

脉冲星脉冲轮廓累积的最小熵方法*

张 华[†] 许录平

(西安电子科技大学电子工程学院, 西安 710071)

(2010年8月23日收到; 2010年9月29日收到修改稿)

获取脉冲星辐射信号累积轮廓的传统方法是先识别信号周期, 再按周期对信号进行折叠累加, 这种方法依赖于周期识别的准确性, 对累积轮廓质量也无评价标准. 本文分析了脉冲星累积轮廓和周期的强相关性, 提出了周期未知时直接进行轮廓累积的最小熵方法, 并对其进行证明, 在此基础上, 讨论了该方法在周期识别中的作用. 利用仿真数据和 RXTE 卫星的实测数据实验验证了该方法的有效性.

关键词: 脉冲星, 最小熵, 累积轮廓, 周期识别

PACS: 97.60.Gb, 98.70.Qy, 96.60.tk

1. 引 言

脉冲星是一类高速稳定自旋的中子星, 其辐射信号覆盖了射电、红外、可见光、紫外、X 射线及 γ 射线等波段^[1]. 脉冲星累积脉冲轮廓是通过来自脉冲星的大量形状各异的、具有稳定周期的极弱脉冲信号序列按周期折叠累加平均得到的一类包含丰富细节和特殊结构的波形, 对特定脉冲星的特定频段而言, 它是独一无二的并且极其稳定^[2], 因此它在脉冲星磁层结构^[3]、辐射机理^[4]以及脉冲星定时、深空定位^[5-8]等研究中都起着至关重要的作用. 脉冲星轮廓累积传统方法包括近似模板相关法及其改进方法^[9], 它们的应用前提是精确识别脉冲星周期. 在脉冲星周期识别方面, 常用基于 FFT 的频域周期检测法^[10, 11], 但这些方法仅在高斯白噪声前提下适用, 而频域变换也会导致信号能量分散于谐波中, 当噪声包含频率成分与谐波接近时, 谐波能量得到增强, 这将导致信号及其周期被误检测^[12]. 比如, 在利用 FFT 对 ROSAT 望远镜观测的光变曲线数据进行周期检测时就存在失效情况^[13]. 在周期识别错误或精确情况下, 上述近似模板法将得不到正确的轮廓, 因此这也限制了这种方法的推广应用. 那么能不能摆脱周期限制, 而直接以轮廓特征为准则进行累积呢? 为此, 本文提出了脉冲星轮廓

累积的最小熵方法. 文中, 基于叠加接收脉冲这一最为简单、基本的方式, 先讨论累积轮廓和周期的强相关性, 再给出轮廓累积最小熵方法的具体实现并证明其有效性, 然后讨论了最小熵方法在周期识别中的作用, 最后通过仿真和实测数据实验进行了进一步说明和验证.

2. 脉冲星轮廓累积的最小熵准则

既然累积脉冲轮廓可以通过周期叠加得到, 不难想象, 若周期识别不准确, 累加过程中, 前后两个周期的脉冲信号不能严格对齐, 包含脉冲信号的观测值将被分散于整个周期内, 造成信号区的分辨率下降, 轮廓剖面的波形钝化. 如图 1 所示, 图中以脉冲星 B0329 + 54 为例(仿真条件见 5.1 节), 给出了不同周期误差下的累积轮廓(为了便于区分, 各累积轮廓加上了偏置量), 可见周期无误时可以得到清晰、尖锐的轮廓, 周期误差 $\Delta T/T$ 为 10^{-5} 时轮廓变得很模糊, 周期误差 $\Delta T/T$ 为 5×10^{-5} 时已经看不出明显轮廓, 这说明累积轮廓波形是轮廓累积质量的直接反映且与累积周期具有很强的相关性. 为此可以利用累积脉冲轮廓的锐化度作为对齐的准则, 即对一段观测信号, 依次改变累积周期, 求该信号的累积脉冲轮廓, 轮廓锐化度最大时, 就认为得到了最优轮廓, 此时的周期为真实信号周期, 这个过程

* 国家高技术研究发展计划(863)(批准号:2007AA12Z323)和国家自然科学基金(批准号:60772139)资助的课题.

[†] E-mail: zhanghua@mail.xidian.edu.cn

正好与先识别周期再累积得到轮廓的传统方法相反. 如何评价累积轮廓的锐化程度呢? 可以利用熵的概念.

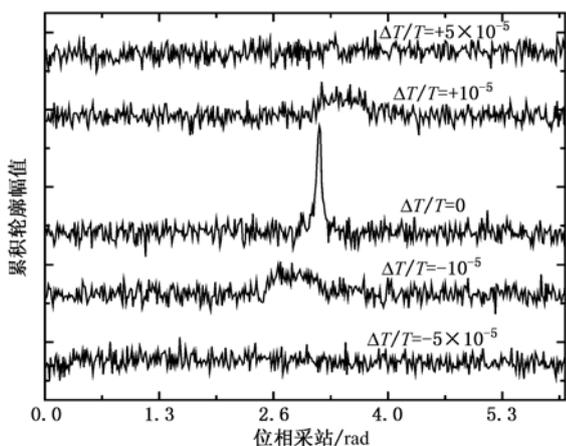


图1 累积轮廓与周期误差的关系示意

以香农熵为分析对象, 它可以表示为概率密度函数的对数的期望. 设随机变量 x 的密度函数为 $p(x)$, 其熵为^[14]

$$H(x) = E\{-\ln p(x)\} = -\int_{-\infty}^{\infty} p(x) \ln p(x), \quad (1)$$

其离散形式为

$$H(x) = -\sum_i p_i \log p_i, \quad (2)$$

其中 $p_i = P\{x = x_i\}$. 熵用于衡量随机变量的不确定性, 从随机变量分布角度看, 也可以认为是对平均分散程度的测量. 以脉冲星累积轮廓为例, 若脉冲信号平均分配于整个周期, 轮廓波形幅度将处处相等, 熵最大; 反之, 波形锐化程度越高, 熵越小. 由此可见, 利用轮廓最小熵作为脉冲星轮廓累积的准则是可行的. 本文将利用最小熵作为脉冲星信号轮廓累积准则的方法称为脉冲星轮廓累积的最小熵方法.

3. 脉冲星轮廓累积的最小熵方法及其证明

第2节结合脉冲星轮廓累积过程简要说明了最小熵方法的可行性, 也就是说利用最小熵方法累积出的轮廓应该是最优的. 本节给出具体分析方法, 并对最小熵方法进行证明.

如图2所示, 为一段 X 射线脉冲星观测信号, 它由 N 个采样构成, 采样间隔为 δt , 起始时刻为 t_0 ,

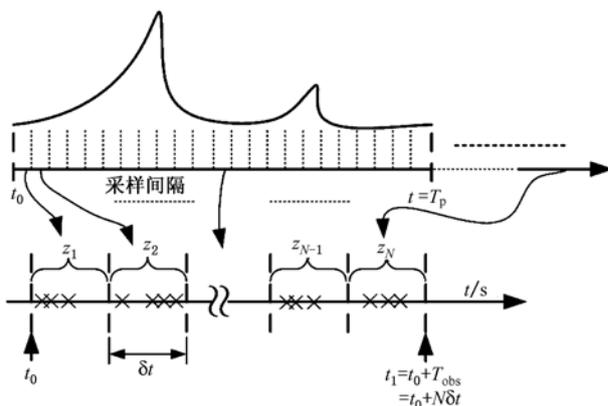


图2 脉冲星信号采样序列示意

周期为 T_p , $N\delta t = mT_p$, m 为观测信号中的周期数, $z_i (i = 1, \dots, N)$ 表示每个采样间隔内光子计数, 写成向量形式为 $\mathbf{z} = [z_1, z_2, \dots, z_N]^T$, 令 $\|\mathbf{z}\| = \sum_{i=1}^N |z_i|$. 为了区别观测信号, 用 $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$ 表示累积后的信号, 令 $p_i = |x_i| / \|\mathbf{z}\|$, 根据(2)式, 将 \mathbf{x} 的熵定义为

$$H(\mathbf{x}) = -\sum_{i=1}^m p_i \log p_i. \quad (3)$$

若将脉冲星标准累积轮廓分为辐射区和非辐射区, 在通过周期折叠方法进行累加过程中, 当每个周期都严格对齐时, 信号被集中在辐射区, $H(\mathbf{x})$ 较小; 反之, 当周期出现偏差时, 前后两个周期将出现错位, 前一次周期折叠过程还会不断影响后续的周期对齐, 此时辐射区的部分信号会随着累积过程分布到非辐射区, 使 \mathbf{x} 的各分量趋于相等, $H(\mathbf{x})$ 增大. 这一过程等价于证明 $H(\mathbf{x}) < H_d(\mathbf{x})$, $H_d(\mathbf{x})$ 表示周期未严格对齐的 \mathbf{x} 的熵. 不失一般性, 令 $x_j (j \in [1, m])$ 为信号 \mathbf{x} 中辐射区分量, 它是 m 个周期的观测信号 \mathbf{z} 的辐射区分量 y_k 叠加的结果, 即 $x_j = \sum_{k=1}^m y_k$. 从 x_j 中任取 $y_a, a \in [1, m]$, 假设 y_a 在累加过程中被分配到非辐射区 $x_w, x_w \in [1, m]$, 且

$$x_w < x_w + y_a \leq x_j - y_a < x_j,$$

令 $\phi(\alpha) = -\alpha / \|\mathbf{z}\| \log(\alpha / \|\mathbf{z}\|)$, 先证明

$$\phi(x_j) + \phi(x_w) < \phi(x_j - y_a) + \phi(x_w + y_a). \quad (4)$$

根据拉格朗日中值定理, 存在 $\xi_1 \in (x_w, x_w + y_a)$ 和 $\xi_2 \in (x_j - y_a, x_j)$, 使

$$\begin{aligned} & \phi(x_w + y_a) - \phi(x_w) \\ &= \phi'(\xi_1)(x_w + y_a - x_w) \end{aligned}$$

$$= \phi'(\xi_1) y_a \quad (5)$$

和

$$\begin{aligned} & \phi(x_j) - \phi(x_j - y_a) \\ &= \phi'(\xi_2)(x_j - (x_j - y_a)) \\ &= \phi'(\xi_2) y_a. \end{aligned} \quad (6)$$

将(5)式和(6)式代入(4)式,简化后,欲证(4)式,则只需证明

$$\phi'(\xi_2) < \phi'(\xi_1), \quad (7)$$

因为 $\phi'(\alpha) = -1/\|z\| \log(\alpha/\|z\|) - 1$, 欲证(7)式,则只需证明

$$\log\left(\frac{\xi_1}{\|z\|}\right) < \log\left(\frac{\xi_2}{\|z\|}\right). \quad (8)$$

由于 $\xi_1 < \xi_2$, 显然(8)式成立,则(4)式得证;再利用(4)式证明(3)式,设 x_j 中 y_a 错位到 x_w , 此时 \mathbf{x} 的熵表示为 $H_d(\mathbf{x})$, 即

$$\begin{aligned} H_d(\mathbf{x}) = & - \sum_{i \neq j, w; i, j, w \in [1, m]} p_i \log p_i \\ & + \frac{x_j - y_a}{\|z\|} \log\left(\frac{x_j - y_a}{\|z\|}\right) \\ & + \frac{x_w + y_a}{\|z\|} \log\left(\frac{x_w + y_a}{\|z\|}\right). \end{aligned} \quad (9)$$

对(9)式简化得到

$$\begin{aligned} H_d(\mathbf{x}) = & - \sum_{i \neq j, w; i, j, w \in [1, m]} p_i \log p_i \\ & + \phi(x_j - y_a) + \phi(x_w + y_a) \\ > & - \sum_{i \neq j, w; i, j, w \in [1, m]} p_i \log p_i + \phi(x_j) + \phi(x_w) \\ = & H(\mathbf{x}). \end{aligned}$$

这就证明了只有辐射区严格对齐时,脉冲星信号按周期累加才能最小化 $H(\mathbf{x})$, 从而使累积轮廓达到最优. 脉冲星信号也可以按多个周期进行折叠累加,这并不影响用最小熵准则衡量对齐程度,反而能减小周期识别时位相分辨精度引起的误差,这一点在第4节中说明.

4. 利用累积轮廓最小熵确定脉冲星周期

第2,3节中说明了周期不准确时脉冲星累积轮廓波形会钝化,实际上脉冲星轮廓累积时还存在另一个问题:脉冲星信号是离散的位相采样序列,在累积过程中,采样序列只能按照位相采样间隔最小单位的整数倍折叠,如果实际周期不是位相采样间隔最小单位的整数倍,按单周期逐次累积时,前后

两个周期就不能严格对齐而出现误差,随着累积周期数的增加,误差也会逐渐增大. 有两种方法可以减小这种误差:一是逐次累加过程中,修正累积时前后两个周期参考点^[9];二是按多周期折叠累积. 前一种方法可以累积出单周期轮廓,后一种方法信号能量会分散到多周期中,轮廓相对模糊,此时也可以进一步采用第一种方法进行累积得到单周期轮廓. 相比而言,多周期累积更适用于最小熵方法,有些情况下能提高周期识别的精度. 为说明这一点,本节先从理论上解释,然后在实验中进一步分析.

为表述方便,将轮廓累积时参照的长度称为一个累积窗,仅包含一个周期的累积窗称为单周期窗,包含多个周期的累积窗称为多周期窗. 简单起见,忽略周期演化的高阶项和其他天文因素,脉冲星信号累积过程可表示为

$$C_m(w\delta t) = \sum_{n=1}^{[N/m]} h(n(mT_r - w\delta t), (n+1)(mT_r - w\delta t) - 1), \quad (10)$$

式中, T_r 为累积时参考周期, h 表示脉冲星序列, $h(o, p)$ 表示序列中从采样 o 开始到 p 结束的一段样本, N 为 h 序列的总脉冲周期数, m 表示累积窗的周期数,单周期窗累积时 $m = 1$, $[\cdot]$ 表示取整, $w\delta t$ 表示偏差, δt 为位相采样间隔, w 为自然数. 通过改变 $w\delta t$, 利用最小熵确定周期方法为

$$T = T_r - \frac{1}{m} \min_{w\delta t} \text{Ep}[C_m(w\delta t)], \quad (11)$$

式中, $\text{Ep}[\cdot]$ 表示用(3)式求熵. 第3节中已经指出了利用多周期窗轮廓累积的最小熵方法确定脉冲星周期,可以提高周期识别的精度,(11)式说明了这一点. 直观看来,(11)式中,在 $w\delta t$ 误差一定情况下,累积窗长 m 越大,所得真实周期 T 误差越小. 实际上,累积过程中累积窗的误差最大可达 $0.5 \delta t$, 此时周期 T 的最大误差为 $\Delta T = \delta t/(2m)$, 因此利用最小熵方法确定周期时,使用 m 个周期长的累积窗进行累积,其误差应该为使用单周期窗累积时最大误差的 $1/m$. 特别当实际周期不是位相采样间隔最小单位的整数倍时,按一定长的多周期窗累积可以消除小数部分引起的周期识别误差. 比如 $m = 10$ 时,可以完全消除1位小数引起的误差. 但累积窗并不是越长越好,累积窗越长,累积轮廓的熵随累积窗变化时分辨率越低. 实验部分将进一步说明.

5. 实 验

5.1. 仿真数据准备

选取 B0531 - 21, B1937 + 21, B0329 + 54 和 B0950 + 08 四颗平均轮廓具有代表性的脉冲星为例,主要参数如表 1 所示. 为方便对比和分析,不失一般性,以 X 射线脉冲星辐射信号为例,并设 X 射线探测器有效探测面积为 1 m^2 ,将仿真数据中信号辐射强度设置为 $5 \times 10^{-2} \text{ ph/cm}^2 \cdot \text{s}^{-1}$,平均背景辐射强度为 $5 \times 10^{-2} \text{ ph/cm}^2 \cdot \text{s}^{-1}$. X 射线脉冲星信号,可用泊松分布对光子到达事件进行建模. 由于 X 射线脉冲星辐射信号强度呈周期性变化,光子到达探测器的密度随时间分布是不均匀的. 定义落在区间 (t_1, t_2) 内的光子数等于 k 的概率为

$$P(k/(t_1, t_2)) = \frac{\left[\int_{t_1}^{t_2} \lambda(t) dt \right]^k}{k!} \exp \left\{ - \int_{t_1}^{t_2} \lambda(t) dt \right\}, \quad (12)$$

式中, $\lambda(t)$ 表示 t 时刻光子流量密度, $\lambda(t) = \lambda_s(t) + \lambda_n$, 其中 $\lambda_s(t)$ 为信号辐射强度,当脉冲周期为 T 时,有 $\lambda_s(t) = \lambda_s(t + T)$, λ_n 为各种噪声强度的总和,包括背景辐射噪声、系统热噪声等; $t_2 - t_1 = \Delta T$ 为观测时间. 根据脉冲星周期、标准累积轮廓(简单起见,采用 EPN 数据库轮廓^[15])、平均流量和每周期采样次数计算每两次采样之间信号流量 $\delta(t)$ 和背景辐射流量 α , 单周期仿真信号表示为 $b_{\text{sim}}(t) = \sum_{n=1}^S \{ \text{poissrnd}(\delta(t - n\Delta t)) + \alpha \}$, 其中 S 为采样点数, $\text{poissrnd}(\cdot)$ 为(12)式的泊松模型;将单周期仿真信号顺序排列,模拟出观测序列.

表 1 仿真用脉冲星参数

脉冲星	周期/s	周期采样数	采样间隔/s	脉冲宽度/s
B0531 - 21	3.34×10^{-2}	62	5.3×10^{-4}	1.7×10^{-3}
B1937 + 21	1.56×10^{-3}	1024	1.5×10^{-6}	2.1×10^{-5}
B0329 + 54	7.145×10^{-1}	474	1.5×10^{-3}	6.7×10^{-2}
B0950 + 08	2.5306×10^{-1}	472	5.3×10^{-4}	5.8×10^{-2}

5.2. X 射线脉冲星累积轮廓的熵分析

对模拟的观测序列按单周期窗进行折叠累加,其中周期较短的 B0531 - 21 和 B1937 + 21 分别累积 10^4 次和 1.5×10^4 次, B0329 + 54 和 B0950 + 08

均累积 10^3 次,利用(3)式求其熵,熵随累积窗误差变化情况如图 3 所示. 图中横坐标为累积周期窗与真实周期窗在采样点数上的差. 鉴于各个脉冲星信号周期、采样率不同以及累积窗宽只能是采样间隔的整数倍,横坐标统一用采样间隔 δt 作为单位;纵坐标为熵值,为了对比方便,进行了归一化处理. 从图中可以看出,当累积周期窗与真实周期窗有差异时,熵值保持在较高的水平,波动较小;当累积周期窗与真实周期窗相同时,熵值变小,且有明显谷底. 所选四颗星表现一致,这就说明了利用最小熵作为信号累积准则的可行性和有效性.

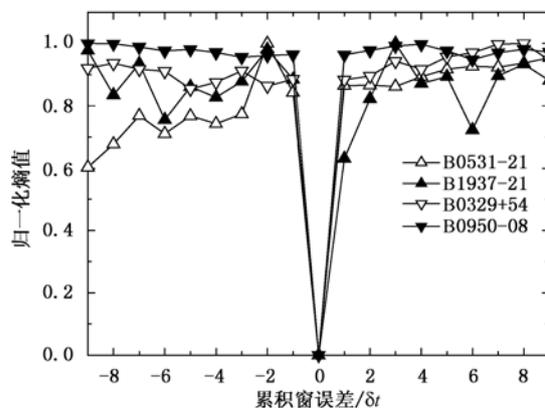


图 3 单周期累积轮廓的熵

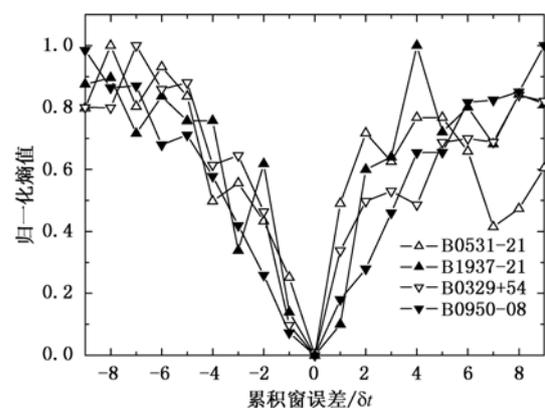


图 4 多周期累积轮廓的熵

利用最小熵分析累积脉冲轮廓,并不限定累积窗周期数,相反,按照多周期窗进行折叠累加,在累积效果上能得到一些额外信息. 为了说明这一点,对模拟的观测序列按多周期窗进行折叠累加,其中周期较短的 B0531 - 21 按窗长 1000 个周期进行折叠累加 10 次, B1937 + 21 按窗长 1500 个周期累积 10 次, B0329 + 54 和 B0950 + 08 均按窗长 100 个周期累积 10 次,利用(3)式求熵,熵随周期变化情况

如图 4 所示. 这里, 将累积窗设计的较长, 折叠次数设计的较少, 是出于两点考虑, 其一, 保证所用的观测序列长度与单周期窗累加时一样, 不引入附加信息; 其二, 较长的折叠周期和较少的累积次数能充分说明按多周期窗折叠累积不影响累积时最小熵方法的应用. 从图 4 中可看出, 当周期窗准确时熵值最小, 说明按多周期窗累积时使用最小熵准则是可行和有效的. 与图 3 中不同的是, 按多周期窗累积时, 随着累积窗长与真实窗长误差的变化, 熵值变化平缓的多, 且有明显的递增或递减趋势. 之所以存在这种不同, 原因在于多周期窗累积时累积次数变少了. 根据第 2 节的分析, 当脉冲区的观测值在累积过程中被错误地分配到非脉冲区时, 会导致熵变大, 这种错误分配越多, 熵值也越大, 而累积窗误差增大或累积次数增多都会导致这种错误分配增多, 只是不同的情况下二者作用程度不一样; 单周期窗累积时, 累加次数起主导作用, 熵值只在正确累积窗下最小, 因而熵值随累积窗长误差变化就不明显了, 如图 3 所示; 反之, 多周期窗累积时, 累积次数较少, 窗误差起主导作用, 此时熵值随累积窗误差变化就存在明显递增或递减趋势, 如图 4 所示. 图 4 中的 V 字形变化趋势的拐点能够作为是否需要进一步改变窗长再累积的判据, 与单周期窗相比, 在最小熵位置附近的分辨能力有所下降, 但其影响并不大.

5.3. 利用累积轮廓最小熵确定脉冲星周期的性能分析

结合第 4 节结论, 通过实验对最小熵方法在周期确定方面的性能进行比较分析. 不失一般性, 仍以脉冲星 B0329 + 54 为例, 设置其采样间隔 $\delta t = 1.5 \times 10^{-3} \text{ s}$, 将周期修改为 $475.5 \times \delta t$. 这里周期不再是采样单位的整数倍, 这种情况下, 经多次实验表明, 按单周期窗累加次数过多, 将得不到明显的最小熵, 不便于对比分析, 为了减小累积次数, 将信号辐射强度设置为 $5 \times 10^{-1} \text{ ph/cm}^2 \cdot \text{s}^{-1}$, 其他仿真数据生成条件及方法与 5.1 节相同. 先生成模拟的观测序列, 然后进行累积, 累积时参考周期设置为 $475 \times \delta t$, 先按单周期窗累积 100 次, 再按窗长为 10 个周期的多周期窗累积 10 次, 分别求熵并归一化, 结果如图 5 所示.

从图 5 中可见, 单周期窗累积时, 有两个熵值较小, 且非常接近, 它们所对应的周期分别为 $475 \times \delta t$

和 $476 \times \delta t$, 而真实周期为 $475.5 \times \delta t$ 时, 理应有这种结果. 若以最小熵为判据, 单周期窗累积方法所得周期应该为 $475 \times \delta t$, 此时误差为 $0.5 \times \delta t$. 与单周期窗累积相比, 多周期窗累积时, 只有一个值明显较小, 利用最小熵为判据所得累积窗与实际所用累积窗的长度差为 $5 \times \delta t$, 由于窗长为 10 个周期, 设真实周期为 T_p , 有 $10 \times T_p = 10 \times 475 \times \delta t + 5 \times \delta t$, 计算得到 $T_p = 475.5 \times \delta t$, 与真实周期相等. 这就说明了利用最小熵方法确定脉冲星周期时利用多周期窗累积的优势.

仍以脉冲星 B0329 + 54 为例, 为进一步分析窗长对周期识别的影响, 将周期修改为 $475.5314157 \times \delta t = 0.7104 \text{ s}$, 利用 4.1 节的仿真条件和方法得到 1000 个周期长的序列, 保持序列长度不变, 利用最小熵方法求周期, 分析周期误差随累积窗长的变化情况, 结果如图 6 所示. 从图 6 中可见, 总体上随着累积窗周期数的增加, 周期识别误差有减小趋势; 累积窗长小于 100 时, 误差随窗长变化浮动较大, 在某些特定窗长位置, 误差接近零, 但并不明显; 当累积窗长大于 100 时, 周期识别误差就有明显减小了, 仔细观察能够发现, 随着累积窗长的增加, 周期识别误差逐渐平滑减小. 当然, 这里只给出了个例, 通过相当多实验我们发现, 实验所得结论是基本类似的, 不同的是误差随累积窗长变化趋势的快慢以及误差出现明显改善的位置不一样, 这和真实周期小数部分的具体值有关系.

作为对比, 对同样的序列利用 FFT 方法求幅度谱, 如图 7 所示. FFT 方法一直是周期信号分析的常用方法, 其不足在于其受谐波和色噪声的影响, 不可避免出现多解情况, 造成真实解判定的混乱. 其中, 谐波来自于轮廓中的高频成分, 如脉冲星脉冲轮廓的多峰、子脉冲等特征, 因此频域谐波即使在没有噪声的情况下也会存在; 色噪声可能来自于探测器的电子读出设备, 它会增强与其同频谐波的能量, 造成误判. 从图 7 中可以看到, 幅度谱中包括了真实频率成分, 但同样包含丰富的谐波成分, 有些谐波的幅度甚至大于真实频率的幅度, 很容易造成误判. 相对而言, 最小熵方法在周期识别中所得结果要清晰得多, 一般都只有唯一解, 图 3、图 4 和图 5 都说明了这一点. 而在周期识别准确性方面, FFT 方法所得周期为 0.7095 s , 误差 0.001 s ; 累积窗长分别为 200 和 500 周期时最小熵方法所得周期分别为 0.7154 s 和 0.7124 s , 误差分别为 0.005 s 和

0.002 s, 可见, 二者周期识别精度都较高, 相比之下, FFT 方法精度略微高一点.

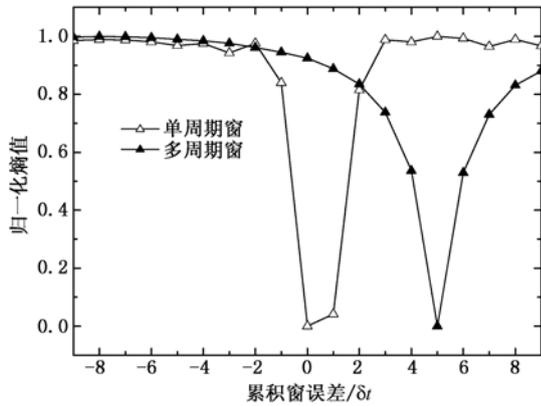


图5 两种累积窗周期确定比较

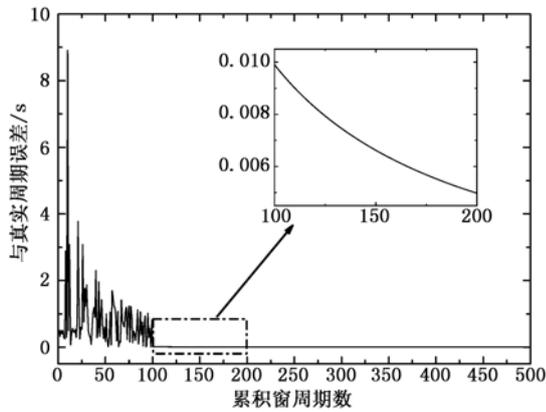


图6 周期识别误差随累积窗长变化情况

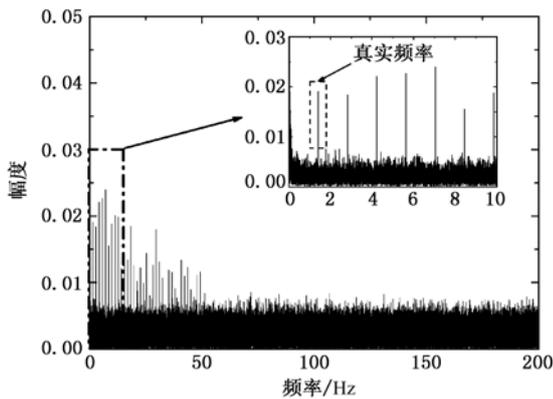


图7 FFT 幅度谱

5.4. RXTE 实测数据实验

从 HEASARC 数据库的 Public Data Access 链接进入, 以 B1855 +09 为关键字搜索出以该脉冲星为观测对象的观测号, 然后从其 FTP^[16] 上下载与该

测号对应的数据, 利用 FTOOLS 工具处理得到该脉冲星 X 射线光变曲线. 经过修正, 剔除掉不良数值, 得到 256 个周期长的信号, 对该信号经周期累积去均值归一化后的轮廓如图 8 所示. 对这 256 个周期长的信号按单周期窗和窗长为 50 个周期的多周期窗累积求熵随窗误差变化情况, 并归一化, 结果如图 9 所示. 观察图 9, 当累积窗正确时能得到明显的最小熵值, 可见利用最小熵方法处理实测数据和仿真数据的结果一致, 这就进一步说明了最小熵方法的有效性. 图 9 中多周期窗累积结果与图 4 的仿真结果基本一致, 所不同的是没有明显的趋势, 这是由于这里所用数据采样率低、周期数少的原因, 与图 4 的结论并不矛盾.

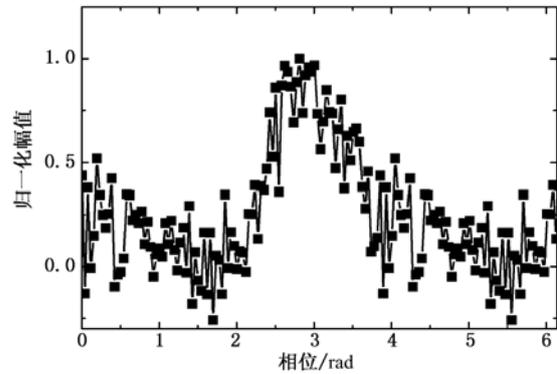


图8 归一化累积轮廓

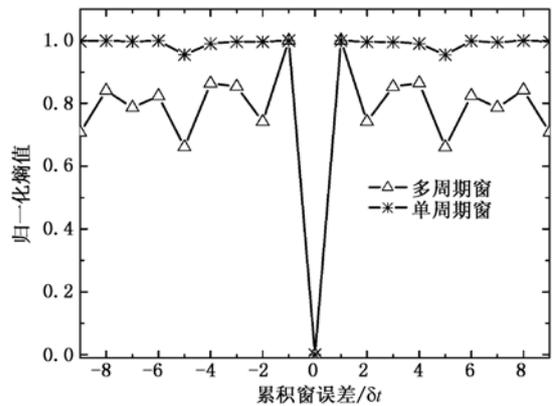


图9 累积轮廓熵随累积窗误差变化情况

6. 结 论

本文所提脉冲星轮廓累积最小熵方法与传统方法实现过程完全不同, 先在最小熵意义下得到累积轮廓, 再提取周期, 这为脉冲星轮廓累积和周期识别提供了新的思路. 对最小熵方法的可行性和有效性, 文中给出理论上的证明, 并利用仿真和实

测数据进行了实验验证;分析了单周期窗和多周期窗在周期识别中不同的作用和意义,其中多周期窗在周期识别精度方面表现更好;作为对比,分析了最小熵方法和 FFT 方法在周期识别中的性能差异。实验结果显示,二者在精度方面相当,都能达到较高精度,但最小熵方法所得结果具有唯一性,避免

了 FFT 方法中经常出现的多解情况。本文在简化条件下已处理了一些仿真例和实例,更详细的分析有待今后继续进行。直接利用轮廓质量作为累积轮廓评价标准的最小熵方法性能较好,只需要加法运算,计算量并不大,是实际意义和推广能力都较好的方法。

- [1] Ji P Y, ZHU J 2008 *Chin. Phys. B* **17** 356
- [2] ESAMDIN A, Wu X J, Zhang J, Wang Na, Kang L S, Jin S Z, Zhang H B, Wang W X 2000 *Acta Astrophysica Sinica* **3** 59 (in Chinese) [伊·艾力、吴鑫基、张晋、王娜、康连生、金声震、张洪波、王维侠 2000 天体物理学报 **3** 259]
- [3] Kramer M 1994 *A&AS* **107** 527
- [4] Xu X B, Wu X J 2002 *Science in China Ser. A* **12** 1134 (in Chinese) [徐轩彬、吴鑫基 2002 中国科学(A辑) **12** 1134]
- [5] Zhong C X, Yang T G 2007 *Acta Phys. Sin.* **56** 6157 (in Chinese) [仲崇霞、杨廷高 2007 物理学报 **56** 6157]
- [6] Taylor J H 1992 *Phil. Trans R. Soc. Lond.* **341** 117
- [7] Xie Z H, Xu L P, Ni G R 2008 *Acta Phys. Sin.* **57** 6683 (in Chinese) [谢振华、许录平、倪广仁 2008 物理学报 **57** 6683]
- [8] Sheikh S I 2005 *Ph. D. Dissertation* (Maryland: University of Maryland)
- [9] Li J X, Ke X Z 2008 *Acta Astronomica Sinica* **4** 394 (in Chinese) [李建勋、柯熙政 2008 天文学报 **4** 394]
- [10] Seward F D, Harnden F R, Helfand D J 1984 *Astrophysical Journal* **287** L19
- [11] Li J X, Ke X Z 2010 *Acta Phys. Sin.* **59** 8304 (in Chinese) [李建勋、柯熙政 2010 物理学报 **59** 8304]
- [12] Zhang H, Xu L P, Xie Q, Luo N 2011 *Acta Phys. Sin.* **50** (in Chinese) (in press) [张华、许录平、谢强、罗楠 2011 物理学报 **50**] (已接受)
- [13] Gregory P C, Loredo T J 1996 *ApJ* **473** 1059
- [14] Papoulis A, Pillai S U (Translated by Bao Z, Feng D Z, Shui P L) 2004 McGraw-Hill (Xi'an Jiaotong University Press) p530 (in Chinese) [帕普里斯. A, 佩莱. S. U 著 保铮, 冯大政, 水鹏朗译 2004 概率、随机变量与随机过程 (西安交通大学出版社) 第 530 页]
- [15] Epn. <http://www.jb.man.ac.uk/research/pulsar/resources/epn/browser.html>
- [16] Rxte. FTP: legacy.gsfc.nasa.gov.

Minimum entropy cumulation method of pulsar profile^{*}

Zhang Hua[†] Xu Lu-Ping

(School of Electronic Engineering, Xidian University, Xi'an 710071, China)

(Received 23 August 2010; revised manuscript received 29 September 2010)

Abstract

To obtain the cumulation profile of pulsar signal, the traditional method needs to identify its period first, and then fold the signal with the period. This method depends on the accuracy of the period, and it has no evaluation standard for the profile. In this paper we analyze the strong correlation between the cumulation profile and the period, and present a minimum entropy method of pulsar cumulation profile and its verification. Therefore this method is applied to the identification of pulsar cycle. The RXTE data and the simulation data are used to verify the effectiveness of the method.

Keywords: pulsars, minimum entropy, cumulation profile, cycle identification

PACS: 97.60.Gb, 98.70.Qy, 96.60.tk

^{*} Project supported by the National High Technology Research and Development Program of China (Grant No. 2007AA12Z323) and the National Natural Science Foundation of China (Grant No. 60772139).

[†] E-mail: zhanghua@mail.xidian.edu.cn