

基于掌上电脑的农田信息快速采集与处理系统的研究

方 慧, 何 勇

(浙江大学生物系统工程与食品科学学院, 杭州 310029)

摘 要: 为了解决精细农业中空间信息及属性信息的快速获取与分析处理问题, 提出并实现了基于掌上电脑的农田信息快速采集与实时处理系统。分析了具体实现系统功能的若干关键技术, 如地理信息系统(GIS)、空间索引、Windows CE 数据库技术、采样控制等因硬件环境的不同而产生的问题及相应的解决方案, 又针对农场环境下的特殊性, 提出并实现了一个面向对象的农场模型, 以便于更好的描述农场环境及更方便的确定采样方案。作为一种实时系统, 该系统可结合全球定位系统(GPS)和各种传感器, 快速地测量和分析农田信息。应用该系统对两种 GPS(Ag132 与 MAP330)进行了精度比较试验, 快速而方便地证实了 Ag132 有更好的精确性与稳定性, 更适于农业生产要求。试验的过程与结果也间接表明了该系统具有较好的实用性, 为精细农业田间信息快速采集和处理设备的研制与开发奠定了基础。

关键词: 农田信息采集; 掌上电脑; 实时处理; 存取方法; GIS; GPS

中图分类号: TP274⁺ 1.2; TP368.33

文献标识码: A

文章编号: 1002-6819(2004)06-0124-05

0 引言

准确可靠的农田信息是实施精细农业的基础。农田信息可分为两类: 一是农田的位置信息, 包括地块的边界信息, 采样点的经度、纬度, 地块中沟壑的位置等地理位置信息。只有这些位置信息确定了之后, 才能在相应的位置上进一步采集精细农业所需的另一类信息: 田间相应点的属性信息, 如, 水、有机质、氮、磷、钾等土壤属性信息及籽粒产量、小穗数、生物质产量等相关属性。传统的信息采集是收集样本到实验室化验, 但这些方法费钱费时费力。为开展精细农业的研究, 需要有适用于精细农业的农田信息快速采集技术。近年来, 传感器技术的发展使田间实时快速数据采集成为可能。而数据采集过程还有存储数据、控制及实时修改采样计划等的需要, 本文主要进行数据采集系统软件控制部分的研究。

由于是在田间采集信息, 台式机不适于作为软件控制的平台。笔记本电脑虽然相对轻便, 但存在价格较高、质量较大、供电时间不够长等缺陷。近年来, 掌上电脑(Palm sized PC, PPC)的发展十分迅速, 不但操作系统的功能日益增强, 相关的开发工具也越来越强大, 使得开发基于掌上电脑的农田信息采集控制系统成为可能。

目前, 基于掌上电脑的农业应用系统的研究还较少, 国内主要有农业专家系统方面的研究^[1,2], 而在基于掌上电脑的信息采集方面, 在国外, 一些公司已根据现有的精细农业的成果开发了初步的产品, 如 Trimble 公司开发的系列产品, 如 AgGPS170 Field Computer 是一种软硬件一体化的专用于农业的计算机平台, 它可应

用到大田作业中, 能通过 RS232 接口接收并发送 ASCII 码的传感器信息、GPS 位置信息, 并能按精细农业要求的作业方式, 记录下农机的运行轨迹及变量作业实施过程。另一种较好的 Trimble 公司的农用掌上电脑解决方案是基于 Windows CE 的 AgGPS EZmap 软件。AgGPS EZmap 提供了地图功能与简单的采样控制, 可完成如实时计算并显示农机当前施肥的覆盖面积、移除非种植面积、土壤采样点的确定等功能。单独的 AgGPS EZmap 在采样控制上主要是手动输入属性及条码的自动输入, 但上述技术和装备的主要的缺陷在于, 它们是为专用的目的而设计的, 因此, 缺少足够的灵活性, 只能按它现有功能操作, 无法增加新的方案, 功能受到限制。自行开发了一个功能强大, 通用性较好的信息处理系统。

本文首先描述了系统的框架与功能, 然后重点介绍系统实现中的关键技术, 最后借助该系统完成的 GPS 精度分析试验及试验结果。

1 系统的框架与功能描述

系统由采样控制、空间数据管理、属性数据库管理 3 个主要的功能模块组成。各大模块的功能如图 1 所示。其中, 串口通讯控制部分主要负责接收 GPS 数据、传感器数据, 并能手工输入无法从传感器测得的数据及其它属性信息, 还能根据采样流程来控制具体的采样数据的读取; 采样点的计算主要是根据采样范围, 计算采样点推荐位置; 地图浏览功能主要是放大、缩小、移动等地图操作; 空间属性查询与分析主要是某地块的面积、周长计算、线段的长度度量等空间属性分析。系统的总体实现框架图如图 2 所示。

2 若干关键技术的研究

2.1 地理信息系统的集成

本系统在本质上是一个专用于农田信息采集的小型专用 GIS 系统, 由于精细农业生产的特殊性, 它与当前其它专用的基于掌上电脑的 GIS 系统(如地籍测量用 GIS 及出租车导航用 GIS)有很大的不同。常用的地

收稿日期: 2003-12-11 修订日期: 2004-07-01

基金项目: 国家自然科学基金资助项目(30270773); 高等学校优秀青年教师教学科研奖励计划资助项目; 浙江省自然科学基金资助项目(301270); 浙江省自然科学基金人才基金资助项目(RC02067)

作者简介: 方 慧(1973-), 女, 浙江杭州人, 博士, 从事计算机与信息管理方面的研究。Email: xffh@citiz.net

通讯作者: 何 勇, 博士, 教授, 博士生导师, 杭州市凯旋路 268 号浙江大学生物系统工程与食品科学学院, 310029。Email: Yhe@zju.edu.cn

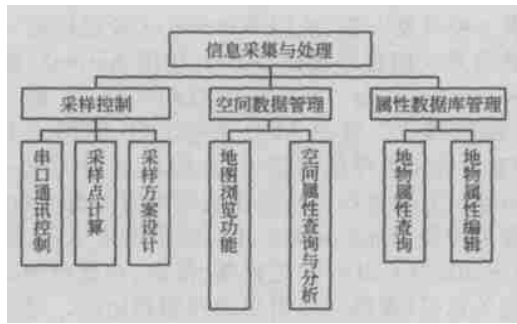


图 1 系统的基本功能模块框图

Fig 1 Basic functions of the system

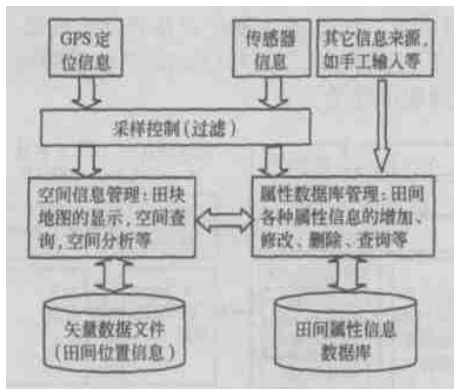


图 2 系统基本框架图

Fig 2 Architecture of the system

籍测量 GIS 往往只要采集来自 GPS 接收机的位置信息,再手工输入其它属性信息,而本系统除了位置信息,还需要自动采集来自传感器的农田属性信息,有些不能用传感器采集的信息必须实时或后期用手工输入。因此,为了满足农田信息采集的灵活性要求,系统必须设计多种数据采集方案,从而产生了较复杂采样控制流程。出租车用 GIS 主要用于导航,大部分的地理数据及属性数据都是预先存储在计算机中而不是动态生成的,在内存管理上会比本系统更简便一些。

2.2 面向对象的数据模型

本系统采用面向对象的方法来描述各种事物。用面向对象的方法,可把地物抽象为点、线、面等矢量,如本系统的最初采用的数据模型。

用总点集对象以链表形式来存储所有点的具体信息,坐标位置相同的点不能重复存储,其它具体的对象只是指向所在点;用点对象描述被抽象为点的地物,如移动的 GPS 信号源;用点集对象描述被抽象为点集的地物,整个集合是一个对象;用线对象描述被抽象为线的地物集合,其中一串坐标点组成的一条连续的线是线对象的一个“部分”;用多边形对象组成一个或若干个封闭区域,其中,每个封闭区域是多边形的一个“部分”;矩形:用于描述被抽象为矩形的地物,一种特殊的多边形。

然而在实际的设计过程中,笔者发现这种表示方法过于强调把地物抽象为点、线、面,而忽视了地物的特殊性及其地物之间的关系,使农场中地物的特殊性较难表达,各种采样方案较难合理表示。

为解决这一问题,本系统采用了设计模式方法。设计模式的就是对于一类重复出现的问题的一种可重用的解决方案,在软件工程中一个设计模式能解决一类软件设计问题。根据文献[10]中提出的各种模式,针对本系统存在的这一问题,本系统提出了一种基于对象的农场模型。该模型结构如图 3 所示。

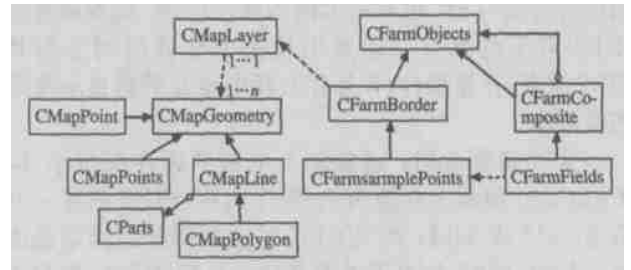


图 3 农场模型

Fig 3 Architecture of fam model

在这个结构中: CFamObjects 主要是一个公共接口类,为所有的子类向外部提供公共的接口。CFamBorders 是一个 CFamObjects 的子类,它是所有边界对象的抽象,任何一个地物都有边界,一个点状对象可以认为是一种特殊边界的对象。CFamComposite 是一个组合类,在这个类中有一个指向 CFamObjects 的指针列表,这样 CFamComposite 的对象可以是多个不同类型的边界对象对组合。CFamSamplePoint 是 CFamBorder 的子类,由于采样点有特殊的采样方案,所以,CFamSamplePoint 是专为采样点对象而设计的。它从 CFamBorder 继承,是一种特殊的边界对象。CFamFields 是 CFamComposite 的子类,代表了农场中的地块,除了 CFamComposite 的功能,它还能实例化 CFamSamplePoint、CMapGridLayer,其中 CMapGridLayer 是一个负责地块栅格数据的类,这表示一个农场地块是由一个采样点子层、栅格属性子层及大家都有的几何特征子层构成,即两个矢量图层及一个栅格图层叠置后组成一个地块。

这个农场模型实际上是组合方法与原型方法的结合。组合方法是将对象组合成树形结构以表示“部分-整体”的层次结构。在本系统的农场模型中,CFamObject、CFamBorder、CFamComposite、CFamSamplePoints、CFamField 是组合模型部分。原型方法主要用原型实例指定创建对象的种类,并且通过拷贝这些原型创建新的对象。它主要用于当要实例化的类是在运行时刻指定或当一个类的实例只能有几个不同状态组合中的一种时。CMapLayer、CMapGeometry 与 CMapPoint、CMapPoints、CMapLine、CMapPolygon 组成了原型部分,通过这种原型模式,CFamBorder 动态生成不同几何对象的层,以代表不同类型的边界。

2.3 数据的管理

2.3.1 空间索引结构

空间索引结构的性能在很大程度上决定了 GIS 的

大多数空间查询的性能。以前针对空间存储结构的研究都是在资源非常丰富的 PC 机或服务器这样的硬件环境下展开的, 大部分的研究基于这样的前提^[3-5, 12]: 总体资源比较丰富, 相对地, 内存有限, 外存很大, 外存的读取速度与内存相比很慢。为了加快索引, 要尽可能减少对外存的读取次数。因此, 以前的大部分多索引结构的研究都是针对所谓的外存索引展开的。而掌上电脑的硬件环境与 PC 机有很大的不同。与 PC 机或服务器相比, 掌上电脑与空间索引相关的硬件限制主要是 CPU 速度、存储量(内存外存)、耗电量及屏幕显示量等四大限制。

从存储量考虑: 通常掌上电脑的内存都很小, 以 iPAQ 3700 型掌上电脑为例, 程序内存与对象内存一共只有 32M 或 64M, 而且 CPU 对这两部分的读取速度是一样的。因此, 如果不考虑其它辅助存储设备, 而只用掌上电脑内部的硬件来实现整个系统, 只能用主存存储结构。在 PC 机上比较常用的主存索引的结构有: K-D 树、BSP 树、四叉树、Cell 等^[3-5, 11]。从耗电量及 CPU 的限制考虑, 掌上电脑耗电量的多少与 CPU 工作负荷有关, 如果程序的数据结构与算法设计得比较合理, 使 CPU 的计算量相对较少, 则耗电量也减少。在掌上电脑的环境下, 如 R 树^[13]、四叉树、Buddy 树^[14], 在不同的应用中(如点查询、范围查询、最邻近查询等)往往各有所长, 没有哪种结构有绝对的优势, 采用哪一种需根据系统实际应用来确定。从屏幕显示内容考虑: 与 PC 机相比, 掌上电脑的屏幕通常比较小, 如果仅从矢量地图的显示角度考虑, 则采用分层的形式来表示空间数最为合理。文献[6]提出了专门针对显示问题的适用于掌上电脑的矢量数据分块存储数据结构, 这种是一种非主存存储结构, 仅读取参与显示的数据, 因而其显示速度较快。笔者认为, 这种数据结构是建立在已被分级的矢量数据之上, 如果矢量数据全部是动态插入生成的, 则还须解决合理而有效的对动态数据进行自动分级, 因此, 这种结构更适用于静态矢量数据的表达。

根据以上的分析, 系统的所设计的空间索引方案为: 考虑以 Window CE 为操作系统的掌上电脑, 应用程序只有 32M 的虚拟内存, 可以通过使用内存映射文件或分配大于 2M 的虚拟内存, 这时由系统管理虚拟内存, 从而可以使用超过 32M 的虚拟内存。因此, 对于本系统来说, 当程序的数据量很小时, 可以只用主存索引, 所有的数据都读入程序内存, 并在程序内存中建立主存索引。当数据量比较大时, 可以在程序内存建立主存索引, 而数据文件保存在对象存储 RAM 上, 通过建立内存映射来读取, 这仍属于主存索引。当空间数据量非常大时, 则可用内存映射来实现空间索引, 而数据文件保存在 CF 卡等外存上, 这是类似于磁盘索引的空间索引方式。

2.3.2 数据库的选择与应用设计

本系统的开发是基于 Window CE 为操作系统的掌上电脑。在 Window CE 3.0 平台上使用数据库技术,

主要有 3 种开发方案: 使用 Windows CE 提供的一组独特的数据库应用程序接口 (API)、使用 ActiveX 数据对象 (ActiveX Data Objects, ADO) 接口或使用 OLEDBCE 接口。其中 ADO 是基于 OLEDB 之上的技术, 使客户应用程序能够使用 OLEDB 提供的统一的数据访问接口方法来存取与操作数据库服务器的数据。

如果使用 Windows CE 独有的数据库 API, 则只能使用 CEDB, 但 CEDB 缺乏窗体、报表、存储查询以及表之间的关系, 只提供对一组表的存储和访问。而如果用 ADOCE 或 OLEDBCE 则可使用 CEDB, 还可以管理 SQL Server CE 及其它由第三方提供者提供的各种数据库, 如: Sybase SQL Anywhere, Oracle Lite 等。ADO 与 OLEDB 相比, ADO 的主要优点是开发难度低, 但灵活性也较低, 且 ADOCE 与 OLEDBCE 的选择主要与所用的编译程序有关。

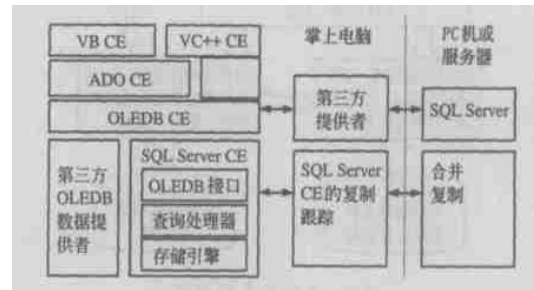


图 4 Windows CE 的数据存取结构

Fig. 4 Windows CE data access architecture

如图 4 所示, VBCE 不能直接调用 OLEDB, VBCE 必须通过 ADOCE 来调用 OLEDBCE 的功能, 而 VC++ CE 可以直接调用 ADOCE 或 OLEDBCE, 且用 OLEDBCE 可以保持较好的灵活性。由于没有现成的支持类, 用 OLEDBCE 时所需的支持类必须由用户自己开发, 或使用第三方开发的现成的类。

由于采用了面向对象的数据模型, 所以采用 C++ 语言是较好的选择。因此, 数据库接口选择 OLEDBCE 较适合, 而数据库则采用 SQL Server CE。

3 采样流程的设计

本系统根据精细农业的特点、农田信息采集的特点与可能性, 设计了以下 5 种采样流程: 边界点、手动事后、手动即时、全程自动、定点自动。其各自的定义为:

1) 边界点: 用于确定地物的边界, 如农田中的建筑、地块、池塘、水沟等。由于是边界点, 所以只应读取 GPS 数据, 且边界的确定可分为两种情况, 一种是直接沿着边界走, 一种是在边界的几个转折点上重复测量若干次, 取平均作为一个新的点。

2) 手动事后: 用于如果所测的信息目前还不能用传感器实测从田间获得, 必须在某点采样后用其它手段测得的情况, 传统的田间数据采集就是这一情况。这时的操作主要是找到某一点, 然后把一个记号写入数据库。

3) 手动即时: 用于测量无法从传感器获得, 而直接

从田间测得的一些信息, 需要用 GPS 定位后, 即时测量或观察, 并即时记录。

4) 全程自动: 用于类似于在运动的农机上采样的情况。所测的田间信息是从传感器传入袖珍计算机, 实现自动的 GPS 数据与传感器信息的读取、识别、配对并存储到数据库中。

5) 定点自动: 对某一地块中的若干点, 用 GPS 导向到相应的位置, 根据采样要求, 每点读取若干次传感器数据。与全程自动有些类似, 但这时还要加上一个求平均的处理。

除了以上 5 种控制采样的情况, 还有一种特殊情况就是仅观察 GPS 的点的运动, 也就是说, 只是在跟踪 GPS 点的信号。

整个控制过程用 2 个线程来完成: 主线程与串口线程。串口线程只负责从串口把信号读入, 如果是 GPS 信号, 并检查信号的完整性, 把正确的信号写入缓冲区, 并通知主线程。而主线程在得到有信号到达缓冲区时, 则一方面, 把信号传给跟踪层, 由它来更新跟踪的数据, 另一方面, 把消息传递给当前活动层, 由当前活动层根据自己的采样方案自动做出正确的响应。比如要采集传感器信息时, 首先选择当前采样方案为方案 4 或方案 5, 此时当前的活动层必然是采样点层, 采样点层是前面提到过的 CFam SamplePoint 类的一个实现, 具体的采样是通过该类中的 DoSample 函数根据方案 4 或方案 5 的具体参数完成的。

4 系统的实现与测试

笔者已经设计完成了基于掌上电脑的农田信息快速采集处理系统, 系统平台采用 PocketPC 2002, 开发工具为 Microsoft Embedded Visual C++ 3.0, 试验用的掌上电脑是 COMPAQ 公司生产的 H3700 型。系统界面如图 5 所示。

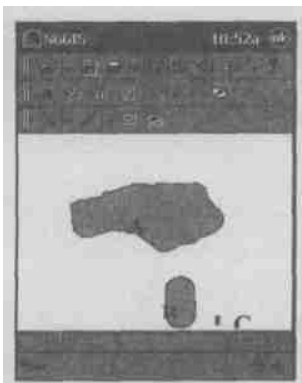


图 5 系统实地测量和计算结果图示

Fig 5 Test results produced by the system

图 5 是用掌上电脑测量的 3 个地块的边界, 是用本系统进行的 1 次试验结果图示。试验中, 使用了两个 GPS: Magellan 公司的 MAP330 与 Trimble 公司的 Ag132。其中 MAP330 GPS 精度比较低, 不能接收差分信号, 但体积很小, 便于携带, 价格低, 比较适于野外作业等, 而 Ag132 有较高的精度, 它将 GPS 接收机、信标

差分接收机、卫星差分接收机集成在一起, 其差分定位精度为亚米级。图 5 中的 A、B 是用 Ag132 测试两个地块的结果, C 是换用 MAP330 后测的另一个小地块的结果。表 1 是用本系统计算 A、B、C 三地块的面积与周长值(表中的 GK 表示采用高斯-克吕格投影)。

表 1 用本系统计算的结果列表

地块	面积(GK)/m ²	周长(GK)/m
A	58504.856	1075.655
B	10751.630	404.266
C	261.482	68.574

表 2 是用这个系统进行 GPS 精度的比较测试的结果。在这个试验中, 用两个 GPS 分别测了面积为 3630.537 m² (54.9 m × 66.13 m) 的长方形地块, 周长为 242.06 m, 从表 2 可见, Ag132 的精度大大高于 MAP330, 其面积误差 1.30% 左右, 周长误差只有 0.29%, 能满足农业测量的要求。

表 2 AG132 和 MAP330 的测量结果

GPS 型号	面积(GK)/m ²	与实际的面 积偏差/%	周长(GK)/m	与实际的面 积偏差/%
MAP330	3339.594	8.71	236.866	2.14
AG132	3677.842	1.30	242.752	0.29

5 结 语

提出并实现了用掌上电脑进行农田信息的采集、控制与管理系统的, 分析了该系统所涉及到的若干关键技术如 GIS、数据模型、空间索引、Windows CE 的数据库技术、采样控制, 并针对每一种技术在掌上电脑环境下产生新问题提出了一些方案与建议。在系统测试试验中, 结合 GPS 和各种传感器快速地测量和分析农田信息, 有较好的实用性。本系统的研究与开发为精细农业田间信息快速采集和处理设备的研制与集成化信息采集平台的开发提供了基础。

[参 考 文 献]

- [1] 王东辉, 涂云华. 基于 Windows CE 的 HPC/PDA 农业专家系统开发平台的研究与开发[J]. 高技术通讯, 2000, 10: 28-31.
- [2] 陈立平, 王东辉, 赵春江, 等. 掌上电脑农业专家系统开发平台的研究与开发[J]. 农业工程学报, 2002, 18(3): 142-145.
- [3] 顾 军, 吴长彬. 常用空间索引技术的分析[J]. 微型电脑应用, 2001, 17(12): 40-42.
- [4] 刘 东, 李 琦, 承继成. 主存空间对象的索引方法[J]. 环境遥感, 1996, 11(4): 302-308.
- [5] 陈 菲, 秦小麟. 空间索引的研究[J]. 计算机科学, 2001, 28(12): 59-62.
- [6] 李鲁群, 李成名, 林宗坚. 服务于 PDA 的矢量数据分块存储数据结构的研究[J]. 测绘学报, 2002, 31(2): 170-174.
- [7] 何 勇, 方 慧, 冯 雷. 基于 GPS 和 GIS 的精细农业信息处理系统研究[J]. 农业工程学报, 2002, 18(1): 145-

- 149
- [8] 姜山, 程君实 Windows CE 的实时性分析[J]. 测控技术, 2000, 19(1): 62- 64
- [9] 肖乐斌, 钟耳顺, 刘纪远 面向对象整体 GIS 数据模型的设计与实现[J]. 地理研究, 2002, 21(1): 34- 44
- [10] Erich Gamma, Richard Helm, Ralph Johnson, et al Design patterns Elements of Resuable Object-Oriented Software[M]. 1995. Addison Wesley Longman, Inc
- [11] Ning An, Anand Sivasubramaniam, Narayanan Vijaykrishnan, et al Analyzing energy behavior of spatial access methods for memory-resident data[J]. The VLDB Journal, 2000 411- 420
- [12] Sam et H. Spatial data structures[M]. Modern Database Systems ACM Press 1995, 361- 385
- [13] Gutman A. R-trees: a dynamic index structure for spatial searching [A]. In: Proceedings of the ACM-SIGMOD Conference[C], 1984, 47- 57
- [14] Seeger B, Kriegel H P. The buddy-tree: an efficient and robust access method for spatial data base systems[A]. In: Proceedings of the VLDB [C], 1990, 590- 601

Field information collection and processing system based on palm-sized personal computer

Fang Hui, He Yong

(College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China)

Abstract To quickly and effectively acquire and analyze field information, a Palm-sized Personal Computer (PPC) based field information fast collection and real-time processing system was designed. Compared with desktop computer, the operation system of this system is noticeably different in many aspects of system development and design, which is brought by windows CE OS. So the difference which was generated by the key technologies of the system, due to different hardware environments, was analyzed and debated. These key technologies include geographic information system (GIS), spatial access methods, database technology on windows CE, sampling control plan, etc. Based on the particularity of farms, an object-oriented farm model was developed to describe farm objects properly, to edit sampling plan conveniently and to conduct sampling control easily. As a kind of real-time system, the system can be used with global positioning system (GPS) and kinds of sensors to acquire and analyze field information instantly. The precision test of two GPS systems (AG132 and MAP330) was finished by the system. The result of experiment validated higher accuracy and stability of AG132 and more suitable to be used in precision farming. The experiment process and result also indicated good practicality of the system. The research and development of the system provided key technology for the development of equipment for precision farming field information fast collection and real-time processing.

Key words: field information collection; palm-sized PC; real-time processing; access method; GIS; GPS