Intelligent Interaction Room

技术可行性分析

2011-10 writer:钟王偲 甄羿

目录

1.整体技术架构分析………………………………………………………………..2

2.开发环境……………………………………………………………………………....2

3.具体分析

GUI层……………………………………………………………………………..…..3

图形识别

语音识别

输入数据抽象

输出

网络层………………………………………………………………………….……..4

加密与解密包

Web层面浏览解析包

网络安全包

网络组织包

内核层…………………………………………………………………………….…..5

I/O内核

主DNA内核

集成DNA内核

通信内核

必要存储内核

硬件层………………………………………………………………………….……..6

4.参考文献…………………………………………………………………………...…..7

5.法律支持与运营体制…………………………………………………………………….8

**整体技术架构分析：**

硬件层与网络层

GUI用户层

内核层（系统层）

基本层次

IIRoom系统本身含有3个层次：GUI层、内核层、网络层。同时必须要硬件层辅助（硬件层非重点内容）。IIRoom包含特殊的文件系统，以及特殊的应用程序组织（有DNA特征）。其中实现的主要难点在于算法和数据结构的选择，必要时必须自主开发新的算法，现有算法均是成熟的可行的算法 。本系统的智能主要体现在主DNA内核和集成DNA内核，这两个内核以及客户-服务端的模式构成了IIRoom的主体。层次中具体架构图式如下（虚线表示一定条件下非必需）：

加密与解密包

网络安全包

集成DNA内核

外部物联网/局域网

输出设备

图形识别

语音识别

I/O内核

主DNA内核

必要存储内核

WEB层面浏览解析包

通信内核

**开发环境：**

1. 编程语言和数据库：

C/C++ PHP/MySQL/JAVA 汇编；

1. 硬件环境与系统环境：

**系统环境**：FreeBSD/Ubuntu Linux/Red Hat(服务器版)/Mac(放弃Windows环境，需要Mac强大的图形处理能力)

**CPU**: Intel Xeon 6核 或 AMD同等级别处理器。

**RAM**:32GB DDR5

**GPU**:FX3800级别以上

**存储**：1TB(不需要太大容量，要求速度)15000 r/s以上传统硬盘，或者基于DRAM芯片的SSD。

**带宽**：10Gbp/s(LAN)

\*用户使用时的配置低于开发环境的配置，配置要求主要在全息投影和体感设备上。

**下面我们对所涉及的技术与算法做具体分析：**

1. GUI:

GUI

视觉

听觉

感应操作

内核

1. 图形识别（动作识别）：

**使用微软kinect作为我们的识别设备，下面是我们假想的一些算法：**

输入摄像头端口具备红外感应，可以扫描到红外视图，然后识别时执行以下几步：

1. 扫描筛图，把红外感应的图层从整体的影像中筛选出来，可能涉及到的算法有离散FFT.
2. 撒点，已被扫描的图像进行整体抽象，依据精度把图形抽象为给定坐标中的点。必须为离散型方便计算机处理。
3. 凸包围图：如果是2D的平面，使用Graham扫描法，对于3D空间可以扩展立体的凸包。此种扫描主要使用叉乘，核心是判断两个线段的相对位置。需要一开始找到原点。
4. 近似匹配：利用近似搜索把动作做成离散的分解，由以上几步，模式匹配后近似搜索动作组合（可能依靠k-d树），并最终抽象成真正的指令。模式匹配可能用到的算法有KMP（基本），AC-自动机，后缀树或者后缀数组。

**技术难点**：模式匹配，凸包，近似搜索，K-d树的实现。

1. 语音识别：

语音识别负责识别用户的语音，并做出指令，识别后的处理放式与图形识别非常类似，具体如下：

1. 录入语音，整合波形，抽象成离散的点，点的密度视识别精度而定，之所以抽象是为了匹配时方便处理，计算机仅可处理离散的数据，与图形识别撒点相同。
2. 模式匹配类似的语音指令，涉及到的算法有：KMP，AC-自动机，后缀数组等。

**技术难点**：模式匹配。

1. 输入数据抽象：

**我们仍然在这个步骤使用kinect:**

语音识别后的数据和图形识别后的数据或者直接输入的数据等统称为输入数据，输入数据分段为各个数据组，数据组再次抽象为斐波那契堆，堆是一种优先队列，可以很好的满足我们的处理需求。而斐波那契堆具有相当好的特点：删除是O(logN)，

除了删除，其他的插入和查找的时间复杂度理论为O(1)的，具体实现时可以使用Pair-Heap，Pair-Heap是基于斐波那契堆的一种实用性实现，具体实现时效果优于斐波那契堆。

**技术难点**：配对堆（Pair-Heap）。

1. 输出：

使用裸眼3D技术。（视差障壁技术或更为先进的MLD技术）

使用眼镜屏幕隔空操作（具体参见麻省理工学校学生普拉纳夫发明第六感装置）

输出时将信息流译码，然后交与硬件层，如全息投影设备等进行输出。本部分不涉及创新内容，与当先的技术没有太大区别，无非是全息投影。

**技术难点**：涉及硬件，有全息投影等输出设备即可。

1. 网络层：

网络层在IIRoom中具体指网络套件(与TCP/IP协议中的网络层意义不同)，是具有具体网络功能的一些套件：包括加密与解密包，WEB层面浏览解析包，网络安全包，网络组织包这4个基本包，其他包功能可以由主DNA内核在基本包基础上由用户自行进行重组。包中的内容含在各个内核中。

具体包功能技术可行性分析：

1. **加密与解密包：**

结合**RSA密钥**，**MD5**，**SHA1**等加密技术对“脑波”加密、做数字签名等，加密后的流称作“密脑波”与“脑波签名”、非加密的脑波叫“原脑波”，解密后的流叫“失密脑波”。

流中的数据先被分块，被切割成有树状结构的数据组，数据组再被加密，并被赋予随机分阶的权，然后传输。解密时，必须按照数据组的树状结构的顺序解密，否则将失败，因此“脑波”具有很高的安全性。而且支持流的随机化算法加密，即同样的信息可以被不同流程的加密传送，再由公钥给出指示，做具体解密。

失密脑波与原脑波能表示完全一样的信息，但是可能具有不同的结构，解密时按照树状结构的顺序，查看随机分阶的权，在不破坏信息流的情况下被重组为等价的树结构，能保证不破坏原信息的指标就是先前的随机分阶的权，但是如果结构特殊，可能只有唯一的树满足，那么失密脑波与原脑波就会完全一致。

**技术难点**：各种加密、Hash和随机化算法。

1. **WEB层面浏览解析包：**

我们已经具备的web2.0技术和基于http与ftp等相关协议传输的信息，均可以直接整合为“脑波”，系统接收到这种脑波后，对于脑波先用加解密包做出解密，然后将这种信息流先经由主DNA内核处理整合，最后输出到GUI层。

**技术难点**：暂无。

1. **网络安全包：**

本包直接采取Linux/Unix式安全方式，这是由系统本身的机制所决定的，\*nix式的安全机制使得无需安装任何杀毒工具，因此对于用户权限的把控IIRoom非常严格。我们采取软防火，这个模块可以使用开源的（GNU协议，法律支持）Linux/Unix机制，而Linux/Unix的机制也是相当成熟的。但是我们并不是架构在\*nix体系上的东西，网络安全包将封闭化，这是保证安全的一个必要渠道。

**技术难点**：与\*nix安全机制的整合。

1. **网络组织包：**

这是网络层与主DNA直接嫁接的包，他隶属于通信内核。在超级服务端可以自由组织私有类神经的网络信道（用以共享DNA）。网络组织也是完成加密会议的必要，网络网络通路需要检查组织的可行性和最短链路。实现时，对于可行性采用网络流算法，因为网络可以被抽象成隐式图，所以使用HLPP预流-推进算法和匈牙利--二分图算法，这两种算法主要做匹配以快速检查可行性。对于最短链路使用D\*算法，D\*是A\*的强化，一种启发式搜索算法，组织式这两步都非常必要。

**技术难点**：网络流，二分图，D\*.

1. 内核层：

I/O 内核

***主DNA内核***

必要存储核

通信内核

集成DNA内核

IIRoom内核体系架构

1. I/O内核：

I/O内核要对输入的抽象数据与主DNA内核进行交互，首先要分析堆的内容，然后依次弹出到主DNA内核层，通讯内核层等。而接收到主DNA内核层的指示与信息流时，他会依照指令与具体的数据类型进行最适当的输出，是GUI层的直属功能内核。

**技术难点**：对于堆的信息分析。

1. 主DNA 内核：

这是IIRoom的“首脑”，首先他作为整个系统的“领导”，负责最高权限的指令下达，直接控制其他任何内核，但是不能垮层控制。对于任何指令其内嵌堆、栈结构，拥有检索模块，可以直接对存储内核检索。并支持超精度计算。

**DNA的定义**：DNA是指IIRoom的文件制式和程序组织，IIRoom的文件结构不再是线性的，而是类似超文本同时又有独立的拓扑结构的。程序的组织中每个子段均有特征DNA,特征DNA的构造又主DNA内核采用Hash完成，同时不同类型的文件和程序满足相统一的加密标准，并流化。因此程序可以自由依照DNA重组。首先按照特定的Hash抽取特征DNA.然后将文件结构信息与DNA整合，加密DNA并被组织到文件中，文件按照DNA的结构组织顺序。解密时先解密DNA，获取DNA中的信息读取出文件的结构顺序，接着对于程序或文件进行执行即可。

**DNA重组**：这是由用户自选的重组，首先分析至少2个软件的对等模块是否存在并反馈，如果存在，主DNA内核先解密DNA，更改并交换DNA，加密DNA。这样新的软件便被组织完成。

**DNA自重组**：这是IIRoom自行的智能行为，它本身具备一定的智能性。每当用户运行服务或者软件时，都有一个指标测试对于同类服务进行比较，如若这次的服务优于之前的则自行进行改良，改良方式是神经网络算法。而对于IIRoom本身对用户的操作反馈和指令执行的能力，用户可以直接反馈自己对IIRoom的理解指令的意见，IIRoom也能自行发觉，发觉后改良方式是遗传算法，这样再结合神经网络能做到定向进化，是系统的智能性自我提高。

**技术难点：**遗传算法，神经网络，Hash。

1. 集成DNA内核：

集成DNA内核就像一个DNA复制工厂，但是与主DNA内核不同的是，他可以并有权限去整合并记录很多个“DNA流”，并使用类似“病毒”的超速进化方式，加速进化并大规模淘汰，这种方式与遗传算法不同，是非定向不可控的指数级进化，因此“淘汰”至关重要，而且对于DNA以基因突变为主：先随机一个子代，然后做淘汰检验，这样就加速了“进化”，培养出更智能的超级服务端的人机交互功能。

**技术难点**：类似于加入随机化的遗传算法。

1. 通信内核：

通信内核由很多功能包间接的组成一个通信子系统，而通信内核就是直接控制各个功能包，同时它被主DNA内核直接控制，并无选择的执行主DNA内核的各种指令，指令被接受后，通信内核就并行的控制和协调各个功能包，因此它的主要算法是并行计算和懒惰计算。并行计算是对于网络端子的协调，懒惰计算则是协调问题的主要手段。

**技术难点**：并行计算，懒惰计算。

1. 必要存储内核：

IIRoom的客户端存储不再是重点，因此本内核主要用于服务端。内部组织采用仿生学原理，模仿人脑中枢神经的类神经的网络，与自身主内核的DNA匹配，采用动态自增自删减的神经元存储模式。用“立体”数据模式存储，相比传统介质有海量提升。存储时数据结构使用B树，AVL和线段树。B树与AVL决定存储的结构，而线段树对神经元做了抽象，专门存储一段“神经区域”的地址。这样，可以很高效的存取数据，时间复杂度对于各个操作均为O(logN).

**技术难度：**AVL，B树，线段树。

1. \*硬件层(Not important)：

kinect:主要的输入设备。抛弃键盘鼠标等传统输入设备。

**＊全息投影仪：**它们将被安装在空间区域里的各个角落，每一个投影仪都通过若干个投影灯泡进行照射，使光线发生干涉或衍射，最终生成3D立体影像。每一个投影仪都被系统同步并接受来自于系统的最新动态数据，投影出最实时的动态场景。理论上讲，全息投影仪的空间分布设计、个数与质量、投影仪上的投影灯泡个数与质量（颜色范围和照射路程）和真实空间（包括被虚拟的空间）的情况会影响投影质量，但是GUI层会智能调节至用户希望的最佳画质。这个硬件支持第三方产家的研发和生产，用户可有多种选择。

**＊全息声音收录及虚拟设备：**可以与全息投影仪合为一体进行全息声音收录和虚拟，也提供用户机会取购买第三方生产的具有更强大功能的全息声音收录及虚拟设备，以自我设计更加合理的位置，使收录和虚拟效果更加舒适。它们将被安装在空间的各个方位形成比一般单方位设备更加立体环绕效果。这个硬件支持第三方产家的研发和生产，有利于把全息声音功能发挥得淋漓尽致。

**超级计算机：**作为区域服务端和总超级服务端的服务器，有强大的计算功能和海量存储功能，能对内核文件进行高密度智能管理和更新，并与网络全时段同步，进行高流量的输入与输出操作。

**带宽：**10G/bps的基本要求到100G/bps或以上的推荐要求，便可满足对各种数据和“脑波”的传输，大大降低网络延迟值，使得同步实时场景时流畅无阻，清晰。

**体感接收设备：**首先，全息摄影仪传输的信息流到了系统内核中是会进行一次分辨并把人体动作信息流感应出来，把它从总的信息流分离出来做动作分析并进行反馈到GUI上，这是最基本的体感接收。而体感接收设备可由第三方开发和生产，以适应客户的各种不同需求，如大型全息游戏，残疾人辅助体感设备，专业人士专用体感设备等，只要在保证设备与系统同步良好的状况下，几乎可以完成任何的体感接收。

**\*技术难点**：以上难点均为物理层面，我们不过多分析，仅知目前都可行，并已有相关技术即可。

参考文献：

1. 《算法导论》 作者： （美）科曼（Cormen,T.H.） 等著
2. 《数据结构与算法分析(C++描述)》 作者：(美)(MarkAllenWeiss)维斯
3. 《操作系统:精髓与设计原理》 [斯托林斯(William Stallings)](http://www.amazon.cn/s?ie=UTF8&search-alias=books&field-author=%E6%96%AF%E6%89%98%E6%9E%97%E6%96%AF%28William%20Stallings%29) (作者)
4. 《计算机组成及汇编语言原理》 [卓拉(Patrick Juola)](http://www.amazon.cn/s?ie=UTF8&search-alias=books&field-author=%E5%8D%93%E6%8B%89%28Patrick%20Juola%29) (作者)
5. 《计算机网络与因特网》 [Douglas E.Comer](http://www.amazon.cn/s?ie=UTF8&search-alias=books&field-author=Douglas%20E.Comer) (作者)
6. 《算法:C语言实现(第5部分)图算法》 [塞奇威克(Robert Sedgewick)](http://www.amazon.cn/s?ie=UTF8&search-alias=books&field-author=%E5%A1%9E%E5%A5%87%E5%A8%81%E5%85%8B%28Robert%20Sedgewick%29) (作者)
7. 《计算机程序设计艺术（1-3卷）》 作者：Donald E. Knuth
8. 计算机图形学相关书籍
9. GNU协议

法律支持与运营体制：

本软件完全原创，符合当地法律标准与GNU协议，技术完全自主开创，没有任何侵权行为。IIRoom版权及最终解释权归5th Planet所有，侵权必究。

运营时，IIRoom的客服端完全免费，但不开源，不可复制，我方保留版权。用户只需要购买硬件设备即可。收入来源依靠广告和第三方，并掌管超级服务端等。用户对于IIRoom无需支付升级等任何费用，同时第三方也有免费和收费服务及软件供用户选择。（5th Planet）