

天气雷达组网拼图并行处理方法研究

王志斌¹, 万玉发¹, 沃伟峰²

(1. 中国气象局 武汉暴雨研究所, 湖北 武汉 430074;

2. 国家气象中心, 北京 100081)

摘 要:天气雷达组网拼图并行处理的目的是为了满足不同业务的需求, 通常的串行方法在雷达运行的六分钟周期内不能完成。利用共享存储多核处理器的高性能服务器, 研发了高频度实时全国天气雷达资料处理的并行计算系统, 根据单部天气雷达处理和多部雷达拼图的特点, 提出了进程级别和线程级别细粒度的 OpenMP 共享内存编程的两级并行方法, 给出一个合理的派生线程粒度数, 同时利用共享内存文件的方法提高程序的执行效率。实验结果表明, 使用两级的并行计算等方法可以满足全国天气雷达实时处理的要求。利用并行计算实现实时性要求高的业务系统是重要的手段。

关键词:天气雷达; 共享内存编程; 两级并行; 内存文件共享

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2013)07-0187-04

doi: 10.3969/j.issn.1673-629X.2013.07.048

Research on Parallel Progress Method of Weather Radar Data Mosaic

WANG Zhi-bin¹, WAN Yu-fa¹, WO Wei-feng²

(1. Institute of Wuhan Heavy Rain, China Meteorological Administration, Wuhan 430074, China;

2. National Meteorological Center, Beijing 100081, China)

Abstract: Weather radar mosaic parallel processing purpose is to meet the demand of real-time service, usually serial method cannot achieve in radar running cycle. Using high performance server of shared memory multiprocessor, develop the parallel computing system of high frequency real-time weather radar data processing. According to single weather radar processing and multi-radar mosaic characteristics, put forward the two stage parallel method of the process level and thread level of fine particle size OpenMP shared memory programming. Give a reasonable derived thread size, put forward the shared memory file method to improve the program execution efficiency. The experimental results show that, the use of two stage parallel calculation method to the national weather radar can satisfy the requirement of real-time processing. Use of parallel process is important for real-time system.

Key words: weather radar; OpenMP; two levels parallel; memory file share

0 引言

多普勒天气雷达是强对流天气监测的主要手段, 为了充分发挥多部雷达在联合监测天气中的作用, 应该把来自多部雷达的资料进行组网拼图, 由多个雷达重复取样获得的天气信息要比只由单个雷达取样获得的更加精确, 还可以在很大程度上解决因单部雷达观测的波束几何学原因(例如, 静锥区、波束展宽、波束高度、波束阻挡等)引起的很多问题^[1]。美国天气雷达的数据处理方式是数据质量控制放在雷达观测站上进行的, 对进行了质量控制的数据上传到雷达数据处理中心进行分区域的拼图。而我国的处理方式是把没有进行数据质量控制的数据直接上传到国家(或区域

中心)的数据处理中心后, 用户根据自己的需求进行处理分析^[2-4]。针对省一级的天气雷达组网, 每个省的雷达数量一般都不超过10部, 系统采用串行处理方式基本能够满足业务实际的需求。作者针对大区域邻近省的十几部雷达资料进行了并行处理的初步研究, 取得了较好的加速比。在前期研究的基础上对更大区域的雷达组网拼图进行实时的处理, 通过两级并行处理方式, 充分利用多核处理器, 大内存解决了组网拼图时效问题。多普勒天气雷达的扫描的间隔是六分钟, 某一区域的雷达资料要进行组网拼图一般情况下也应该在6分钟周期内到达资料处理中心, 收集完成后且6分钟周期内要完成一系列前期处理以及最后的组网

收稿日期: 2012-09-24

修回日期: 2012-12-26

网络出版时间: 2013-03-05

基金项目: 国家自然科学基金资助项目(41205086); 国家气象局业务项目; 科技部行业专项(GYHY201006002); 科技部业务专项

作者简介: 王志斌(1965-), 男, 副研究员, 研究方向计算机图像处理、数据库开发。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130305.0819.048.html>

拼图。如果 6 分钟周期内不能完成,一定要放宽处理周期时间,那么资料的完整性得不到保证,同样对生命史非常短的强对流天气系统的捕捉就会大打折扣,缩短雷达资料的处理时间就显得非常重要,采用并行处理是一个非常行之有效的手段。天气雷达组网拼图是在先处理单部雷达资料的基础上再进行组网的,非常适合并行处理的 JOIN-FORK 模式^[5,6]。作者通过两级并行处理方式完成了长江流域和全国天气雷达组网的工作,从而解决了传统的串行处理方式不能完成的工作。

1 天气雷达组网流程

天气雷达组网流程主要包含,各雷达站资料上传,雷达数据的抗错及质量控制,生成统一的雷达资料格式,组网拼图处理和基于组网拼图的产品以及在单站雷达基数据基础上进行的降水估测和组网产品。

1.1 资料传输处理

每一个区域(或省)有多部雷达,雷达站通过通信专线每六分钟实时地传输经过压缩处理的数据到区域(或省)的 FTP 服务器上。由于各个雷达开机时间和通信传输问题,资料到达 FTP 服务器上的时间都是不同时,系统查找所有雷达站的资料中最新的一个资料,取其资料中包含的时间,并以此为准向前一个时间段,在此范围内的资料认为都是同周期内的雷达,且可以进行组网的雷达资料。

1.2 分组处理

对于全国范围内的 150 多部雷达资料到达中国气象局雷达资料服务器上的时间差别大,为此把全国分为不同的区域进行处理,区域内的雷达资料到达时间还是比较统一的,差别不是太大,如把华中区域的雷达分为一组,把华南的雷达也分为一组等,共分了十二组来进行,即使某一个区域的雷达资料出现问题,也不影响到其它区域。分组处理另外一个原因是,通过共享内存处理的线程使用内存是有限制的,一般不超过 2G,由于单个雷达资料处理使用的内存非常大,派生线程个数大于一定的数量时,处理雷达的程序执行就会有问题,如果不分组进行,处理 150 多部雷达是不可能的。分组处理是使用不同的进程来进行的,每组包含多个程序,由每个程序派生出多个线程进行处理。

1.3 质量控制和数据规范存储

天气雷达获取的数据由于受环境和硬件的影响,并非都是有降水回波,它还包括非气象回波,如晴空回波、地物杂波、异常传播下的超折射回波。对于降水估测、自动临近预报来说,可能会对预报员产生误导,必须使用特定的算法进行识别和剔除。同时过滤掉其中的病态及错误数据。如弱回波过滤是根据不同的径

向距离设置相应的强度阈值,若回波强度小于该阈值则设为 0。对反射率场进行噪声过滤、弱回波消除等预处理后,再提取出雷达回波的三维结构特征量,运用模糊逻辑算法识别杂波并进行剔除。由于全国雷达型号有多种,且各种雷达资料原始存储格式都不一样,原始雷达资料读取时要根据雷达资料的型号的不同进行不同的处理,非常麻烦,为了用户使用方便和资料的规范,重新定义了一种通用格式,雷达数据经过质量控制和数据排序填充等过程后存储为通用的格式。

1.4 拼图处理和基于拼图的产品生成

雷达原始基数据资料是按照极坐标储存的,通常要转换为经纬网格的坐标,提供地理信息系统方便使用。这里有大量的计算,而且这种计算是一次性的,并不需要每次对雷达基数据进行转换时都进行,只要通过查表的方式进行相应的转换即可,从而减少了计算量。

通过质量控制后的单部雷达形成的数据是一个较小的范围经纬格方式的三维点场,通过拼接方式把指定区域范围内的多个单部雷达资料进行处理,形成大范围的经纬方式的三维格点场。再在这个经纬格点数据的基础上生成多种衍生产品。

降水估测是在单站雷达资料进行了控制的基础上还要进行特殊处理的模块,它要使用到过去一小时内的雷达基资料,还要通过混合扫描对地形的遮挡进行处理,使数据尽可能地发挥作用,这些方法在受地形影响山区更是离不开。因此估测降水系统对并行化的要求更高。

上述的处理大都是对单站雷达进行处理的过程,同时对多部雷达处理,可采用并行化方式进行,而且特别适合 JOIN-FORK 模式。雷达拼图资料的并行处理可以在具有多核或多 CPU 的单台高档服务器上运行,并行算法主要采用共享内存方式进行。对于多台的集群式的高性能计算机系统,也可以进行处理,笔者在这方面进行实验研究^[7,8],也取得了很好的并行效果,但由于计算机系统的维护和使用成本太高而没有得到推广应用,对于单台的高性能服务器可以满足用户的要求且维护方便。

2 天气雷达组网并行处理分析

2.1 多雷达串行处理分析

程序调度系统会周期性地启动资料处理程序,按照传统的串行处理方法是先查找资料,进行数据的质量控制,单站雷达资料存储和小范围的格点的生成,而后进行大范围的拼图处理和生成产品,再是进行雷达估测降水,这样一环接一环,是典型的串行处理方式。可以通过对目录的实时监控,对每到达一部雷达资料

就进行处理,也可以提高系统对雷达资料的处理速度,但这样处理对雷达资料的规整以及繁重的外设 I/O 处理非常不利,同时对后面的基于拼图处理也有很多不利的影响,典型的串行处理工作示意图如表 1。

如表 1 所示一个雷达运行周期内处理的时间可以进行粗略的估算(单站雷达质量控制的计算时间和回波数据的多少有直接关系)。

2.2 多雷达并行处理分析

雷达资料处理过程就是雷达资料收集处理,数据病态和质量控制处理,单站形成格点场和单站估测降水的循环过程,其中主要运算时间用在单站雷达质量控制和三维格点场大范围拼图以及单站降水估测方面,虽然就某一个雷达来说计算的时间不是太多,但是对 150 多部雷达资料进行处理就会需要进行很长时间。缩短处理时间的理想的方法是对多部雷达数据进行并行化处理。通过分析发现有些处理过程并不能够进行并行处理,如弱回波消除和杂波消除处理是一个顺序的过程,它们之间存在因果关系,单站雷达进行小范围的格点化后才能够进行大范围组网拼图处理,这也是一个串行过程。但并行的部分也比较多,如对单个雷达的处理可以分配到多核上运行,三维的拼图可以分层处理(即可以把多层分为每 10 层一组进行处理),把几层派生到多核上进行处理,因为每层处理过程是互不干扰的,分层处理是粒度比较小的处理,层次划分细了并不能提高运算速度,选择合理的线程粒度也非常重要。

2.3 内存共享编程及文件处理

在雷达资料处理过程中质量控制模块,三维格点场大范围拼图以及单站降水估测方面的计算是最耗时的,对于多部雷达采用共享内存的 OpenMP 进行并行计算,采用了多种实现方法,一种是基于 QT 的并行方法,另外一种是在 WINDOWS 下的制导语句实现,

两种方法实现的效率基本一致,其中 OpenMP 是用在多核处理器上的共享内存并行开发标准,该方法充分利用了多核处理器的特点,相对于进程调度方法,线程派生时消耗的计算机资源少,并行的效率高。在单雷达三维格点场形成后,有两种方法对其中间结果进行存储,一种是磁盘文件方式,另外一种用内存文件方式,单个雷达的三维格点场数据比较大(一般在 10M 左右),通过磁盘进行交换比较消耗时间,比较有效的办法是使用内存文件^[9],形成的单个三维文件保存于内存中,在进行大范围拼图时从内存中读取数据,这样可以提高文件的存取速度。对于现在的高档服务器,内存已经不再是稀缺资源,可以满足要求。

2.4 进程分块调度和线程粒度分析

由于全国雷达拼图采用了分区方式进行,处理某一区域雷达资料用一组程序来完成,对于多个区域雷达资料处理采用分时调度的方式进行(见图 1),每 6 分钟进行一次调度循环,组间间隔 30 秒,这样做减少了计算机资源的竞争,合理分配了 CPU 以及内存资源。系统在实际运行过程中,有时候一个雷达组的运算时间还是比较长,如需消耗 3 分多钟,这和这组程序没有得到需要的计算机资源而暂时等待有关,因此合理选择线程的派生个数非常重要,经过多次实验,如三维分层处理不能分得太细,如果粒度分得太细效率也不会得到提高,因此选择每次派生 2 个线程进行,每

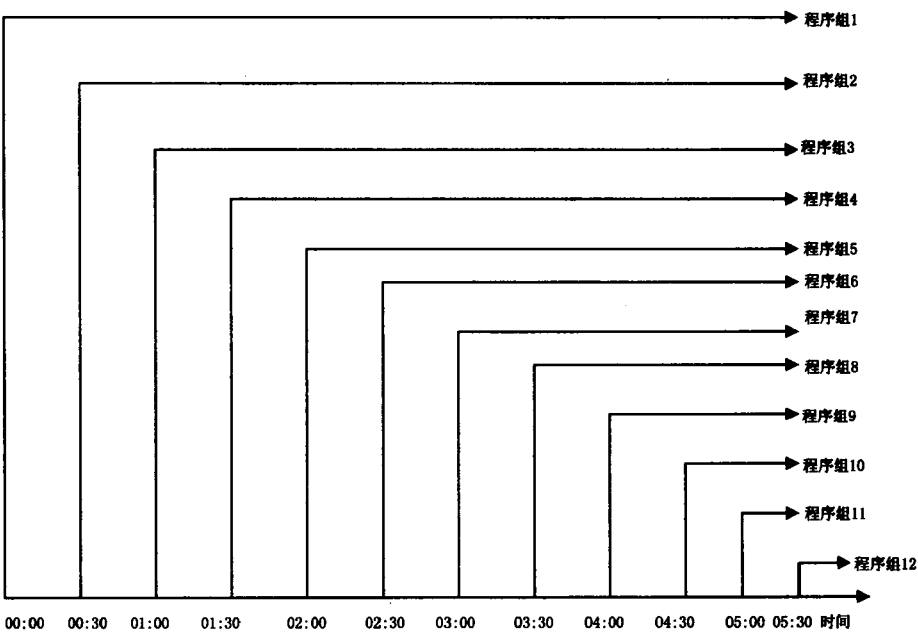


图 1 程序调度时间分配图

表 1 雷达资料处理工作示意图

资料规整和同步处理	数据病态及质量控制处理和存储	单站雷达资料形成格点场	降水估测处理	...	资料规整和同步处理	数据病态及质量控制处理和存储	单站雷达资料形成格点场	降水估测处理
第 1 个雷达				...	第 N 个雷达			

10 层一组,最后一组 11 层。这样的选择和具体计算机配置有关,如果计算机系统的 CPU 个数多,每次派生的线程个数也应该多些,这样程序组被分配到的计算机资源的可能性也大些,等待的时间也较少,另外就是修改 WINDOWS 线程的调度算法,更加适合雷达并行处理的要求。

3 运行结果及分析

在雷达资料处理计算机系统中,雷达数共有 150 多个,所使用的处理平台为 1 台高性能服务器,它包含 2 个独立 CPU,每个 CPU 含 4 核,共有 8 核,每个核又虚拟了 2 个 CPU,实际上共有 16 个 CPU,操作系统采用 WINDOWS2003。程序并行目的是并行方法比相应的串行方法运行得快,为了评估并行程序对雷达资料处理效率,采用加速比作为衡量标准^[10],加速比为串行程序执行时间/并行程序执行时间,对划分为武汉区域雷达耗时较长的拼图部分的运行结果进行了分析,区域分配雷达共 14 部。结果如表 2 所示,串行处理时间为 20917 毫秒,由于进行一次实验的数据有比较大的偶然行,故实验都是进行了 10 次以上,得出平均的计算结果。线程粒度划分结果见表 3。从表 2 可以看出,共享内存并行算法可以提高雷达资料处理速度,当进程数大于雷达个数时处理速度不再提高,也就是说,一个 CPU 处理一个雷达,CPU 再多也处于空闲状态。

表 2 雷达资料并行处理时间

CPU 个数	OpenMP 并行 (ms)	加速比
1	20917	1.000
2	13137	1.592
3	10731	1.949
4	10354	2.020
5	9724	2.151
6	9392	2.227
7	8433	2.480
8	7932	2.637
9	7671	2.767
10	7660	2.731
11	7203	2.904
12	7049	2.967
13	6901	3.031
14	6101	3.428
15	6100	3.429
16	6100	3.429

表 3 线程划分粒度表

三维雷达层数划分 (粒度)	执行时间 (ms)
1	7311
2	4411
3	2816
4	2895
5	2866
6	2695
7	2695
8	2896
9	2292
10	1603
11	3260
12	3904
13	3588
14	3219
15	2843
16	2773
17	2708
18	2286
19	2198
20	2048
21	2005

图 2 和图 3 分别为加速比图和执行时间图。

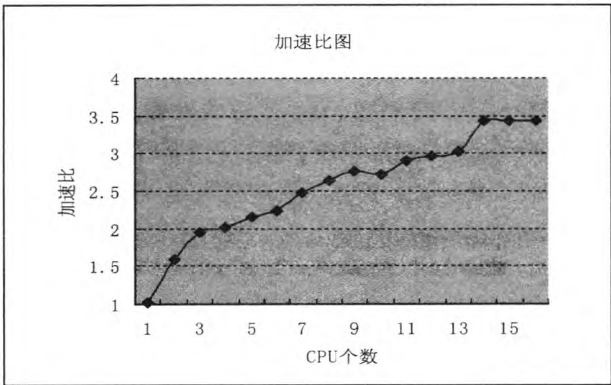


图 2 加速比图

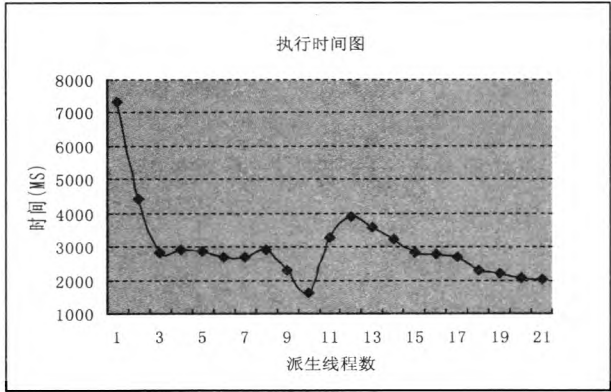


图 3 执行时间图

4 结束语

通过对天气雷达系统特点的分析,研究了在多核环境下,利用 OpenMP 方法对多雷达资料进行并行处理,通过划分区域、内存文件存储、流水线的调度等方法的结合,实验结果表明,多种方法的结合可以较大地提高系统处理资料的速度。满足了全国天气雷达实时组网的要求。同时对在某一特定时间周期内一定要完成任务处理的系统提出了解决问题的思路和方法。

参考文献:

[1] 王红艳,刘黎平,肖艳娇,等. 新一代天气雷达三维数字组网软件系统设计与实现[J]. 气象,2009,35(6):13-18.

[2] Istok M J, Okulski P R, Saffle R E, et al. NWS Use of FAA Radar Data-progress and Plans[C]//Proc. of the 21st International Conference on Interactive Information Processing Systems for Meteorology, Oceanography and Hydrology. San Diego, USA; American Meteorological Society, 2007.

[3] Rasmussen M S, Stuart M B, Karlsson S. Parallelism and scalability in an image processing application[J]. International Journal of Parallel Programming, 2009, 37(3):306-323.

[4] Michalakes J, Dudhia J, Gill D, et al. Design of a Next-generation Regional Weather Research and Forecast Model, Towards Teracomputing[C]//Proceedings of eighth ECMWF workshop on the use of parallel processors in meteorology. Riv-

(下转第 193 页)

首先读取数据包的显示模式标志,如果是SAR模块,则读取数据长度及包编号值,根据包编号值计算出数据在帧缓冲区的地址,然后启动DMA控制器,将数据缓冲区中的数据搬到帧缓冲区对应地址中。

3.3.3 图像处理

图像处理包括图层优先级的处理,像素在帧缓冲区中存储时,每帧图像包括多层图像,根据图像软件中图层优先级寄存器的配置,在多层图像中根据优先级选择最终输出的图像。

3.3.4 图像输出控制

图像输出控制逻辑按照VESA标准产生控制信号,控制信号包括帧同步信号,行同步信号,以及数据有效信号;在数据有效信号有效期间,输出红、绿、蓝图像数据信息。

3.3.5 软件

显示接口驱动软件包括接口底层图像数据处理软件和主机端驱动软件。其中接口底层图像数据处理软件采用Xilinx Software Development Kit环境开发,软件驻留于FPGA的PROM中,运行在FPGA内嵌PPC处理器上。上电后,软件自动加载运行。主机端驱动软件采用Tornado2.2环境开发,软件驻留于主机,主要是提供了一组接口函数用来完成主机端到显示接口数据缓冲的数据传送。

显示接口底层的图像数据处理软件,主要用于图像处理逻辑的配置,以及发送数据的处理,发送数据处理位于中断服务程序中,当主机将数据写入主机接口缓冲区并触发FPGA内部PPC处理器相应中断,图像数据处理中断服务程序需将数据包按约定解析成空操作、显示方式一、显示方式二、自检测、清屏五种工作模式,并分别由发送数据处理程序完成显示数据包的搬运和处理,并最终送往图像帧缓冲区中进行显示。

显示接口主机端驱动软件主要包括显示数据的读写函数,发送显示数据时,首先查询该片缓冲有效标志是否被清除,如果清除则将主机要发送的显示数据发送到RapidIO接口的发送缓冲中,并中断底层PPC处理器。

4 结束语

使用文中所述的硬件架构,包括RapidIO接口,高速图像缓存(QDR)可以解决图像显示过程中数据通路的带宽“瓶颈”问题;通过使用FPGA内嵌的PPC处理器,可以减轻主机的负担,提高处理效率;通过针对不同模式下的显示数据,设置了专用的图像数据硬件处理逻辑,极大地提高了显示效率,实现了雷达数据的高速实时显示。文中所述的方法已应用到了某型雷达中,测试结果表明,该接口功能以及性能可完全满足雷达系统的统一化显示需求,具有较好的应用价值。

参考文献:

- [1] Moir I, Seabridge A. 军用航空电子系统[M]. 吴汉平,译. 北京:电子工业出版社,2008:127-130.
- [2] 李 番,邬双阳,郑永超,等. 合成孔径激光雷达技术综述[J]. 红外与激光工程,2006,35(1):55-59.
- [3] 秦玉亮,王建涛,王宏强,等. 弹载合成孔径雷达技术研究综述[J]. 信号处理,2009,25(4):630-635.
- [4] 吴一戎,朱敏慧. 合成孔径雷达技术的发展现状与趋势[J]. 遥感技术与应用,2000,15(2):121-123.
- [5] 吴晓芳,代大海,王雪松,等. 合成孔径雷达电子对抗技术综述[J]. 信号处理,2010,26(3):424-425.
- [6] 蒋庆全. 有源相控阵雷达技术发展趋势[J]. 国防技术基础,2005,5(4):9-11.
- [7] Wang Wenqin, Cai Jingye. A Technique for Jamming Bi and Multistatic SAR Systems[J]. IEEE Geoscience and Remote Sensing Letters,2007,4(1):80-82.
- [8] Schumacher R, Schiller J. Non-cooperative target identification of battlefield targets classification results based on SAR images[C]//IEEE International Radar Conference. USA: IEEE,2005:167-172.
- [9] Moses R L, Ertin E, Austin C. Synthetic aperture radar visualization[C]//Thirty-Eighth Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: [s. n.], 2004:13-17.
- [10] 杨露菁,王德石,李 煜. 多视角SAR图像的静态建模与显示仿真[J]. 系统仿真学报,2010,22(2):506-507.
- [11] 黄世奇,刘代志,禹春来,等. SAR图像目标方位角估计与分析[J]. 系统仿真学报,2008,20(7):1795-1799.

(上接第190页)

- er Edge, New Jersey: World Scientific, 1999.
- [5] 董仁举,祝永志. 并行编程模型的研究与发展[J]. 计算机技术与发展,2011,21(1):92-94.
- [6] 王惠春,朱定局,曹学年,等. 基于SMP集群的混合并行编程模型研究[J]. 计算机工程,2009,35(3):271-273.
- [7] 王志斌,陈 波,万玉发,等. 天气雷达资料实时并行处理方法[J]. 计算机工程,2009,35(23):255-257.
- [8] 单 莹,吴建平,王正华. 基于SMP集群的多层次并行编


程模型与并行优化技术[J]. 计算机应用研究,2006,23(10):254-258.

- [9] 杨宁学,诸昌铃,聂爱丽. 内存映射文件及其在大数据量文件快速存取中的应用[J]. 计算机应用研究,2004,21(8):187-188.
- [10] 李 帆,冯西安,阎振华. 多目标检测并行处理软件的设计[J]. 计算机工程与应用,2010,46(10):226-228.

天气雷达组网拼图并行处理方法研究

作者：王志斌, 万玉发, 沃伟峰, WANG Zhi-bin, WAN Yu-fa, WO Wei-feng

作者单位：王志斌, 万玉发, WANG Zhi-bin, WAN Yu-fa(中国气象局武汉暴雨研究所, 湖北武汉, 430074), 沃伟峰, WO Wei-feng(国家气象中心, 北京, 100081)

刊名：计算机技术与发展

英文刊名：Computer Technology and Development

年, 卷(期):2013, 23(7)

本文链接: http://d.wanfangdata.com.cn/Periodical_wjfz201307048.aspx