

基于地磁水平分量序列的磁暴自动识别

易世华 李夕海 李志刚

(中国西安 710025 第二炮兵工程学院)

摘要 应用小波包分析技术提取了地磁水平分量序列的能量特征,用 Fisher 方法进行识别,研究了磁暴与非磁暴的识别率在不同时间长度 t 下的变化规律。当 $t=1$ h 时,对磁暴与非磁暴的平均识别率达到 83.8%;在确定为磁暴的前提下,进一步将磁暴按强度分为强磁暴(磁情指数 $K \geq 8$)和非强磁暴($K \leq 7$),平均识别率为 72.9%。在对地磁水平分量进行连续实时监测的情况下,用这种方法可实现对磁暴的自动识别。

关键词 磁暴;小波包分析;特征提取;识别

中图分类号:P318.2⁺2 **文献标识码:**A **文章编号:**1003-3246(2007)01-0037-05

引言

随着人类空间活动的增多以及航天事业的发展,人们越来越重视对空间环境的监测和预报。由于磁暴是空间环境扰动的主要现象之一,当地磁场强烈扰动形成磁暴后,将产生重大灾害性事件,可导致通信卫星、气象卫星等各类航天器的电子仪器故障甚至失去地面联系,同时,磁暴也会对电力系统、通讯系统、输油管道等地面技术系统造成危害(徐文耀,2003)。因此,对磁暴的预报成了空间科学领域一个重要的研究课题。Lundstedt 最早将神经网络用于日地活动预报,提前一天预报 A_p 指数(Lundstedt H,1992);Samuel Kugblenu 等提出了一种带有一个隐含层和误差反向传播学习的神经网络用于磁暴预报,实验表明该网络能以很好的准确性重现恢复相(Samuel Kugblenu et al,1999);刘振兴等提出用等效电路模型来预报亚暴和磁暴的指数(沈超,刘振兴,2000);魏奉思等综合行星际扰动过程的物理模型,提出一种预报由太阳风暴吹向地球而引起的地磁扰动的“ISF”方法(魏奉思等,2003);周晓燕等研究了地磁扰动及其预报的系统辨识建模方法(周晓燕等,1994);研究磁暴预报的神经网络方法还有顾左文(1995)、周晓燕(1998)。纵观上述方法,对磁暴的预报大都以太阳活动的卫星观测数据以及日地系统中各环节的物理模型为基础进行预报,一旦缺少卫星观测数据预报就难以实现,本文则从地磁分量序列出发,运用信号处理和模式识别技术进行磁暴自动识别和强度类型预报。

从地磁分量序列出发进行磁暴自动识别和强度类型预报,基本思路是通过在地面上实时监测地磁各分量的变化,在此基础上进行实时的特征提取,然后进行分类识别,以便尽早地识别出磁暴,并对识别出的磁暴进行强弱分类,以最终实现对磁暴的警报。由于磁暴发生时,所有地磁要素都发生剧烈的变化。其中,水平分量 H 变化最大,能代表磁暴过程的特点,尤其在中低纬度地区, H 分量的变化表现得较为突出,磁暴的大部分形态学和统计学特征都是依据中低纬度 H

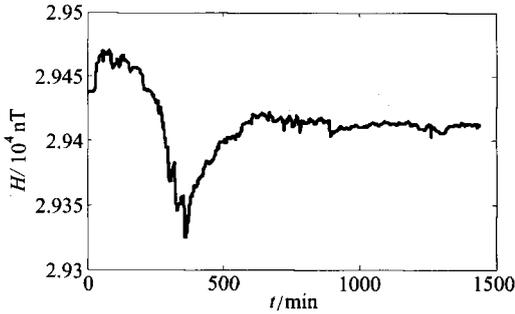


图1 1988年8月21日22时至22日21时
磁暴期间H分量的变化曲线

分量的变化得到的。所以,本文选用H分量的数据进行特征提取和识别实验。所采用的数据是从北京白家疃地磁台获得的1979年到1999年的地磁绝对观测分均值数据。图1就是白家疃地磁台记录的1988年8月21日22时至22日21时发生磁暴时H分量的变化情况。

1 地磁信号特征提取

地磁场的变化磁场可分为平静变化和扰动变化两大类,太阳静日变化和磁暴分别是最主要的平静变化和扰动变化类型。为了消除太阳静日变化对磁暴识别的影响,我们将21年所有的地磁H分量数据按月进行叠加然后求平均,从而分别得到一年1到12月份的通日变化曲线,以此作为每月的太阳静日变化曲线,用地磁H分量数据减去这个太阳静日变化即得到可能包含有磁暴变化的磁扰变化序列。我们的特征就是从这个磁扰变化序列提取出来的。

由于磁暴是在某些特定的太阳风和行星际磁场条件下,输入磁层的能量急剧增加,引起磁层-电离层系统的扰动而产生的,所以在发生磁暴时地磁分量序列会包含一些不同的频率成分,或者说相同频带内信号的能量有较大差别,在发生磁暴时某些频带内信号能量会相对减小,而另外一些频带内信号能量会相对增大。同样,不同强弱类型的磁暴在相同频带内信号的能量也会有较大差别。因此,在各频率成分信号的能量中,包含着丰富的磁扰信息,某种或某几种频率成分能量的改变,即代表了一种磁扰情况。利用这一特征就可建立能量变化到静扰类别和磁暴类型的映射关系,从而得到表征静扰类别和磁暴类型的特征向量。根据以上的分析,对地磁信号进行小波包分析提取特征的具体算法如下。

(1)取时间长度为 t 的H分量数据,按前述方法消除太阳静日变化,同时,为了消除直流分量的影响,减去序列均值,得到零均值的序列,然后采用db1小波对其进行3层小波包分解

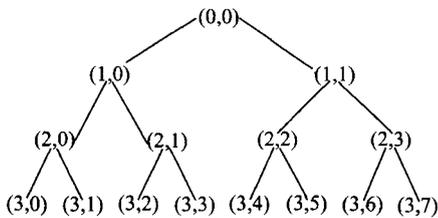


图2 小波包3层分解树结构

(胡昌华等,2004),分解结构见图2。图2中 (i,j) 表示第 i 层的第 j 个结点,其中 $i=0,1,2,3,j=0,1,\dots,7$,每个结点都代表一定的信号特征。其中, $(0,0)$ 结点代表原始信号, $(1,0)$ 代表小波包分解的第一层低频系数, $(1,1)$ 代表小波包分解的第一层高频系数, $(3,0)$ 表示第3层第0个结点的系数,其他依此类推。

(2)对小波包分解系数进行重构,提取第3层各频带范围的重构信号。分别以 $S_{30}, S_{31}, S_{32}, S_{33}, S_{34}, S_{35}, S_{36}, S_{37}$ 表示第3层从低频到高频8个频带范围的重构信号,则总信号可以表示为

$$S = S_{30} + S_{31} + S_{32} + S_{33} + S_{34} + S_{35} + S_{36} + S_{37} \quad (1)$$

(3)求各个频带信号的总能量。设 $S_{3j}(j=0,1,\dots,7)$ 对应的能量为 $E_{3j}(j=0,1,\dots,7)$,则有

$$E_{3j} = \sum_{k=1}^n |h_{jk}|^2 \quad (2)$$

其中 $h_{jk}(j=0,1,\dots,7,k=1,2,\dots,n)$ 表示重构信号 S_{3j} 的离散点幅值。

(4)构造特征向量。由于发生磁暴时,各频带内信号的能量有较大的变化。因此,以能量

为元素构造一个特征向量。特征向量 x 可以如下构造

$$x = [E_{30}, E_{31}, E_{32}, E_{33}, E_{34}, E_{35}, E_{36}, E_{37}] \quad (3)$$

由于8个频带范围信号的能量有时会有较大差异,造成特征向量各元素取值范围彼此差异很大,因此,需要对特征向量做一改进。通过实验我们发现,取 E_{3j} ($j = 0, 1, \dots, 7$) 的4次方根为元素,构造一个新的特征向量 x' 能达到较好的识别效果,即

$$E'_{3j} = (E_{3j})^{\frac{1}{4}} \quad (j = 0, 1, \dots, 7) \quad (4)$$

$$x' = [E'_{30}, E'_{31}, E'_{32}, E'_{33}, E'_{34}, E'_{35}, E'_{36}, E'_{37}] \quad (5)$$

2 磁暴与非磁暴的识别

采用 Fisher 分类器进行分类识别。Fisher 分类方法的思想就是找到一个最佳的投影方向,把模式样本在高维的特征空间里投影到这个方向的直线上,使样本最容易分开。Fisher 分类器最终将多维分类转化为一维分类问题,并且运算量小,速度快,能够满足实时系统的需要,其简单、容易实现的特性得到了广泛的应用。

在获得的21年地磁数据中,总共发生了441个磁暴,其中有32个磁暴数据记录不全,故总共有409个磁暴样本,取每个磁暴起始 t 时间长度的 H 分量数据提取特征,得到409个磁暴的特征向量。同样,取每个磁暴发生前 t 时间长度的 H 分量平静数据提取特征,得到409个非磁暴的特征向量,总共得到818个样本特征数据,实验在这些数据的基础上完成。在这些特征向量中每类分别随机选取50%作为训练集,用来训练 Fisher 线性判别式,其余的50%作为测试集,用来测试识别率。这个过程重复1000次,每次选取不同的50%作为训练集,1000次分类的平均结果作为我们最终的识别率。图3为不同时间长度 t 下磁暴与非磁暴平均识别率变化曲线。从中可以看出,随采用数据时间长度 t 增加,识别率上升,表明采用数据越长,所包含的信息就越丰富,识别的正确率也越高。另外,当时间长度 $t < 1$ h 时,识别率上升很快,当 $t > 1$ h 时,识别率上升较慢,因此,选 $t = 1$ h 的情况作进一步分析。

图4为 $t = 1$ h 时,前100次重复实验对应的识别率(做了1000次重复实验,为了图示清楚只取前100次绘图)。表1给出了平均识别结果及标准差,1000次两类平均识别率的标准差为0.0133,表明识别具有较高的稳定性。

3 磁暴强弱类型的识别

在总共409个磁暴中,有强磁暴(磁情指数 $K \geq 8$)24个,非强磁暴($K \leq 7$)385个,同样得到409个特征向量(样本),采用与磁暴和非磁暴识别同样的实验方法,即在这些特征向量中每类分别随机选取50%作为训练集,用来训练 Fisher 线性判别式,其余的50%作为测试集,用来测试识别率。这个过程同样重复1000次,1000次分类的平均结果作为我们最终的识别率。图5为不同时间长度 t 下强磁暴与非强磁暴平均识别率变化曲线。从中同样可以看出,随采用数据时间长度 t 增加,识别率上升,当 t 接近1小时时,识别率已经变得比较稳定。这表明两类磁暴的类型信息已经基本包含在磁暴起始1小时长度的数据中,这正好与我们的要求相符,因为在磁暴起始1小时内,磁暴扰动变化还不大,如果能利用1小时长度的数据即能把磁暴的强弱类型判断出来对我们的磁暴短时预报很有意义。因此,下面同样选 $t = 1$ h 的情况作进一步分析。

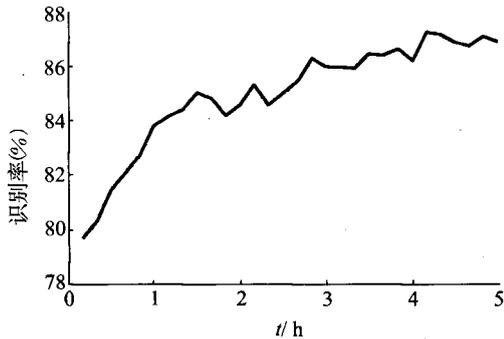


图3 不同时间长度 t 下磁暴与非磁暴平均识别率变化曲线

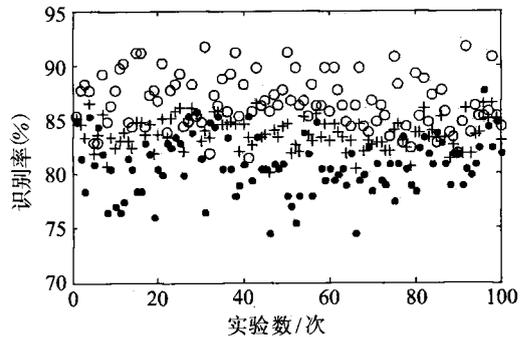


图4 前100次重复实验对应的识别率
•、○和+分别代表磁暴、非磁暴和平均的识别率

图6为 $t=1$ h时,前100次重复实验对应的识别率(同样做了1000次重复实验,为了图示清楚只取前100次绘图),从中可以看出,每次随机选的训练集和测试集得到的识别率差异较大,尤其是强磁暴的识别率差异最大。由于我们提取的特征可分性较差,另一方面由于强磁暴的训练和测试样本数太少,这就要求在应用中选取尽可能多的样本进行训练,或者选典型的样本进行训练,以提高可靠识别性。表2给出了 $t=1$ h的识别结果,强磁暴重复1000次识别率的标准差为0.1215,定量说明了强磁暴识别的稳定性。

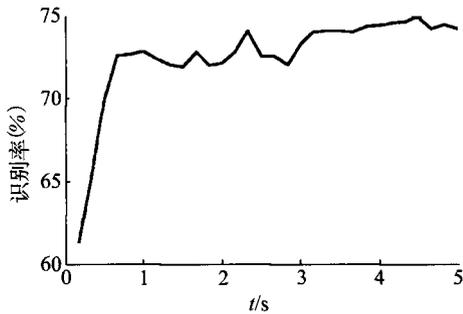


图5 不同时间长度 t 下强磁暴与非磁暴平均识别率变化曲线

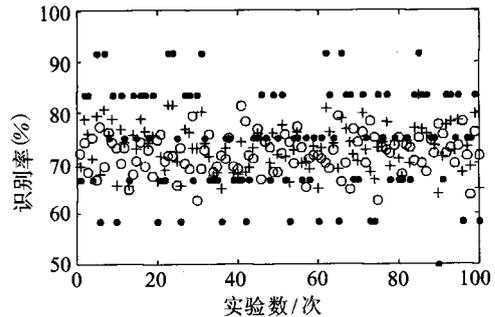


图6 前100次重复实验对应的识别率
•、○和+分别代表强磁暴、非磁暴和平均识别率

表1 磁暴与非磁暴的平均识别结果

	磁暴	非磁暴	两类平均识别率
识别率	80.8%	86.9%	83.8%
标准差	0.0249	0.0222	0.0133

表2 磁暴强弱类型的平均识别结果

	强磁暴 $K \geq 8$	非强磁暴 $K \leq 7$	两类平均识别率
识别率	71.7%	74.0%	72.9%
标准差	0.1215	0.0372	0.0569

4 结束语

本文的目的试图通过地面上实时监测地磁水平分量的变化,进行实时的特征提取,然后进行分类识别,以尽早识别出磁暴,并对识别出的磁暴进行强弱分类,以实现磁暴的警报。从以上结果来看,以地磁 H 分量序列各频段的能量分布为特征,对磁暴与非磁暴有较好的识别效果,说明发生磁暴时地磁信号所包含的频率成分确实发生了一定的变化,即各频段的能量分布发生了变化,因此,具有较好的识别效果。但以同样的特征,对磁暴的强弱类型进行识别,平

均识别率仅为 72.9%,可以看出,不同强弱类型的磁暴所包含的频率成分虽然有所差异,但差异不是很明显,即各频段的能量分布变化不大,以此作为特征识别效果不理想。

值得说明的是,文中所采用的消除太阳静日变化的方法是极其简略的,只要是相同的月,即使是不同的年、不同的日都认为具有相同的太阳静日变化,而实际上太阳静日变化是有逐年变化和逐日变化的。因此,可以考虑采用其他方法,以更好地消除太阳静日变化。当然,我们还可以从磁暴发生机理的角度,提取其他更好的特征,以得到高的磁暴分类识别率。在此基础上,结合磁暴发生的重现性等规律,实现对磁暴的警报。

参考文献

- 顾左文,高玉芬,范国华. 1995. 人工神经网络在磁情预报中的应用[J]. 空间科学学报, 15(1): 59~65
- 胡昌华,李国华等. 2004. 基于 MATLAB6. x 的系统分析与设计(第二版)[M]. 西安电子科技大学出版社
- 沈超,刘振兴. 2000. 磁暴和亚暴的理论预报模式[R]. 中国基础科学·国家重点基础研究项目, 5, 8~13
- 徐文耀. 2003. 地磁学[M]. 北京:地震出版社
- 魏奉思,徐亚,冯学尚,范全林. 2003. 利用“ISF”方法于地磁扰动事件的预报试验[J]. 中国科学(E辑), 33(5): 447~451
- 周晓燕,朱岗昆. 1994. 太阳风引起地磁扰动的辨识建模[J]. 地球物理学报, 37(3): 291~297
- 周晓燕. 1998. 利用人工神经网络预报大磁暴[J]. 空间科学学报, 18(3): 228~233
- Lundstedt H. 1992. Neural network and Prediction of solar-terrestrial effects[J]. Planet Space Sci, 40(4): 457~464
- Samuel Kugblenu, Satoshi Taguchi, and Takashi Okuzawa. 1999. Prediction of the geomagnetic storm associated Dst index using an artificial neural network algorithm[J]. Earth Planets Space, 51, 307~313

Recognition of magnetic storms at the early stage with the horizontal component of geomagnetism

Yi Shihua, Li Xihai and Li Zhigang

(The Second Artillery Engineering Institute, Xi'an 710025, China)

Abstract

In this paper, we extract the energy feature of the horizontal component of geomagnetism with wavelet packet analysis. Then we have investigated the variation of recognition rate at different time length $t=1$ h, and identified by use of Fisher's method. When, the average recognition rate of magnetic storms comes to 83.8%. If known being magnetic storm, the average recognition rate is 72.9%, at the same time it is classified into strong magnetic storm (geomagnetic activity index $K \geq 8$) and other magnetic storm ($K \leq 7$). If the horizontal component of geomagnetism is observed continually and in real time, the recognition of magnetic storm at the early stage can be realized in this way.

Key words: magnetic storm, wavelet packet analysis, feature extraction, recognition