

人机交互

Human-Computer Interaction

夏榆滨
北航计算机学院

第五讲 交互技术

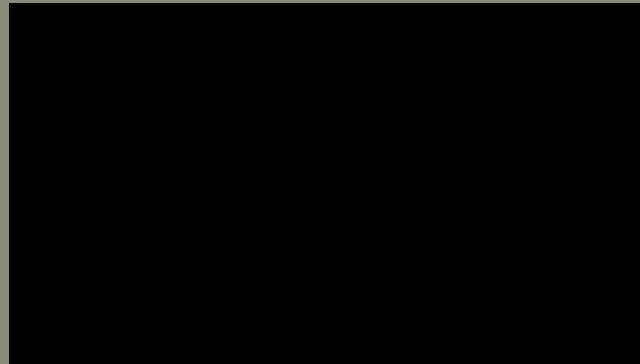
问题

- 交互 (interaction) ?
- 交互范式 (paradigms:
frameworks/model/typical example) ?
- 交互技术?

1. 交互?

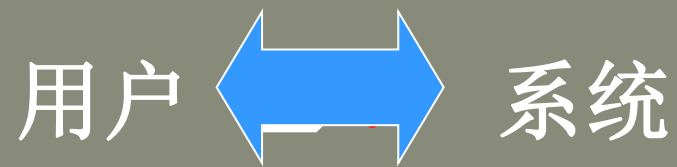
1. 交互

理想的交互



1. 交互

通信



1. 交互

- 交互模型
 - 用户与系统间的信息交换过程
- 人机工程
 - 人机交互中的物理特征
- 交互样式
 - 人机对话的形式方法
- 交互背景
 - 社会, 组织, 动机

1. 交互

- Donald Norman模型

1. 交互

● Donald Norman模型

➤ 七个阶段：

- 用户建立其目标：例如，创建一个红色的三维实体
- 任务规划：
- 定义界面操作：例如，具体的画实体动作
- 执行操作：
- 感知系统状态：
- 解释系统状态：
- 面向目标评估系统状态：

- 唐纳德·A·诺曼（Donald A. Norman）：《设计心理学》、《情感化设计》、《未来产品的设计》

1. 交互

● Donald Norman模型

人机交互系统的设计缺陷，例如：

执行中的问题：

用户规划的操作

≠

系统允许的操作

评估中的问题

用户所期望的系统状态改变

≠

实际展现的状态

1. 交互

● Donald Norman模型

➤ 人的问题：不准确或错误

✓ 不准确：

- 对系统和对任务目标的理解
- 对操作过程的执行
- 某个操作的执行

✓ 错误

- 直接的任务目标错误

✓ 改进？

- 不准确 – 好的界面设计
- 错误 – 对系统的正确理解

1. 交互

- 人机工程

1. 交互

● 人机工程

- 研究人机交互中的物理特征
- 强调“人的因素”
- 定义系统设计的标准与/或设计规则

1. 交互

● 人机工程

➤ 控件与显示:

- 例如，以功能、频度或动作顺序进行分组

➤ 环境

- 例如，座位的设计应适合绝大多数的人

➤ 健康要素

- 例如，光强、噪声等

➤ 颜色的使用

- 例如，警示用红色，正常用绿色，兼顾色弱者

➤ 认知要素:

1. 交互

- 交互样式

1. 交互

● 交互样式

- 命令行方式: **Unix, DOS/cmd**
 - 菜单方式
 - 自然语言: 语音识别
 - 问答与查询对话: **SQL**
 - 填表/电子表格方式: **Windows Form/ Lotus 1-2-3/MS Excel**
 - **WIMP**: 窗口**W**/图标**I**/菜单**M**/指点**P**, **Windows, Motif**
 - 指点与单击: **hypertext/Web Browser/Multimedia**
 - 三维交互界面: 虚拟现实, 三维按钮
 - 大脑与电脑直接交互? ...
 - **AI+?...**
- 每一种样式都会基于某种交互**范式**(交互原理框架)实现

1. 交互

● 交互样式

➤ 交互性: **interactivity**

通常重视外观、功能

还要重视全面的心理感受

外观 + 行为 → 看起来 + 感觉

人机 “和谐” ----- “共振”

1. 交互

- 交互背景

1. 交互

● 交互背景

➤ 对交互的影响因素

✓ 社会/组织：其他人的影响

- 希望表现, 竞争, 担心失误

✓ 动机

- 害怕, 守规矩, 抱负, 成就感

✓ 缺陷系统

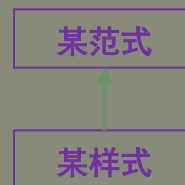
- 导致挫折感, 缺乏动机的驱动

2.交互范式

2. 交互范式

● 什么是范式(paradigms)?

- 是一种处于支配地位的理论框架或世界观: **Predominant theoretical frameworks or scientific world views**
 - 例如, 亚里士多德哲学(Aristotelian), 牛顿学(Newtonian), 爱因斯坦理论(Einsteinian, relativistic))都是物理学中的范式
- 学习HCI的发展过程, 主要就是去理解一系列的范式演化
 - 候选的“范式”
 - HCI发展过程将最终决定多数专家所“公认”的“范式”。



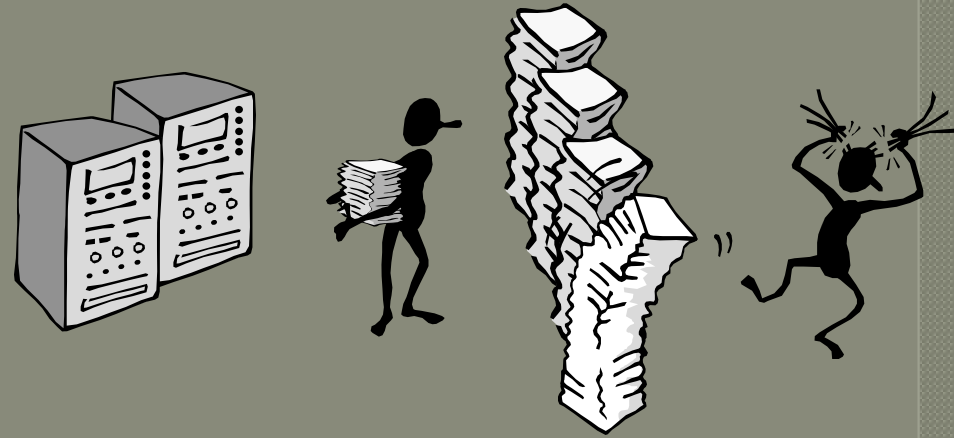
2. 交互范式

● 为什么要研究范式？

- 关心HCI的可用性：共性特点
 - 可用性开发方法：usability
 - 评估可用性
- 交互系统设计与发展
 - 可用性设计范式：继承与创新

2. 交互范式

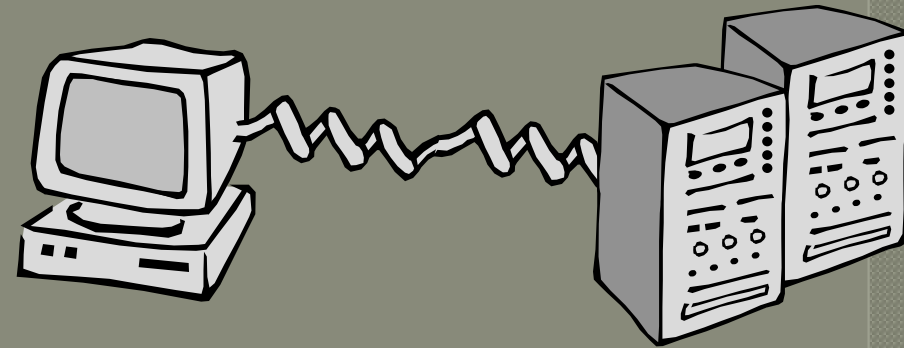
- 批处理: Batch processing



冷酷的计算

2. 交互范式

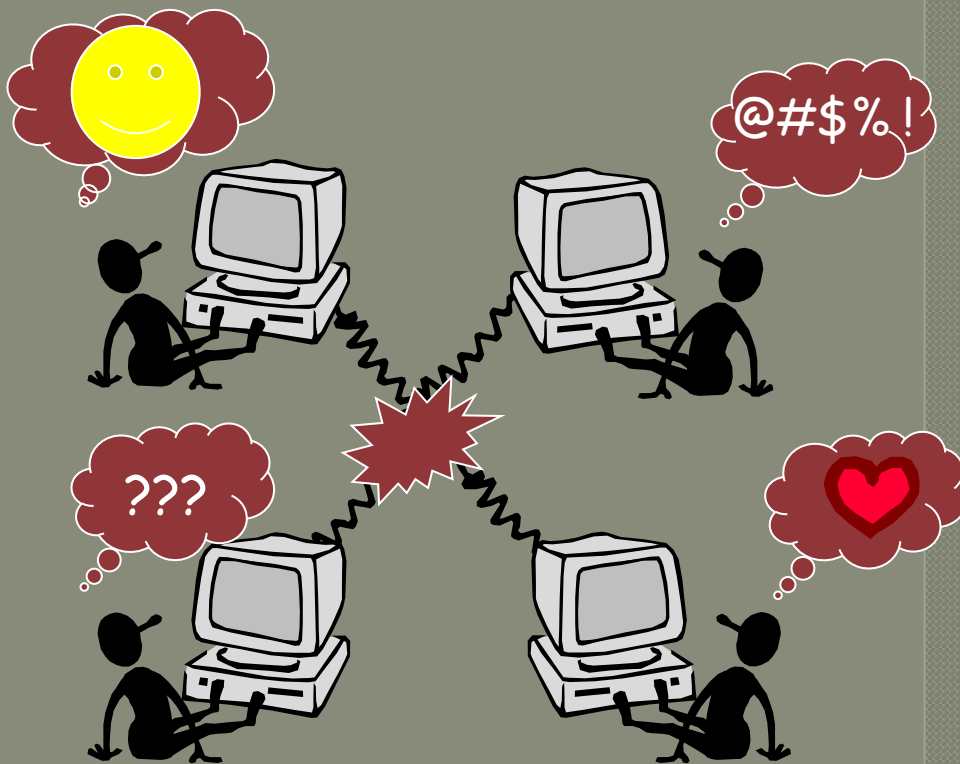
- 批处理
- 分时处理



交互式计算

2. 交互范式

- 批处理
- 分时处理
- **网络处理**



社区计算

2. 交互范式

- 批处理
- 分时处理
- 网络处理
- **图形显示**



直接处理

2. 交互范式

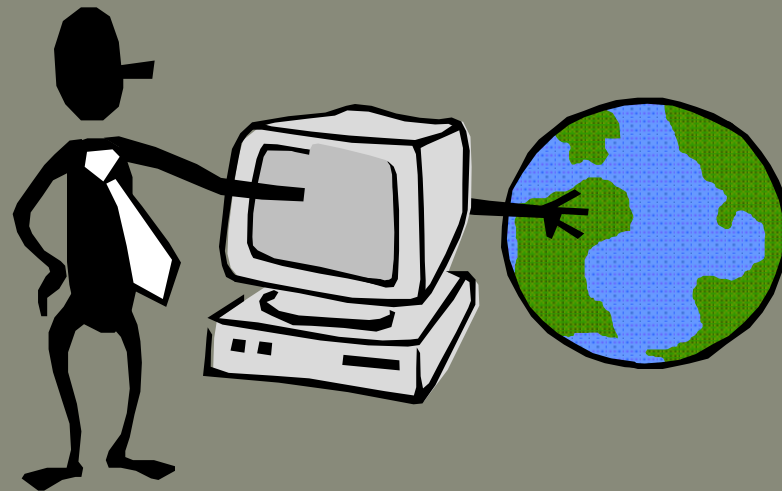
- 批处理
- 分时处理
- 网络处理
- 图形显示
- **微处理器**



个人计算

2. 交互范式

- 批处理
- 分时处理
- 网络处理
- 图形显示
- 微处理器
- **WWW**

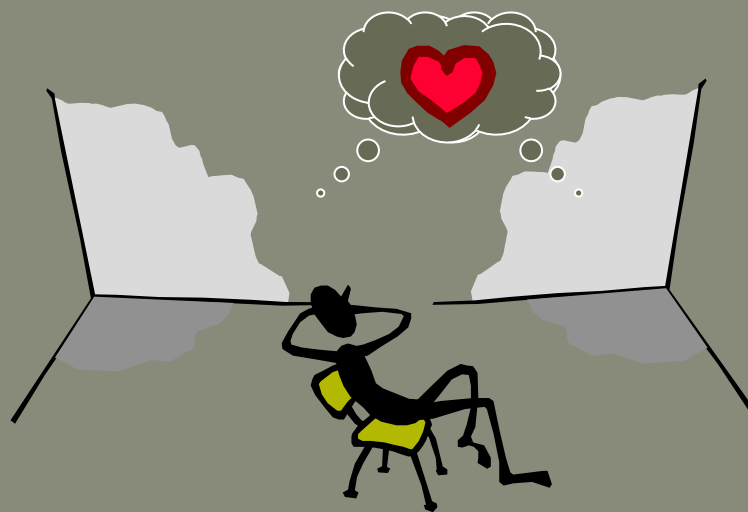


全球化信息

2. 交互范式

- 批处理
- 分时处理
- 网络处理
- 图形显示
- 微处理器
- WWW
- 普适计算、IoT
- AI+...?

物理世界与电子世界在现实生活中实现服务的融合.



3.交互技术

3. 交互技术

- 输入模式
- 基本技术
- 图形与三维交互
- 自然交互、AI+

3. 交互技术

- 输入模式

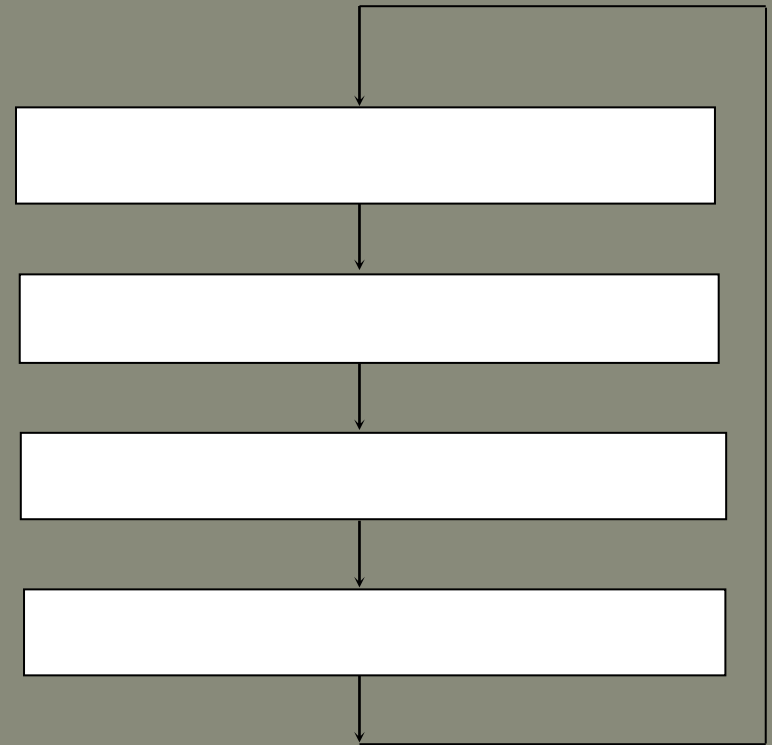
3. 交互技术

● 输入模式 (较底层)

- 请求模式(Request Mode)
 - 采样模式 (Sample Mode)
 - 事件模式 (Event Mode)
-
- 输入/程序：串行、并行

3. 交互技术

- 请求模式
 - 应用程序请求
 - 等待输入
 - 输入数据后继续执行

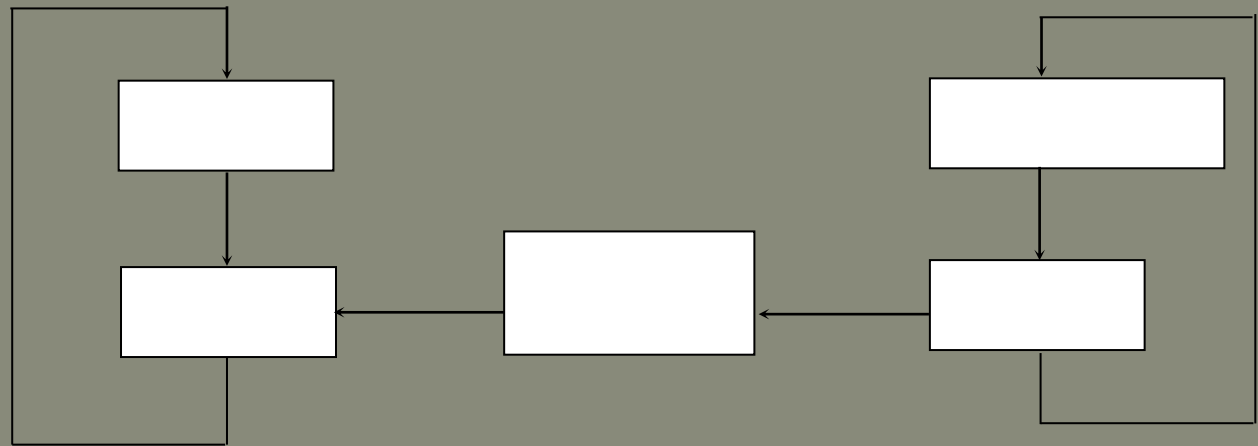


请求模式的工作过程

3. 交互技术

- 采样模式

- 输入设备连续不断地输入信息
- 应用程序遇到取样命令时，读取当前保存的输入数据
- 优点：方便，可同时处理多个输入设备
- 缺点：应用程序的处理时间较长时，可能会失掉某些输入信息

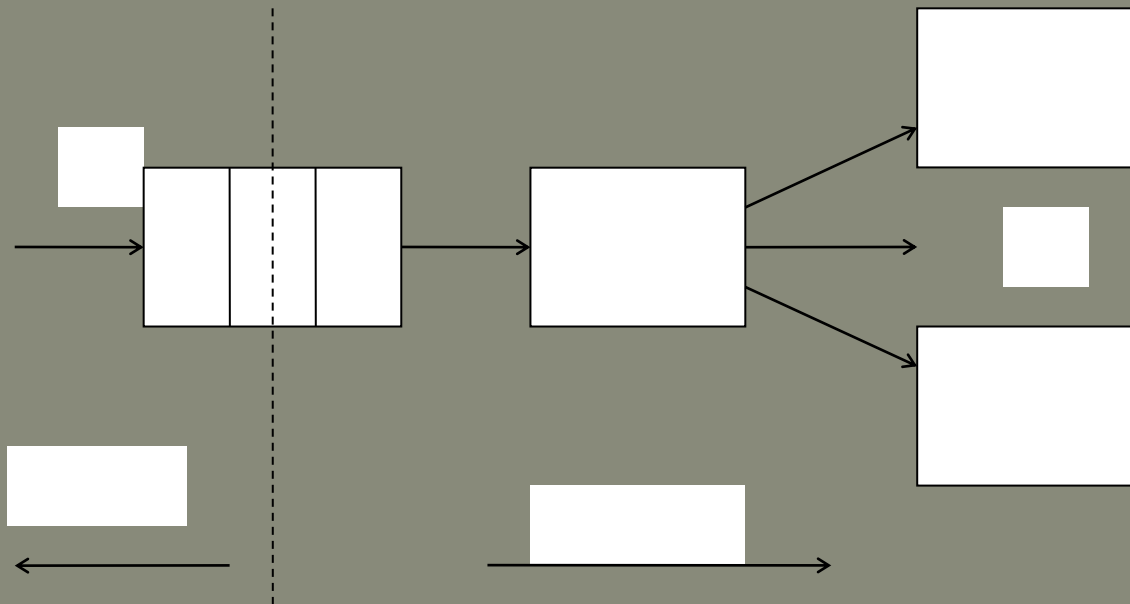


采样模式的工作过程

3. 交互技术

- 事件模式

- 输入设备把数据保存到一个输入队列，也称为事件队列
- 应用程序处理队列中的事件，或删除队列中的事件

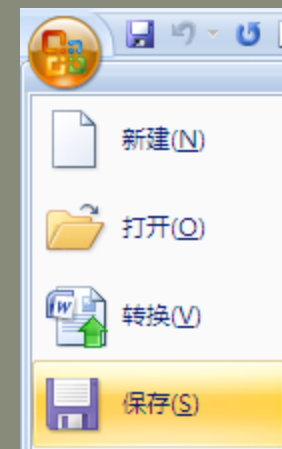


3. 交互技术

- 基本技术(例, 较高层)

3. 交互技术

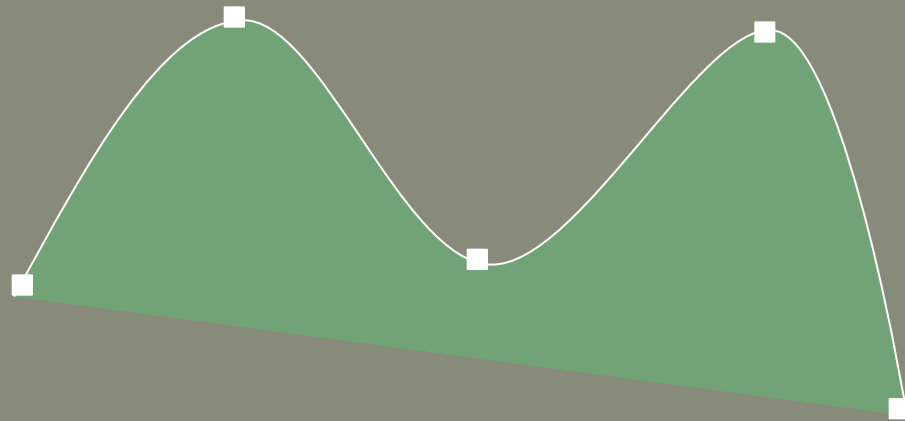
- **定位：**确定点的坐标
 - 直接定位：直接指定坐标值
 - 间接定位：非精确定位方式。鼠标等



3. 交互技术

- 笔划

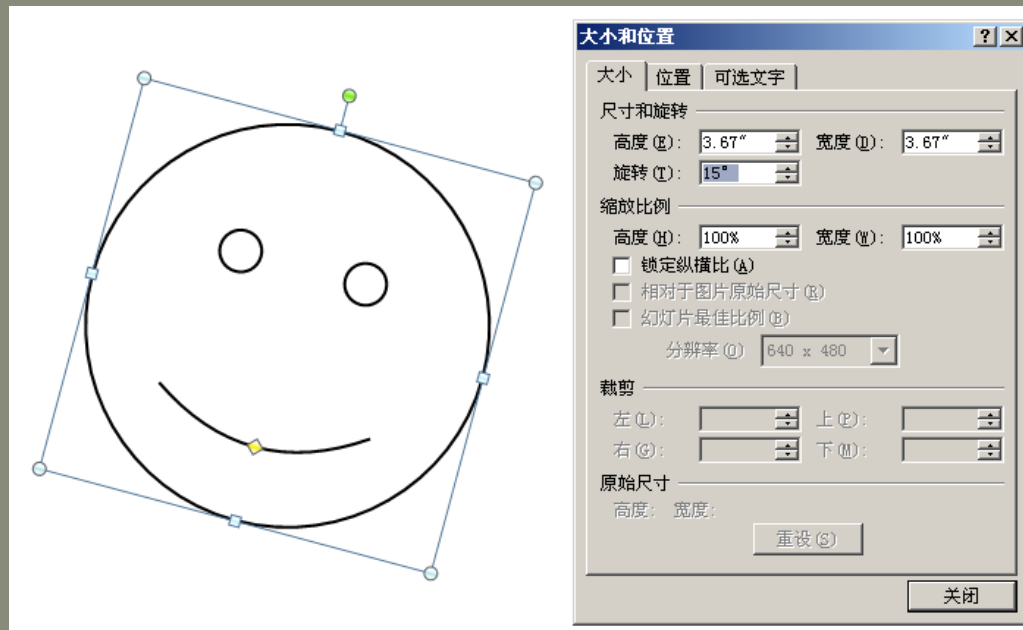
- 一组顺序的坐标点。折线或曲线的控制点。



3. 交互技术

- 定值

- 定值（或数值）输入：旋转角度、缩放比例因子等

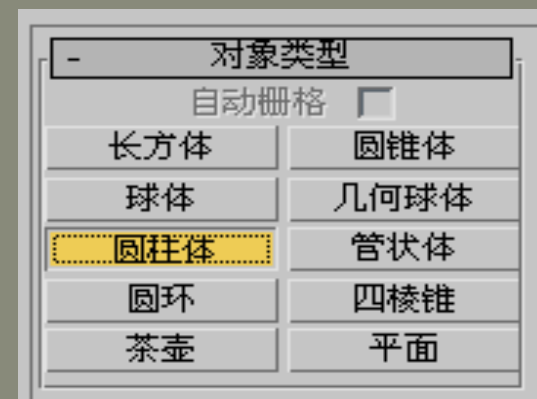
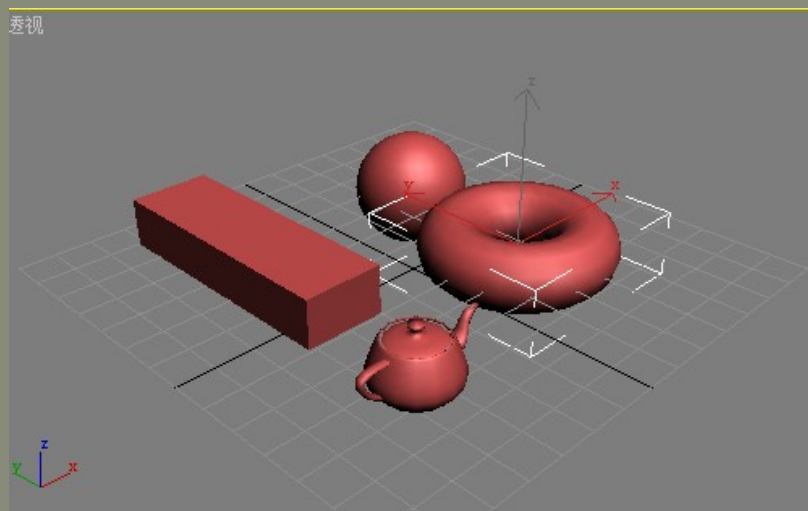


3. 交互技术

- 选择

- 选出一个元素。注视、指点或接触

- 键盘 Ctrl+A
 - 鼠标



3. 交互技术


- 字符串
 - 文本框
 - LoV: List of Value



ID: _____ REGISTER

Password: _____ | FORGET

LOG IN EXIT



查找

查找内容 (F):

▼

☐ 区分大小写 (C)

☐ 全字匹配 (W)

☐ 区分全/半角 (M)

查找下一个 (F)

关闭

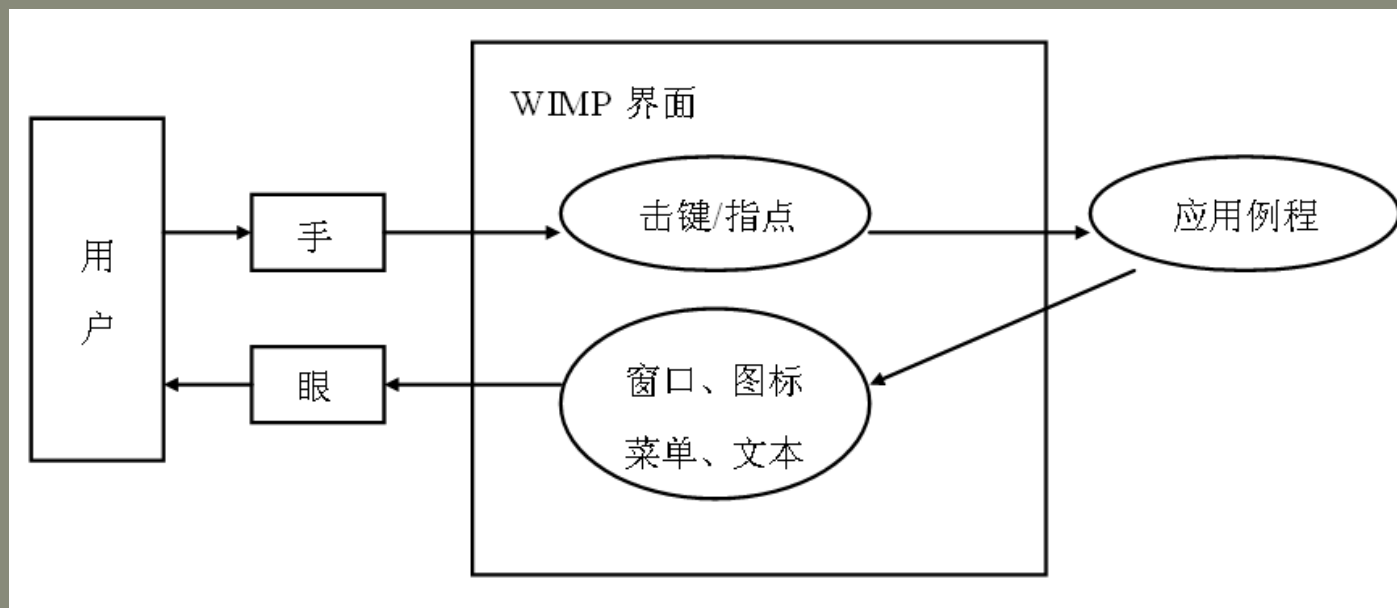
替换 (R)...

3. 交互技术

- 图形与三维交互(例)

3. 交互技术

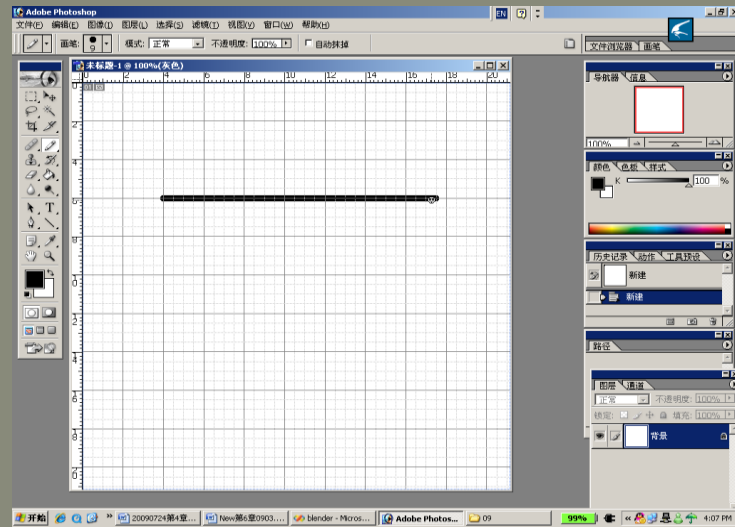
- **WIMP界面：Windows, Icons, Menus, Pointing Device**



3. 交互技术

- 几何约束

- 对图形的方向、对齐方式等进行规定和校准。
- 对定位的约束（网格吸附）

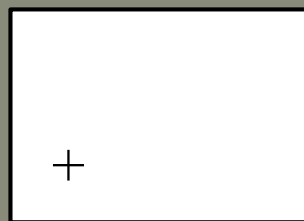


3. 交互技术

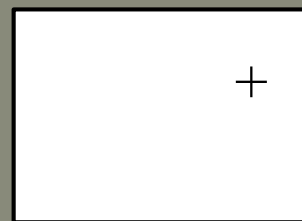
■ 几何约束

— 方向约束

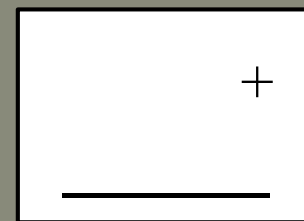
- 例如，绘制印刷线路板、管网图或地籍图时非常有用。



移动
光标



按下按
钮开关



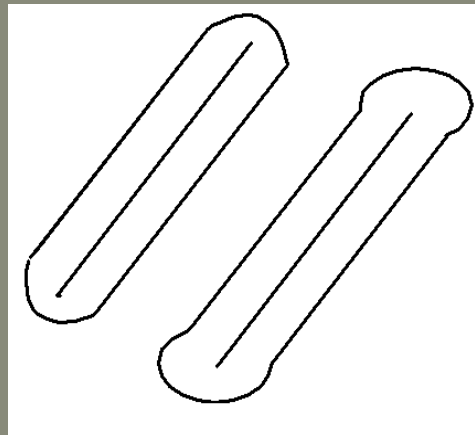
- 例如，在Word绘图中，锁定纵横比，在拖动线段一个端点时，线段只是沿原来方向放缩



3. 交互技术

- 引力场

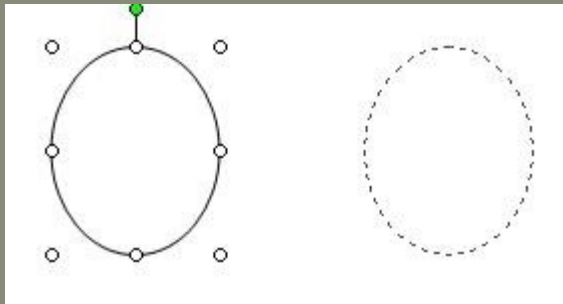
- 光标中心落在区域内时，自动地被直线上最近的一个点所代替
- 引力场的大小要适中



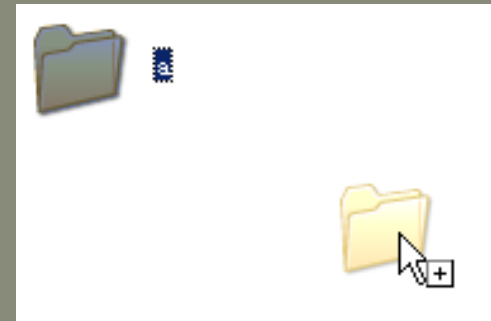
3. 交互技术

- 拖动

一 图形模式



图像模式

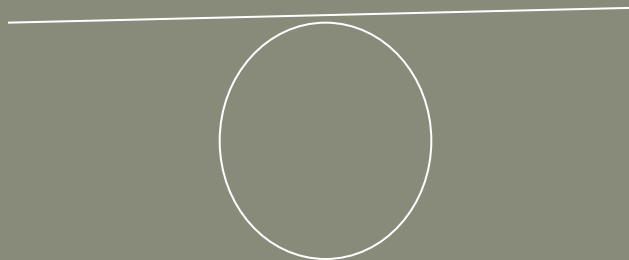


- 图形模式：可在移动的位置上按特定的象素操作模式(如异或方式)进行重新绘制，被拖动的图形不会破坏扫过图形。
- 在图像模式：可进行图像的整体移动。首先按拖动图像大小将屏幕图像保存，当拖动图像离开该位置而移动到下一个新位置时，再恢复该位置上保存的屏幕图像。

3. 交互技术

- 橡皮筋技术

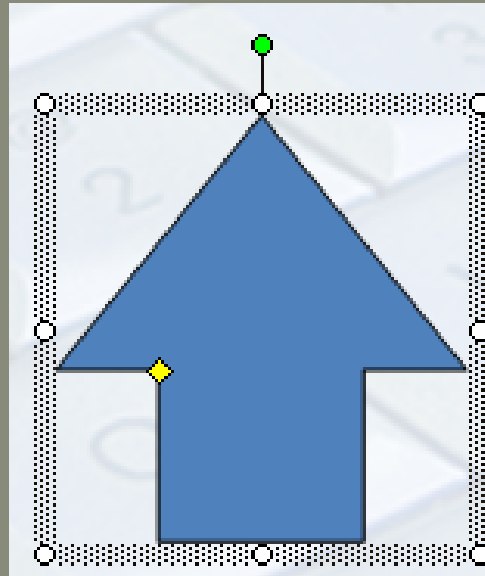
- 不断地进行画图—擦除—画图的过程：可基于双缓冲技术
 - (1) 从起点到光标中心点 (x, y) 处画图；
 - (2) 擦除起点到光标中心点 (x, y) 处的图形；
 - (3) 光标移动到新的位置： $x=x+\Delta x$, $y=y+\Delta y$ ；
 - (4) 转第(1)步，重复这个过程，直到按下确认键为止



3. 交互技术

- 操作柄技术

- 缩放、旋转、剪切等几何变换。图形对象的周围会出现操作柄



3. 交互技术

- 三维交互技术

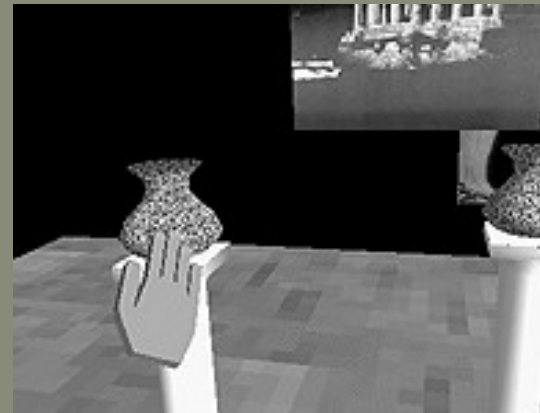
- 面临问题

- 六自由度输入：流行的鼠标、轨迹球、触摸屏等只有两个自由度(沿平面X、Y轴平移)。
 - 窗口、菜单、图符和传统的二维光标在三维交互环境中，难以区分屏幕上光标选择到对象的深度值
 - 三维鼠标

3. 交互技术

- 直接操作

- 三维光标必须有深度感
- 光标在遇到物体时不能进入到或穿过物体内部
- 可以采用半透明三维光标
- 三维光标可以是人手的三维模型



3. 交互技术

- 三维Widgets

- 三维交互界面中的一些小工具

- 漂浮的菜单、拾取物体的手的三维图标、平移和旋转指示器等。

- 1992年美国Brown大学计算机系提出三维Widget设计原则

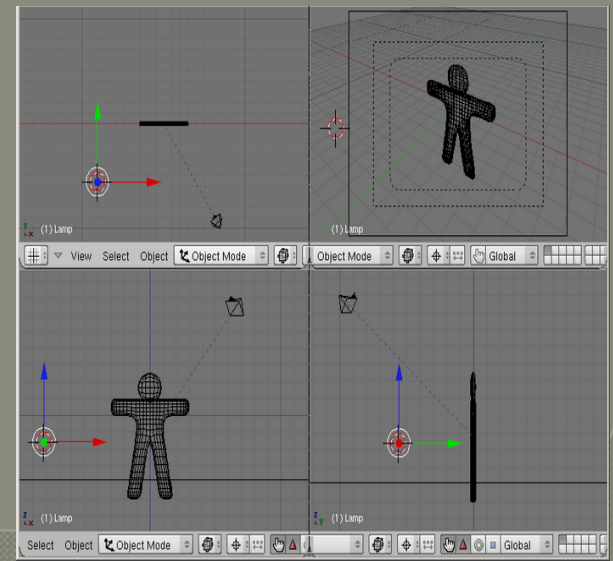
- 几何形状应能表示其用途（扭曲物体的Widget，最好就是扭曲的）
- 适当选择Widget控制的自由度，以简化控制
- 根据用途确定Widget的功能。例如，用于工业设计的用户界面，必须保证精确性

3. 交互技术

- 三视图输入

- 二维输入设备实现三维的输入。

- 三维点：两个视图上的点，唯一确定了三维空间中的一个点
 - 直线段：两端点决定三维空间的一条直线
 - 面：各顶点唯一确定三维空间的一个面
 - 体：各面唯一确定三维空间的一个多面体



3. 交互技术

- 自然交互、AI+

3. 交互技术

1) 语音合成

2) 语音识别

(1) 识别和理解：信号处理、模式识别、概率论和信息论、发声机理和听觉机理、人工智能等。

(2) 主流技术：基于统计, 深度学习

- a) 语音特征提取
- b) 声学模型与模式匹配
- c) 语言模型与语义理解

3. 交互技术

a) 语音特征提取

■ 预处理：

- 预滤波、采样和量化、加窗、端点检测、预加重等

■ 基于语音帧的特征提取：

- 将语音信号分为若干帧，对每一帧提取语音特征

3. 交互技术

b) 声学模型与模式匹配

- 语音到音节概率的计算：
 - 语音特征同声学模型进行匹配
- 较广泛的建模技术：
 - ① 隐马尔科夫模型
 - ② 上下文相关建模
 - ③ 深度神经网络

3. 交互技术

① 隐马尔科夫模型HMM建模

— 两个假设

- 内部状态的转移只与上一状态有关
- 输出值只与当前状态(或当前的状态转移)有关

— HMM建模顺序：

- 常用从左向右单向、带自环、带跨越的拓扑结构
 - 一个音素就是一个三至五状态的HMM
 - 一个词就是构成词的多个音素的HMM串行
 - 连续语音识别的整个模型就是词和静音组合起来的

3. 交互技术

② 上下文相关建模方法

— 建模时考虑协同发音的影响

- 指一个音受前后相邻音的影响而发生变化。
就是人的发声器官在一个音转向另一个音时只能逐渐变化，从而使得后一个音的频谱与其他条件下的频谱产生差异

— 英语的上下文相关建模通常以音素为基元

3. 交互技术

③ 深度神经网络DNN

- 卷积变换
- 双向长短时记忆网络（LSTM, LongShort Term Memory）

3. 交互技术

④ 太极模式识别算法

3. 交互技术

C) 语言模型与语义理解

■ 语言模型计算音节到字的概率。

- 规则模型
- 统计模型
 - 用概率统计的方法来揭示语言单位内在的统计规律：例如，N-Gram模型简单有效，被广泛使用。
 - » N-Gram模型假设：第 n 个词的出现只与前面 $n-1$ 个词相关，而与其它任何词都不相关，整句的概率就是各个词出现概率的乘积。这些概率可以通过直接从语料中统计 n 个词同时出现的次数得到。常用的是二元的Bi-Gram和三元的Tri-Gram。

■ 深度网络+NLP

3. 交互技术

语音特征提取 + 声学模式识别 + 语义理解
综合优化



HCI语音交互通道

3. 交互技术

3) 开发工具实例

Microsoft Speech SDK

- Speech API (SAPI)

- API for Text-to-Speech把文本转化为语音输出

- 金山词霸的单词朗读功能

- » `ISpVoice * pVoice = NULL;`
 - » `HRESULT hr = CoCreateInstance(CLSID_SpVoice, NULL, CLSCTX_ALL, IID_ISpVoice, (void **)&pVoice);`
 - » `pVoice->Speak(L"Hello world", 0, NULL);`

- API for Speech Recognition语音识别应用程序接口,声波转换成文字

其他的SDK或基于AI的语音识别（开源）系统

3. 交互技术

3) 笔式用户界面

实例：

- PIBG范式:P(Physical object)、IB(Icons,Buttons)、G(Gesture) (中国科学院软件所)
- WIMP范式:W(windows)、IM (icons, menus)、P(pointing systems) 相对应。

3. 交互技术

(1) 联机识别

- 一边写，机器一边进行识别

(2) 脱机识别

- 已经写好或印刷好的静态的语言文本图像的识别

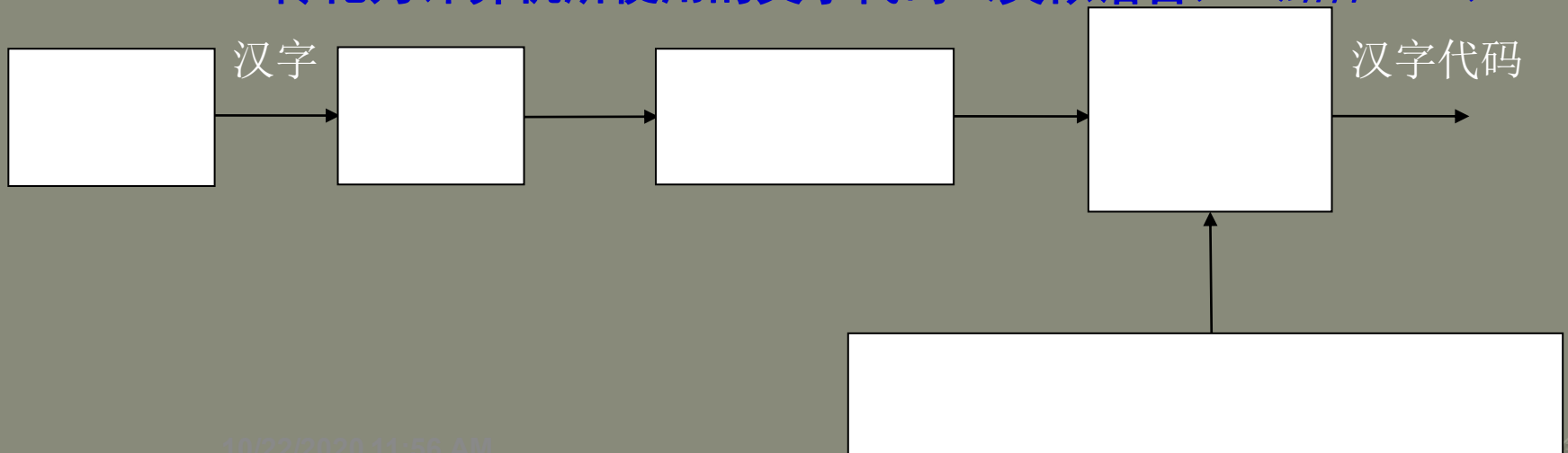
(3) 手写体识别开发技术

- 识别率取决于对手写约束的层次
 - 手写的类型、用户的数量、词汇量的大小以及空间的布局。

3. 交互技术

(1) 联机手写识别

- 识别：
 - 文字图像抬笔、落笔、笔迹上各像素的空间位置，以及各笔段之间的时间关系等信息
- 提取信息特征
- 特征比较（机器学习与模式识别）
- 转化为计算机所使用的文字代码（类似语音）（结合NLP?）



3. 交互技术

(2) 脱机手写识别

- **困难：**

- 点阵图像：要得到笔段的点阵通常需要细化运算。细化会损失一些信息，并且难以得到时间顺序信息（动态参数）。
- 笔画与笔画之间经常粘连：很难拆分，而且笔段经过与另一笔段交叉分成两段后，也难以分清是否应该连起来。

- **结构识别、统计识别以及神经网络方法**

3. 交互技术

(2) 脱机手写识别

- 结构识别方法

- 汉字的组成结构:由笔划(点、横、竖、撇、捺等)、偏旁、部首构成
- 通过把复杂的汉字模式分解为简单的子模式直至基本模式元素, 对子模式的判定以及基于符号运算的匹配算法, 实现对复杂模式的识别。
- 优点是区分相似字的能力强, 缺点是抗干扰能力差。

3. 交互技术

(2) 脱机手写识别

- 统计识别方法

- 将汉字看为一个整体，其所有的特征是从整体上经过大量的统计而得到的，然后按照一定准则所确定的决策函数进行分类判决。
- 统计识别的特点是抗干扰性强，缺点是细分能力较弱。

3. 交互技术

(3) 手写体识别开发技术(例)

数字墨水技术

- 通过三阶贝塞尔曲线来描述笔输入的笔迹，存储方式使得数字墨水文件很小，从而可以更加有效地进行存储。
- 数字墨水的处理包括数字墨水的表示、压缩和显示，**智能的墨水分析技术**，墨水标记和注解技术，墨水的智能操作以及墨水存储和搜索等一系列有关技术。
 - 墨水解析技术可以将笔输入的文字串解析成单字，从而将复杂的语句级手写识别化繁为简，分解成语句解析和单字识别两个部分。
- **微软已实现了数字墨水技术**对英文、德文、法文、韩文、日文、简体和繁体中文等语言的支持。
- **Windows XP Tablet PC Edition**拥有强大而简单的数字化墨水控件和API

3. 交互技术

(深度)神经网络

- 具有学习能力和快速并行实现的特点，因此可以通过神经网络分类器的推广能力准则和特征提取器的有效特征提取准则，对手写字符进行识别。
- tensorflow手写体识别实例：
<https://blog.csdn.net/gentelyang/article/details/77341557>

● 作业

分别列举三种自己最喜欢的交互模式，并从人和机（含软件）两方面说明其交互技术的优缺点（每种技术说明限50字以内）。

再见