

• 软件技术与实现 •

文章编号: 1000-3428(1999)10-0008-03

中图分类号: TP 317.4.315

图象档案管理系统C2风格构架的研究

李丽燕, 钱乐秋, 汪 洋

(复旦大学计算机科学系软件工程实验室, 上海200433)

摘 要: 设计和开发图象档案管理系统重点要考虑以下问题: 档案的快速、完整收集, 归档自动化, 大容量存储和实时动态调档。C2是一种基于构件和连接器的构架风格, 对具有图形用户界面的应用有良好的支持。设计了一个C2风格的图象档案管理系统构架, 说明了如何用构件构架理论去指导、监控系统开发, 并通过应用实例来讨论有关具体问题。

关键词: 软件构架; 构架风格; 构件; 连接器

Research of Image Document Management System Using C2 Style Architecture

Li Liyan Qian Leqiu, Wang Yang

(SE Lab, Computer Science Department of Fudan University, Shanghai 200433)

【Abstract】 Designing and development of image file management system focus on these problems: quick and complete files collection, automatic files classify, store and real time inquiry. C2 is a kind of architecture styles based on component and connector. It supports the application with GUI well. This paper advances a C2 style image file management system architecture, and illuminates how to use architecture_based theories to conduct and monitor the development of system. The corresponding examples are presented in the paper.

【Key words】 Software architecture; Architecture style; Component; Connector

1 背景介绍

与一般的文字档案管理系统相比, 图象档案管理系统具有以下特点: (1) 档案收集技术及设备较复杂, 一般需要配备扫描仪、数字摄影机等硬件; (2) 难以自动化建档, 由于难以获知图象代表的涵义, 需要另辟渠道来获得建档信息; (3) 图象档案容量大, 压缩、存储和快速调档问题必须细加考虑。因而, 图象档案管理系统较之一般的文字档案管理系统复杂得多。

该档案系统是基于“黄页电话号码簿档案管理”这个特定领域的。黄页出版有广告流程和列名刊登流程。广告流程为: 签署合同; 广告设计; 生成广告文件; 排版印刷; 合同存档。列名刊登流程为: 发用户函, 询问列名刊登格式; 用户回函; 回函录入; 排版印刷; 用户回函存档。人们可能会问, 回函信息已经录入, 一定程度上可以把这当作文字档案, 为什么还要作为图象档案存档? 这是因为当用户对列名刊登有异议时, 有力的证据只能是这张用户填写的回函。然而回函数量巨大, 难以手工查找。因而要建立档案库, 把合同、回函转换成图象文件存储到计算机中, 以备查询。

为了达到软件重用(Reuse)和替换(Substitutability)的目的, 我们采用基于构架的开发方法, 着眼于寻找特定领域中不同个体的普遍性, 把系统的整体结构组织成相互作用的构件, 最终达到减少开发成本的目的。

基于软件构架的开发方法的重点从传统的代码行转变成大粒度的构架元素--构件和连接器以及它们之间相互联系的结构。它的开发过程一般为: 针对特定领域进行领域分析; 设计并形式化描述参考构架; 面向特定的应用特殊化参考构架, 得到应用构架; 构架的实例化。所谓构架的实例化, 是指构架中的抽象构件与具体构件的绑定, 具体构件可以从现有构件中选择、继承或改进。也可以是新设计的构件。

2 构架描述

2.1 构架的建模特性

构架描述的基本块是构件、连接器及其拓扑结构。构件即系统元素, 一个构件可以小如一个简单的过程或大如另外一个构架; 连接器描述构件之间的相互作用和管理这些相互作用的规则; 拓扑结构则是描述构架结构中构件和连接器的联系图。每一基本块都有其建模特性(见图1):

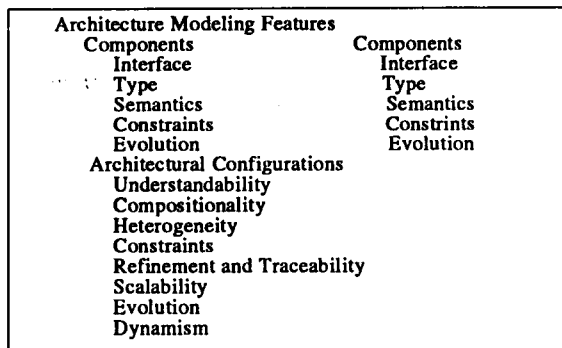


图1 每个基本块的建模特性

2.2 构架描述语言—ADL(Architecture Description Language)

为了支持基于构架的开发, 形式化的建模符号和分析、开发工具是必需的。ADL是一种描述软件系统的概念构架描述语言, 提供具体的语法和概念框架来刻画构架, 着力于系统的整体结构而不是任何特定模块的具体实现。ADL必须包括构件、连接器和拓扑结构3个方面。但是目前几乎没

基金项目: 上海市科委“九五”攻关项目的部分资助

作者简介: 李丽燕(1975~), 女, 研究生, 主研方向: 软件工程

收稿日期: 1999-01-06

有一种ADL能支持这里所列举的所有特性。

2.3 C2简介

C2是一种基于构件和连接器的构架风格,对具有图形用户界面的应用有良好的支持。可以非形式化地把C2总结为由连接器连接起来的构件的层次网。每个连接器连接的构件或其他连接器的数量没有限制。如图2所示。它具有部分可见性(向下独立性)的特点。每个构件都只知道上层构件的接口,并不知道具体实现。其下层构件则完全不可见了。

每个构件都有一个顶层域(Top Domain)和底层域(Bottom Domain)。顶层域定义一组向上传送的请求和可接受的上层发送来的通告集(Notification);底层域定义一组向下载送的通告和可作响应的下层传送来的请求集。构件通过顶层域向上发送请求并接受上层通告,通过底层域接收下层请求并向下载通告。所有构件间的通信都严格地通过一个特定端口的消息交换来进行。图3给出C2构件的例子。从这个例子可以看到:C2构件本身也可以是一个构架。它的内部结构可以完全独立于外部构架。

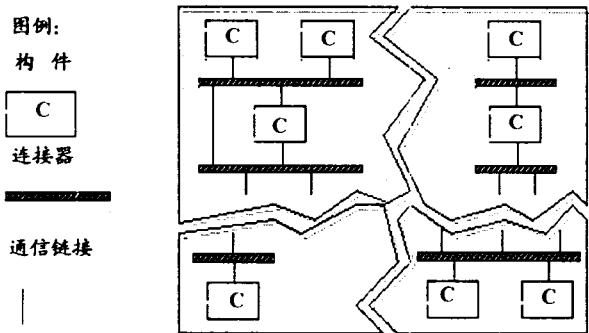


图2 C2风格构架示例

对前面提出的构架建模特性,C2作如下的支持。它完全支持构件方面的特性,对于构件的演变,C2更是提供高级的子类机制(SubTyping)和类型检测机制。C2不提供连接器的接口定义,而是结合到它连接的构件的接口中。由于连接器的演变由复杂的和变化的协议来支配。因而C2只支持构件演变。由上面的例子可以看到,C2的拓扑使得构架结构有良好的可理解性、组装性和可伸缩性。

3 图象档案管理系统的构架化开发

3.1 概念化构架

档案的快速、完整收集和自动化归档、大容量存储和实时调档技术均是图象档案管理系统的关键点。针对其特殊需求,我们提出这样的管理模式:给每份函件打上条形码,用可连续扫描的高速扫描仪扫描转换成图象文件,识别出其中的条形码进行自动化归档。图象档案存储在大容量的存储设备中,其编目信息存储在数据库中。可供用户网上实时查询。图3给出了黄页号簿档案管理系统构架。

组成这个构架的构件被逻辑地分成4组。构架顶层的构件封装了整个系统的数据和状态信息。这些构件只响应请求,传送内部状态变化信息及数据。这里除了文件归档ADT(Abstract Data Type)是领域相关的之外,其余都是领域无关构件。

Control类构件根据系统流程控制询问设备状态和查询数据,解释获得的状态信息及数据,并根据Manager类构件

的指示改变状态,插入、更新或删除数据。例如文件Control控制图象文件的存取和删除。图象Control提供图象处理功能:图象增强、压缩和解压缩、条形码识别。这些构件都是与领域无关的。

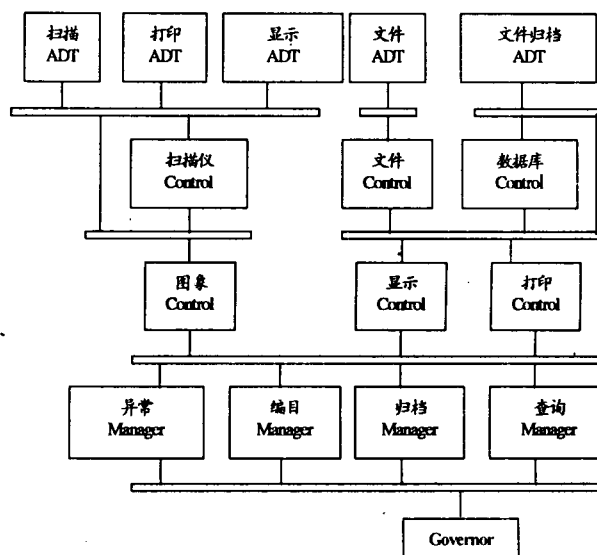


图3 黄页号簿档案管理系统构架

Manager类构件接收经过Control类构件解释过的状态信息及数据,作进一步的分析,然后把处理意见交给Control类构件。异常Manager处理识别不出条形码的异常情况,编目Manager对档案进行编目,归档Manager对档案进行归档,查询Manager对档案进行检索。这些构件都是领域相关的,与档案的编目规则有紧密联系。

Governor构件接收所有用户事件(如按下鼠标左键等),控制全局。它可以是档案管理领域中的通用构件,具有存档、调档和打印的一般功能。

构件之间相互作用是一种非常复杂的行为,连接器设计的好坏直接影响到构架的重用性和构件的可替换性。本构架中扫描仪Control和上层的接口符合图象输入设备的TWAIN界面标准,图象Control和上层的接口为常用的图象文件格式(Bmp, TIFF, JPEG, PCX, CALS等),Manager类构件与上层的接口则是图象档案文件和档案编目信息。

3.2 构架实例化

本系统采用Delphi Client/Server V2.0开发,后台数据库使用Oracle,扫描仪采用富士通的M3096EX/GX高速台式平板扫描仪,装有自动送纸器ADF,连续扫描速度为每分钟20多页,一次可送入50页材料。所有图象档案存储在大容量的光磁盘阵列中。

(1) 构件的选取和改进 扫描仪、图象、显示和打印Control构件分别选用KOFAX ImageControls V2.0 Gold版本提供一系列的VBX构件:Kscan、Kview、Kimgp和KPrnt。并相应作了一定的改进。以图象Control为例:在扫描中,往往出现连续几份回函的条形码是相同的,这是因为有的大公司或企业的刊登信息特别多,有若干张回函。于是设计连续图象构件把多页图象保存在同一文件中。对话构件是提供给上下两层的接口。图象Control(见图3)是多个构件的有机集成,本身就是一个小的框架。(下转第25页)

-9-

善,同步要求过高:主进程要在接受到全部的 n 个从进程的返回值后才发送下一组数据,这样会使较早完成计算的进程被迫耗费许多时间以等待下一轮计算数据的到来。该方案的测试结果分析为:

	1节点机	2节点机	4节点机	8节点机
主进程运行时间	2152	1074	626	432
主进程通信时间	2152	1074	624	430
从进程平均计算时间	2134	917	436.5	228.7
从进程平均通信时间	18	157	186	169.3
加速比	0.74	1.48	2.53	3.69
并行计算成本	2152	2148	2502	4512
运行效率	0.74	0.74	0.63	0.36

5 结论

本文考虑到文献[2]提出的算法简单易实现,而且对

(上接第9页)

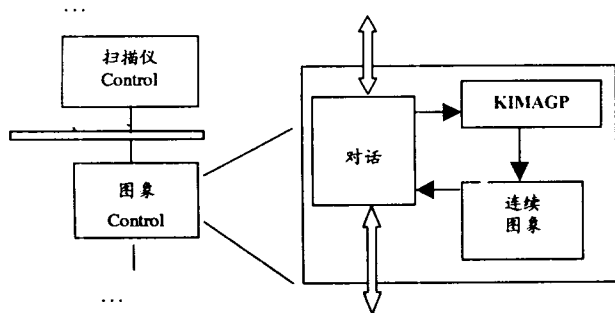


图4 具体的C2构件

(2) 用户构件的编写 与特定应用相关的构件主要是文件归档ADT和Manager类构件。这些构件由于跟档案编目和归档机制直接相关,因而设计好档案编目信息是编写这类构件的关键。一般来说,编目信息得包括档案唯一键和档案文件路径信息。在本领域中的编目信息如下:号簿信息、用户信息、图象文件路径信息。

3.3 构架的多样性和适应性

(1) 构件的多种语言的实现和集成 Kimgp构件有VC和VB两种版本,并可互相替换。图象Control构件中Kimgp构件采用VB版,而对话和连续图象构件用Delphi实现,通过VBX接口集成。

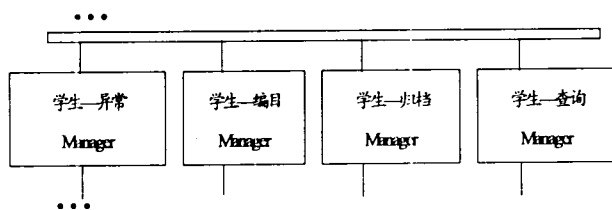


图5 考生图象档案管理系统构架

(2) 应用的变化 前面已经提到与特定应用相关的部分

任何一种二维图象的识别都可以不作太大的改动就可以直接应用,也不需要模板库中的样本作任何预先的处理(与神经网络算法相比),只是其运行速度太慢,易于并行实现,所以本文并没有对文献[2]提出的算法进行改进,只是使其中的循环并行化,目的是希望能从并行角度对离线手写汉字识别问题进行初试。最终得出结论是:由于汉字是大类别模式识别问题,无论如何改进算法,其时间复杂度都太高,要彻底想使手写识别字速度快、识别率高,最佳改进办法是设计一个好的算法并使之并行实现。

参考文献

- 1 Tang D, Gupta G. An Efficient Parallel Dynamic Programming Algorithm. Computers Math Apple, 1995, 30(8): 152
- 2 刘漂,白刚.脱机手写字识别的DP算法的设计,中国图象图形学报,1998,3(3): 58

主要集中在档案编目和归档机制这两个细节上。替换那几个相关构件,就可组装成不同的图象档案管理系统。例如高考学生档案中有不少获奖证书复印件之类的图象档案,可以把档案编目信息定义为:考生地区信息、考生信息、图象文件路径信息。

替换原构架的Manager类构件,见图5,就得到一个高考学生档案管理系统,可供招生人员实时快速查询,大大提高招生效率。

4 结束语

构件组合、替换和重用是大规模软件开发的关键。把软件构架看成具有层次结构的构件和连接器的组合,可以很好地解决这些问题。

本文提出了一个C2风格的图象档案管理系统构架,指导档案自动化管理领域中各种特定应用的开发,对软件的可升级性和构架级、设计级的重用有重要意义。根据此构架开发的黄页号簿档案管理系统,具有以下突出特点:(1)档案的搜集极为简单,齐全,快速而且可靠。每分钟可以扫描20多页回函。公司可以据此向客户提供各种服务,还能促进公司更加准确、高效地完成生产任务。这无疑大大地提高了公司的形象。(2)档案的建立自动化。系统使用条码输入技术,它的信息输入准确可靠,归档快速、高效。(3)所有的档案采用无失真的压缩存储,压缩比一般可达30:1,大大节省了存储空间。

参考文献

- 1 Medvidovic N, Oreizy P, Robbins J E, Taylor R N. Using Object-oriented Typing to Support Architectural Design in the C2 Style. In Proceedings of ACM SIGSOFT'96; Fourth Symposium on the Foundations of Software Engineering (FSE4), San Francisco, CA, 1996-10:24-32
- 2 Luckham DC, Kenney J J, Angustin L M, et al. Specification and Analysis of System Architecture Using Rapide. IEEE Transactions on Software Engineering, 1995, 12(4)
- 3 Monroe R T, Kompanek A, Melton P, Garlan D. Architectural Styles, Design Patterns, and Objects. IEEE Software, 1997-01