



杨学军 蒋艳凰 周海芳 常志明/ 著

分布并行图像

处理 技术

Science & Technology Press

科学技术出版社

本书的出版得到国家863重大软件专项服务
器操作系统内核项目(2002AA1Z2101)资助



分布并行图像 处理 技术

杨学军 蒋艳凰 周海龙 常志明/著

《分布并行图像处理技术》内容提要、特点及出版意图

随着传感器和数据通信等相关技术的发展,数字图像处理的各个应用领域所获得的图像规模越来越大,数据量急剧膨胀。并行处理是提高运算速度最为有效的手段之一,本书以遥感图像处理作为应用背景,基于当前流行的 Cluster 体系结构,对分布并行图像处理技术展开了深入研究,在图像扭曲、图像压缩、图像配准、图像分割、图像分类等方面取得了一系列创新性的研究成果。

本书认真总结了作者的科研成果,内容新颖,突出理论联系实际,通过实例说明原理,富有启发性。本书的出版旨在对从事图像处理、分并行处理、遥感信息处理等领域研究的科技工作者、研究生等提供有益的参考。

分布并行图像处理技术

著 者: 杨学军 蒋艳凰 周海芳 常志明

责任编辑: 郑久平

出版发行: 湖南科学技术出版社

社 址: 长沙市湘雅路 276 号

<http://www.hnstp.com>

印 刷: 湖南航天长宇印刷有限责任公司

(印装质量问题请直接与本厂联系)

厂 址: 长沙望城坡

出版日期: 2005 年 6 月第 1 版第 1 次

开 本: 787mm×1092mm 1/16

印 张: 9.75

字 数: 185000

书 号: ISBN 7-5357-4296-3/TN·97

定 价: 28.00 元

(版权所有·翻印必究)

前 言

数字图像处理是信号与信息处理学科的一个重要组成部分,也是诸多计算机应用领域中一个十分活跃的分支。它的应用成果已遍布医学、工业制造、通讯、刑事侦察、遥感等许多领域。随着传感器和数据通信等相关技术的发展,各应用领域获得的图像规模越来越大,数据量急剧膨胀,迫切需要快速图像处理技术的支持。并行处理是提高运算速度最为有效的手段之一。在已有的并行计算机体结构中,基于分布存储的 Cluster 结构因性能价格比高、可扩展性好,已成为目前最为流行的并行体系结构。因此,研究分布存储环境下的并行图像处理技术不仅顺应了计算机技术的发展,而且可在一定程度上满足大规模图像处理的实际需求。

本书针对上述应用需求,深入研究了分布并行图像处理技术,主要研究内容包括:并行图像扭曲(几何变换)、并行图像压缩、并行图像配准、并行图像分割、并行图像分类等。

图像扭曲涉及到图像中每个像元的几何位置计算和灰度值重采样,属于计算密集型处理,当图像规模较大时,运算时间很长。已有的并行图像扭曲算法主要是对缩放、旋转等简单几何变换进行处理,对于复杂的几何变换,存在通讯开销大、负载分配不均衡等问题,因此这些算法的通用性较差。本书的第二章提出了两种分布并行图像扭曲算法:PIWA-LOC 和 PIWA-LIC。PIWA-LOC 算法通过局部输出区域计算,有效地解决了并行处理中的数据局部性问题,对于几何变换后负载分布基本均衡的图像扭曲问题,能够获得较好的并行性能。PIWA-LIC 算法对负载进行均匀分配,并通过局部输入区域计算保证数据局部性,该方法几乎适用于任意复杂的几何变换。这两种算法不仅适用范围广,而且显著提高了图像扭曲的处理速度。

大图像的传输和存储一直是数字图像处理中的技术难题。零树小波编码方法可以达到任意指定的压缩比,已经得到了广泛的应用。但是,由于在零树编码过程中需要多次扫描变换后的小波系数矩阵,这需要大量的处理时间。为了加快该算法的处理速度,本书的第三章分析了零树小波编码法在小波变换和零树编码过程中的并行性,提出了一种并行零树小波编码算法 PEZW-ID,该算法具有良好的数

据局部性、负载均衡性和可扩展性。

目前对大规模图像自动配准问题的并行算法研究十分有限,尤其是对粗粒度的并行算法研究比较欠缺;无论是自动点匹配方法还是全局配准方法,都尚未系统地针对当前通用高性能并行计算机开发高效的并行图像配准算法。本书的第四章提出了基于网络划分的自动点匹配算法,重点在自动提取特征点的问题上进行了探索。该方法既能提取出图像的明显特征点,又保证了特征点的均匀分布,促进了算法精度的提高。

流域变换是一种有效的图像分割方法,但效率低一直是困扰流域计算发展的问题,虽然已取得一些研究成果,但总的来说对其并行性的开发并不充分,在前期和后期都需要较复杂的变换还原计算。为了弥补已有算法的缺陷,本书的第五章提出了一种改进的基于区域关系图的并行流域分割算法 PGWA,并采用平滑算子解决流域变换存在的过度分割问题。

遥感是数字图像处理技术的重要应用领域,本书的第六章研究了遥感图像的监督分类问题。由于遥感图像规模庞大,导致遥感图像分类的速度难以满足适应的需求。本书对几种监督分类算法的并行性进行分析,并采用分布并行处理技术提高它们的学习与分类速度。

本书的第七章对 YH-PRIPS 系统进行了详细介绍。YH-PRIPS 是我们在研究过程中,针对遥感应应用领域,设计并实现的一个遥感图像并行处理系统。YH-PRIPS 采用基于 B/S 结构的三层应用模型,并运用高性能并行手段实现遥感图像的处理。系统集成本书所研究的并行算法,采用层次结构与模块设计,易于扩展、升级和移植。

全书的编写得到常春喜老师和郑久平编辑的关心和帮助,他们提出了一些很好的建议;本书的出版得到国家 863 重大软件专项服务器操作系统内核项目(2002AAIZ21001)资助,借此机会表示衷心感谢。

限于作者水平,书中的不当之处在所难免,敬请专家和读者朋友指正。

作者

2005 年 6 月

目 录

第一章 绪论

§ 1.1 基本概念	(1)
§ 1.1.1 数字图像处理	(1)
§ 1.1.2 并行处理	(1)
§ 1.2 研究动机	(2)
§ 1.2.1 图像处理对并行技术的需求	(2)
§ 1.2.2 并行计算机体系结构的发展趋势	(2)
§ 1.3 研究中涉及的若干问题	(4)
§ 1.3.1 并行模式	(4)
§ 1.3.2 并行算法研究的几个重要问题	(5)
§ 1.3.3 算法验证平台	(6)
§ 1.4 本书主要研究内容	(6)
§ 1.5 本书的组织结构	(7)

第二章 图像扭曲

(9)

§ 2.1 向后映射图像扭曲算法及其分析	(9)
§ 2.1.1 向后映射算法介绍	(10)
§ 2.1.2 复杂性分析	(11)
§ 2.1.3 并行图像扭曲技术现状	(12)
§ 2.2 PIWA-LOC 算法	(14)
§ 2.2.1 Cluster 环境描述	(14)
§ 2.2.2 PIWA-LOC 算法思想	(14)
§ 2.2.3 输入图像数据划分	(16)
§ 2.2.4 输出图像块存储结构	(17)
§ 2.2.5 局部输出区域计算	(18)

§ 2.2.6	局部重采样计算	(21)
§ 2.2.7	应用示例	(22)
§ 2.3	基于局部输入区域计算的并行图像扭曲算法 PIWA-LIC	(23)
§ 2.3.1	PIWA-LOC 算法的局限性分析	(24)
§ 2.3.2	PIWA-LIC 算法思想	(24)
§ 2.3.3	负载划分	(26)
§ 2.3.4	局部输入区域计算	(27)
§ 2.3.5	局部重采样计算	(31)
§ 2.3.6	应用示例	(32)
§ 2.4	算法性能分析的理论模型	(33)
§ 2.4.1	单任务并行处理性能分析	(33)
§ 2.4.2	多任务流水并行处理性能分析	(37)
§ 2.5	实验结果与分析	(39)
§ 2.5.1	遥感图像几何校正实验设定	(39)
§ 2.5.2	PIWA-LOC(CL)实验结果与分析	(40)
§ 2.5.3	PIWA-LOC(YH) 实验结果与分析	(42)
§ 2.5.4	PIWA-LIC(CL)实验结果与分析	(45)
§ 2.5.5	PIWA-LIC(YH)实验结果与分析	(47)
§ 2.5.6	特殊几何变换的实验结果	(49)
§ 2.5.7	实验结论	(50)
§ 2.6	小结	(51)
第三章	图像压缩	(52)
§ 3.1	图像压缩编码概述	(52)
§ 3.2	图像压缩编码的理论基础	(53)
§ 3.2.1	数据冗余	(53)
§ 3.2.2	信息量	(54)
§ 3.2.3	编码	(54)
§ 3.2.4	率失真定理	(54)
§ 3.3	图像并行压缩编码的研究	(55)
§ 3.3.1	简单压缩编码	(55)
§ 3.3.2	预测压缩编码	(57)
§ 3.3.3	变换压缩编码	(57)

§ 3.4 并行零树小波编码算法的研究	(59)
§ 3.4.1 零树小波编码算法的背景	(59)
§ 3.4.2 零树小波编码的串行算法	(59)
§ 3.4.3 串行算法的分析	(60)
§ 3.4.4 PEZW-ID 算法流程	(62)
§ 3.4.5 PEZW-ID 算法的正确性及性能分析	(63)
§ 3.4.6 算法性能分析	(66)
§ 3.4.7 实验结果	(67)
§ 3.5 小结	(69)
第四章 图像配准	(71)
§ 4.1 图像自动配准技术研究	(71)
§ 4.1.1 图像配准的四个要素	(71)
§ 4.1.2 自动点匹配与全局配准	(75)
§ 4.2 一种基于网格划分的自动点匹配并行算法	(77)
§ 4.2.1 特征点的自动提取算法	(78)
§ 4.2.2 搜索同名点	(81)
§ 4.2.3 并行性分析	(82)
§ 4.2.4 实验结论	(82)
§ 4.3 小结	(84)
第五章 图像分割	(85)
§ 5.1 流域变换的定义	(86)
§ 5.1.1 连续域定义	(86)
§ 5.1.2 离散域定义	(87)
§ 5.2 流域变换的串行实现算法	(91)
§ 5.2.1 模拟“泉涌淹没”的流域算法	(91)
§ 5.2.2 基于地形学距离的流域算法	(92)
§ 5.3 数据划分与复制	(95)
§ 5.4 一种改进的基于区域关系图的并行流域算法 PGWA	(96)
§ 5.4.1 M-R 算法及其缺陷	(96)
§ 5.4.2 改进的算法 PGWA	(98)
§ 5.5 结合地形学距离的流域分割并行优化算法 E-PGWA	(103)

§ 5.5.1 伪种子点和区域关系图	(104)
§ 5.5.2 E-PGWA	(106)
§ 5.6 过度分割问题及其解决方法	(110)
§ 5.7 实验结论	(113)
§ 5.7.1 测试结果与分析	(113)
§ 5.7.2 与相关工作的比较	(116)
§ 5.8 小结	(117)
第六章 图像监督分类	(118)
§ 6.1 概念描述	(118)
§ 6.2 最近邻分类法的并行化	(119)
§ 6.2.1 最近邻分类算法 Knn 的并行化	(119)
§ 6.2.2 并行 Knn 算法性能分析模型	(120)
§ 6.3 几种神经网络算法的并行化	(122)
§ 6.3.1 三种结构化神经网络	(123)
§ 6.3.2 神经网络算法的并行化	(124)
§ 6.3.3 并行神经网络算法性能分析模型	(125)
§ 6.4 实验结果与分析	(126)
§ 6.4.1 实验设定	(126)
§ 6.4.2 基于 CL 平台的实验结果	(128)
§ 6.4.3 基于 YH 平台的实验结果	(130)
§ 6.5 小结	(132)
第七章 遥感图像并行处理系统 YH-PRIPS	(133)
§ 7.1 系统结构设计	(133)
§ 7.1.1 YH-PRIPS 应用模型	(133)
§ 7.1.2 YH-PRIPS 软件运行框架	(134)
§ 7.2 系统功能介绍	(135)
§ 7.3 小结	(138)
第八章 研究展望	(139)
参考文献	(140)

第一章 绪论

§ 1.1 基本概念

§ 1.1.1 数字图像处理

图像是用各种观测系统以不同形式和手段观测世界而获得的,可以直接或间接作用于人眼并进而产生视知觉的实体。常见的图像是连续的,为了便于用计算机对图像进行加工,需经过采样处理,将连续的图像离散化,这种离散化后的图像即为数字图像。数字图像处理是利用计算机和其他电子设备对数字图像进行加工处理的过程,包括图像获取、图像编码、图像变换、图像恢复、图像增强、图像识别等内容[Zhang 1999]。

根据抽象程度和研究方法的不同,数字图像处理可分为三个层次:数据处理层(也称为狭义的图像处理层)、信息提取层(也称为图像分析层)和知识表示层(也称为图像理解层)。数据处理层主要是对图像进行加工以改善图像的视觉效果,为图像识别打下基础,它也包括对图像进行压缩编码以减少存储空间和传输时间,该层主要是像素级的处理。信息提取层主要是对图像中感兴趣的目标进行检测和测量,该层处理的对象是图像,处理后的结果是对目标的特点与性质进行描述的表征数据,如类别符号、目标边缘、纹理或统计信息等。知识表示层则是在信息提取层生成结果的基础上,进一步研究图像中各目标的性质和它们之间的相互关系,形成人们易于理解的知识表示。知识表示层的抽象程度最高,基本上是对抽象出来的符号进行运算,其处理过程和方法与人类的思维推理有许多类似之处。

§ 1.1.2 并行处理

并行处理(Parallel Processing)[Chen 1994]是计算机系统中能同时执行两个

或更多个处理机的一种计算方法,不同的处理机可同时工作于不同的任务或同一任务的不同方面。并行处理的主要目的是节省大型和复杂问题的执行时间,提高处理速度。当代面临着的重大科学技术问题要依赖计算技术协助解决,一方面要作大型计算以得到更精确的解,另一方面要作计算机模拟,以便进一步了解所探讨问题的结构与运动规律。这两个方面都离不开并行处理技术。

为了能够使任务的不同部分在不同的处理机上同时执行,首先需要开发任务的并行性,然后进行并行程序设计[Shen 1997],将任务的各部分分配到不同处理机上。由于任务的各个部分之间常常存在相互依赖关系,因此并行程序设计很难自动实现,往往需要由程序员手动实现。

并行性主要是指同时性或并发性[Li 2000],并行处理是一种相对于串行处理的处理方式,它着重开发计算过程中存在的并发事件。并行性通常划分为作业级、任务级、例行程序或子程序级、循环和迭代级以及语句和指令级。作业级的层次高,并行处理粒度粗。粗粒度并行性开发主要采用 MIMD 方式,而细粒度并行性开发则主要采用 SIMD 方式。开发计算机并行性的方法主要有:资源重复、时间重叠和资源共享三种方法。

并行程序设计方法的一个最基本的观点,就是把一个事物的行为看成是多个事物互相作用的结果。这是一个观念上的根本转变,根据这个观点,并行程序设计方法的核心内容就是并行划分与算法映射。

§ 1.2 研究动机

§ 1.2.1 图像处理对并行技术的需求

数字图像处理是信号与信息处理学科的一个重要组成部分,也是诸多计算机应用领域中一个十分活跃的分支。它的应用成果已遍布医学、工业制造、通讯、刑事侦察、遥感等许多领域。随着计算机、集成电路等技术的飞速发展,图像处理无论在算法上、系统结构上,还是在应用上都取得了长足的进展。但是,图像处理依然面临着许多挑战性的问题[Su 2002]。提高图像处理的速度就是一个重大挑战课题,该课题是由图像数据的特点和图像处理算法的复杂性引起的,是一个长期困扰图像界的棘手问题。并行处理技术是提高处理速度的最有效的手段之一,随着高性能并行处理系统的发展,图像并行处理技术为提高图像处理速度提供了更大的空间。

图像并行处理技术的应用是多方面的,就其紧迫性来看,其应用主要集中在军事(如目标识别、地形匹配)、工业自动化(如实时电路检测、智能机器人)、遥感(如资源探测、地理导航、军事遥感)、刑事侦查(如指纹、人脸识别)等领域。这些应用对图像处理速度的要求是十分强烈的。例如,在军事遥感领域,早在 1984 年 6 月,美国从夸贾林岛试验基地发射了光学制导拦截导弹,在太平洋上空拦截并摧毁了民兵 1 号洲际导弹,安装在拦截导弹上的长波红外线传感器的工作借助于两台计算机,数据处理速度大约是每秒 1.8×10^7 个运算指令[Wang 1995]。而在未来星、空、地一体化的遥感系统中,图像并行处理技术更是不可缺少。本书将以遥感作为主要应用领域,通过对遥感图像的处理来验证我们所研究的算法的正确性。

§ 1.2.2 并行计算机体系结构的发展趋势

按照 Flynn 分类法,计算机体系结构可分为单指令单数据流(SISD)、单指令多数据流(SIMD)、多指令单数据流(MISD)和多指令多数据流(MIMD)四类,其中 SIMD 和 MIMD 是应用最为普遍的两种类型[Li 2000, Zheng 1998, Zeng 1999]。20 世纪 80 年代针对 SIMD 的并行研究与应用较为活跃,而 MIMD 并行系统由于可以实现作业、任务、指令和数据各个级别的并行,已逐渐成为当今并行处理系统的主流。

根据各处理机间的耦合程度的不同,MIMD 系统又可分为共享存储和分布存储两类。共享存储的并行机由于受到存储带宽的限制,可扩展性较差,往往不能适应数据规模很大的并行计算。相反,在采用分布式存储的系统中,各个处理机不仅有自己的局部存储器,而且可以通过互网络访问其他结点上的存储器。由于高性能的互网络可以连接成千上万个处理结点组成大规模并行处理系统(MPP, Massively Parallel Processors),因此这种多机系统具有良好的可扩展性。随着单机性能的提高和互联技术的发展,MPP 技术进一步应用到普通的分布式环境中,产生性能价格比高、可扩展性好、松耦合结构的 Cluster 系统,使得并行应用的空间得到进一步拓展。

正是由于 MIMD 系统具有很强的通用性和可重构性,在图像处理领域,MIMD 系统正逐步取代 SIMD 系统的主导地位[Pitas 1993, Yang 1998],其中松耦合的 Cluster 结构因具有性价比高、可扩展性好等优点,备受用户青睐。鉴于此,本书的研究均是针对主流的分布式 MIMD 系统展开,且主要是基于 Cluster 并行计算环境。

§ 1.3 研究中涉及的若干问题

§ 1.3.1 并行模式

根据图像数据所具有的数据量大、规律性强、相关性强、视频传输速率高等特点以及图像处理算法所具有的一致性、分层性、邻域性、行顺序性等特点[Su 2002], 图像处理的并行性表现为四种类型[Pitas, 1993]: ①几何图形的并行性; ②相邻并行性; ③像素位的并行性; ④操作的并行性。而并行性的实现模式大致有三种: ①流水线并行; ②功能并行; ③数据并行。

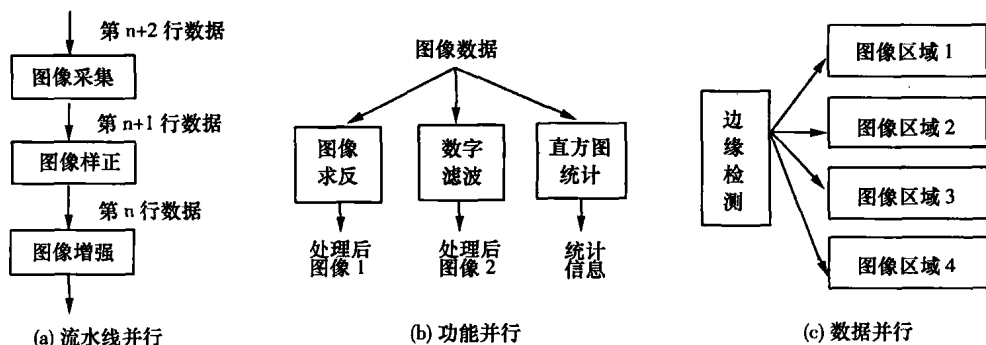


图 1.1 图像处理的三种并行模式示例图

图 1.1 给出了三种实现模式的示例。图 1.1(a)是对图像进行流水线并行的例子,不同的数据行陆续进入流水线,轮流经过各个功能模块。这是一种多功能流水线,其中实现了多个算法;与此相对应的是单功能流水线,即某一功能还可以进一步细分为不同的操作,进而发掘更细粒度的并行。图 1.1(b)是功能并行的例子,它的特点是各功能模块处理的数据是相同的。图 1.1(c)是数据并行的例子,一幅图像被分为 4 块,对每一块数据施加相同的操作,最后可能存在图像块边缘衔接的问题。

对于流水线并行,流水线上的各个步骤所完成的功能不同,各功能所处理的数据也不同,兼具了功能并行和数据并行的特点,并行执行时间取决于执行时间最长的步骤和数据划分的粒度,如果设计得当,这种并行模式能获得很高的效率,但缺点是需要硬件的支持,不适合目前主流的并行处理结构。为了加快某一图像处理算法的运算速度,采用单纯的功能并行难度较大,因为同一算法内部的各个步骤之

间多数是相关的,不易分解。数据并行的思想较为自然,这是因为图像数据具有一致性和邻域性的特点,同时这种并行模式更适合于当前主流的并行计算系统(如MPP、机群系统),因此本书所设计的并行算法均采用数据并行模式。

§ 1.3.2 并行算法研究的几个重要问题

本书中所研究的并行处理算法均需要考虑以下几个重要问题:

(1) 划分(Partition) 设计并行算法的第一步是问题的分解[Chen 1994, Shen 1997]。根据前面对并行模式的讨论,可知分解的方法有功能分解和域分解(数据划分)两种。本书中讨论的算法都是采用域分解的方式,即图像先被划分为若干子图,然后再在子图上施加合适的计算。

(2) 数据分布 这是分布存储结构所特有的问题,任务到处理机的理想映射是满足处理机的利用率达到最大而通信代价尽可能小。在我们的研究中,每个处理机分配一个子域(子图)并使得逻辑上相邻的子图被映射到实际相邻的处理机上。为了避免过多的I/O请求,先由一个处理机加载整个输入图像并在局部存储器内完成必要的的数据划分,然后以消息传递的方式将各个子图分派到其他处理机上去。对于一些专用的软件,如PIPS[Nolle 1994]和PUL-RD[Default 1995],这一操作由系统完成,而对MPI这种通用平台,则需要程序员来负责这部分工作。

(3) 聚合(Agglomeration) 聚合[Moga 1997]一般指组合任务以提高效率,但也指为同一目的进行的数据拷贝或复制计算,实质上是一个数据和计算局部化的问题。数据拷贝是指:由于图像操作多是面向邻域的,因此子域边界的像素常常引起数据的交换。为了增加局部性以避免频繁的通信,我们根据算法的需要在一个子域的相邻处理机上复制一部分边界像素数据,即各个相邻子域之间有一部分重叠的数据。复制计算是指:算法中某些步骤的操作对象是整幅图像或全局数据,且计算时间基本可忽略,如果仍然考虑并行的话,会因为计算和通信的开销比值较小而得不偿失,因此我们让每个处理机都同时执行这些步骤来消除程序的相关性。这两点在后续章节中都有所体现。

(4) 负载均衡 由于并行计算时间取决于系统中执行时间最长的处理机,因此负载是否平衡是影响并行性能的一个重要问题[Chen 1994, Shen 1997]。对图像处理而言,负载是否平衡主要表现在两个方面:①图像分块是否均衡;②对相同或不同图像块实施的计算总量是否平衡。因此,为了达到负载均衡的目的需要仔细设计数据和计算的划分方法。

(5) 开销控制 开销主要来自两个方面,一是计算开销;二是通信开销。开销控制是并行算法设计的关键问题,控制开销的方法有尽量减少冗余计算量、减少通

信次数和通信量、合并和延迟通信、计算与通信重叠等,这些方法在本书的算法设计中均有所体现。

§ 1.3.3 算法验证平台

本书的研究工作均基于主流的 MIMD 结构的并行分布式系统展开,如 MPP 或机群系统。在整个研究过程中,我们主要采用以下两个验证平台对所研究的算法进行性能测试:

①机群系统(以下简称为 CL),主要配置为:16 个同构的处理结点,每个结点的配置为 Pentium4-2G 的 CPU,1GB 的局部存储系统,各处理机通过快速交换式以太网互联,物理带宽为 100Mb/s。软件环境为 Linux 操作系统,基于 MPICH 的消息传递并行库。

②银河并行机(以下简称为 YH),主要配置为:64 个同构的处理结点,每个结点采用 16.6 亿个浮点操作/s 的 CPU,各结点的本地内存为 1GB,网络拓扑结构为胖树结构,点对点带宽为 1.6GB/s,软件环境为 YHUNIX 操作系统,基于 YH-MPI 消息传递并行库。

此外,还需要说明的是,本书所研究的算法均采用数据并行模式和消息传递的并行编程方式,这些算法可以处理大小为 $M \times M'$ ($M, M' \in Z^+$) 的图像,但为了描述的方便,在有些算法的研究中我们设定 $M=M'$,并统一描述为图像规模 M ,但这种简化并不影响算法的一般性。

§ 1.4 本书主要研究内容

本书对数字图像处理领域中复杂度高、处理速度较慢的几个关键问题进行研究,利用并行手段提高它们的处理性能。主要研究内容包括:并行图像扭曲(几何变换)、并行图像压缩、并行图像配准、并行图像分割、并行图像分类等。

(1)图像扭曲涉及图像中每个像元的几何位置计算和灰度值重采样,属于计算密集型处理,当图像较大时,运算时间很长。已有的并行图像扭曲算法主要是对缩放、旋转等简单几何变换进行处理,针对复杂几何变换的研究甚少,因此这些算法的通用性较差。我们针对分布并行环境,研究了并行图像扭曲的数据局部性和负载平衡问题,提出了基于局部输出区域计算的 PIWA-LOC 算法和基于局部输入区域计算的 PIWA-LIC 算法,并对这两种算法的特点、应用范围进行了详细阐述。

(2)对于大图像而言,数据的传输、存储都是亟待解决的问题。零树小波编码

方法可以达到任意指定的压缩比,已经在图像处理中得到了广泛的应用。但是,由于在零树编码过程中需要多次扫描变换后的小波系数矩阵,这需要大量的处理时间。我们分析了算法在小波变换和零树编码过程中的并行性,提出了一种并行零树小波编码算法 PEZW-ID,并对算法的性能进行分析。

(3)目前对大规模图像自动配准问题的并行算法研究十分有限,尤其是对粗粒度的并行算法研究比较欠缺;无论是自动点匹配方法还是全局配准方法,都尚未系统地针对当前通用高性能并行计算机开发高效的并行图像配准算法。我们提出了基于网格划分的自动点匹配算法,重点在自动提取特征点的问题上进行了探索。

(4)流域变换(又称分水岭变换)是一种有效的图像分割方法,但效率低,一直是困扰流域计算发展的问题,虽然已取得一些研究成果,但总的来说对其并行性的开发并不充分,在前期和后期都需要较复杂的变换还原计算。为了弥补已有算法的缺陷,我们提出了一种改进的基于区域关系图的并行流域分割算法 PGWA,并采用平滑算子解决流域变换存在的过度分割问题。

(5)遥感是数字图像处理技术的重要应用领域,图像分类是遥感图像处理的重要内容。由于遥感图像规模庞大,导致遥感图像分类的速度难以满足应用的需求。我们对几种监督分类算法的并行性进行分析,并采用分布并行处理技术提高它们的学习与分类速度。

(6)针对遥感应用领域,我们设计并实现了一个遥感图像并行处理系统 YH-PRIPS。

根据 § 1.1.1 节对数字图像处理的层次划分,我们所研究的图像扭曲、图像压缩、图像配准均属于数据处理层的内容,而图像分割、图像分类则属于信息提取层的内容。

§ 1.5 本书的组织结构

本书的内容以分布并行图像处理为主线,围绕图像处理的若干环节展开讨论,它们分别是:图像扭曲、图像压缩、图像配准、图像分割、图像分类。全书共分八章。

第一章为绪论,首先对本书的研究内容所涉及的几个概念进行介绍,阐述了我们研究分布并行图像处理技术的动机,以及与研究相关的若干问题。

第二章主要研究并行图像扭曲技术,提出两种基于不规则区域计算的并行图像扭曲算法,解决了并行处理过程中的数据局部性与负载平衡问题。研究中,我们对这两种算法所涉及的不规则区域计算问题,以及不规则图像块的处理问题展开了深入讨论,并从理论和实验两个方面对算法的性能进行了详细分析。

第三章对图像压缩技术进行了简单介绍,针对零树小波编码在小波变换和零树编码中占用大量运算时间,导致处理速度缓慢这一问题,提出了一种基于图像划分的并行零树小波编码算法,并对其性能进行分析。

第四章综述了图像配准的相关技术,然后重点介绍了自动图像配准方法,并从新的角度将配准算法分为点匹配和全局配准两大类,归纳总结了这两类算法的特点。提出了一种基于网格划分的自动点匹配并行算法,并对算法的特点进行了详细分析。

第五章主要研究并行流域分割算法,首先分析了 M-R 算法的缺陷和逻辑错误,在修正错误的基础上,面向主流的分布式存储并行系统提出了基于区域关系图的并行流域分割算法;然后在此基础上结合距离定义,还提出了一种引入伪种子点的流域分割并行优化算法,最后进行了性能分析。

第六章研究了分布存储环境下的并行监督学习与分类技术,对最近邻法的分类过程和三种结构化神经网络算法的学习过程进行并行化处理。

第七章对我们开发的遥感图像并行处理系统 YH-PRIPS 的结构与功能设计进行详细阐述。

第八章对下一步的工作进行了展望。