0.1$ 。在 $m=150 \sim 159$ 时出现出现多个局外数据, 同样有 $d_m=-2$ 。

窗宽的设定值 $Q=25$ 。

11 测试结果

当不存在局外数据时, 动窗平均输出的标准偏差约为 0.0199, 动窗加权平均输出的标准偏差约为 0.0204, 动窗中值滤波输出的标准偏差约为 0.0248, 动窗稳健平均输出的标准偏差约为 0.0209。四种平滑方法有相似的处理能力, 处理增益大约等于

$$G=10\lg(Q)$$

经过平滑处理后得到 14dB 的处理增益, 使得输出的标准偏差等于 1/5 输入的标准偏差。

当存在单个局外数据时, 动窗平均输出、动窗加权平均输出、动窗中值滤波输出和动窗稳健平均输出分别如图 2、图 3、图 4 和图 5 所示。图中的细线代表输入数据, 粗线代表输出数据。

显见, 动窗加权平均、动窗中值滤波和动窗稳健平均都能很好地抑制单个局外数据, 而动窗平均则显得较差, 它在降低偏离幅度的同时却增加了偏离数据的个数。

当存在多个局外数据时, 动窗平均输出、动窗加权平均输出、动窗中值滤波输出和动窗稳健平均输出分别如图 6、图 7、图 8 和图 9 所示。图中的细线代表输入数据, 粗线代表输出数据。

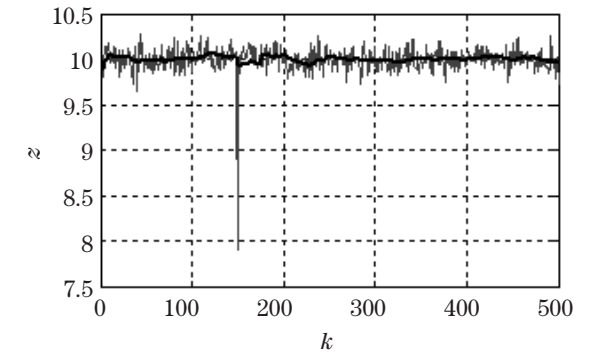


图 2 动窗平均(单个局外数据)
Fig.2 Moving window averaging (single outlier)

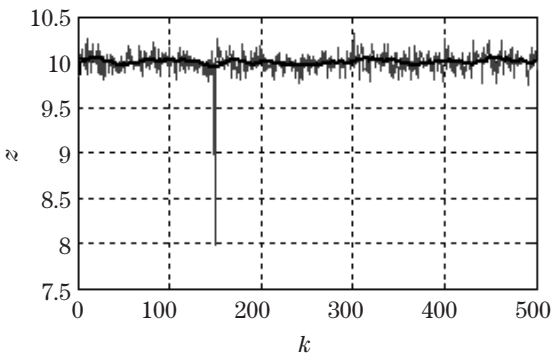


图 3 动窗加权平均(单个局外数据)

Fig.3 Moving window weighted averaging (single outlier)

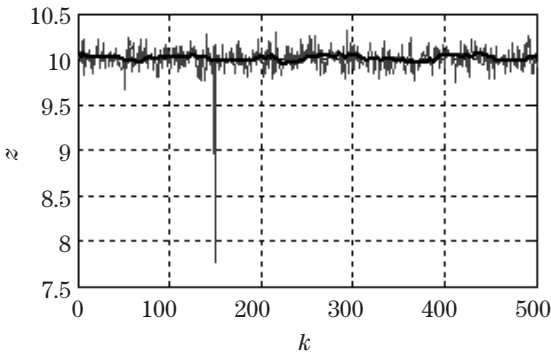


图 4 动窗中值滤波(单个局外数据)

Fig.4 Moving window median filtering (single outlier)

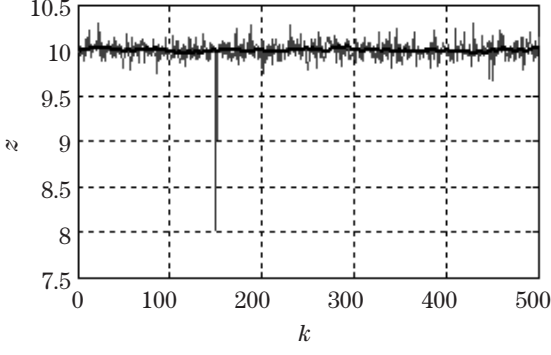


图 5 动窗稳健平均(单个局外数据)

Fig.5 Moving window robust averaging (single outlier)

显见,若按照抑制多个局外数据能力的强到弱来排列,则动窗稳健平均最强,动窗中值滤波次之,动窗加权平均和动窗平均最差。

12 结 论

当处理增益和抑制多个局外数据都很重要时,就应当采用动窗稳健平均,如果计算量受到限制,则可以考虑采用动窗中值滤波。

采用动窗稳健平均的多普勒计程仪,在水库中相对于水层测速时能有效地排除由鱼的运动而引入的局外数据。

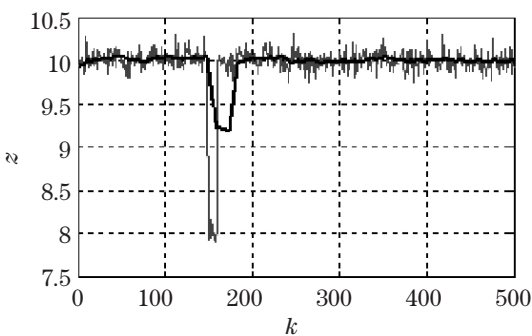


图 6 动窗平均(10个局外数据)

Fig.6 Moving window averaging (10 outliers)

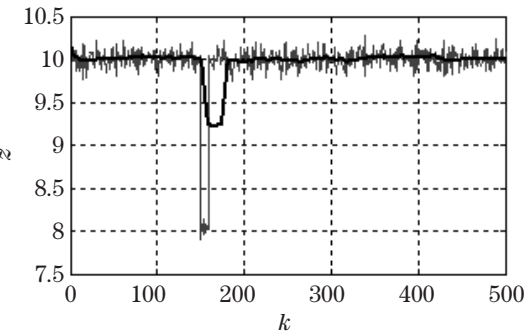


图 7 动窗加权平均(10个局外数据)

Fig.7 Moving window weighted averaging (10 outliers)

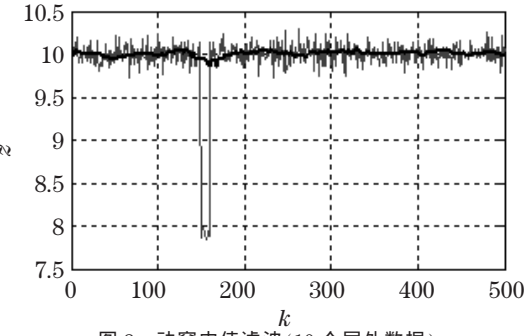


图 8 动窗中值滤波(10个局外数据)

Fig.8 Moving window median filtering (10 outliers)

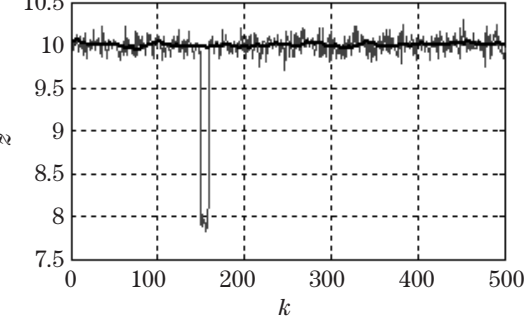


图 9 动窗稳健平均(10个局外数据)

Fig.9 Moving window robust averaging (10 outliers)

参 考 文 献

- [1] David H Titterton, John L Weston. Strapdown inertial navigation technology-2nd edition [M]. The Institution of Electrical Engineers, 2004.
- [2] Sachs L. Applied statistics: A handbook of techniques[M]. Springer-Verlag, 1984.